



This is the **published version** of the bachelor thesis:

Sahuquillo Planas, Pere; Rivas Contreras, Santiago, dir. De les dades a l'acció : monitoratge de l'aigua. 2023. 13 pag. (Grau en Gestió de Ciutats Intel·ligents i Sostenibles)

This version is available at <https://ddd.uab.cat/record/282893>

under the terms of the license

De les dades a l'acció: Monitoratge de l'aigua

Pere Sahuquillo Planas

Sabadell, Juny de 2023

Resum - El treball s'enfoca principalment en analitzar i avaluar el consum de l'aigua en els edificis d'un municipi fictici amb l'objectiu de proporcionar a l'entitat encarregada una eina que li permeti gestionar el consum d'aigua de manera eficient. Per aconseguir-ho, es proposa monitorar el consum d'aigua en temps real i presentar la informació recopilada en dashboards o panells de control, que permeten visualitzar el consum d'aigua en diferents períodes de temps. D'aquesta manera, es podran identificar patrons de consum, detectar possibles fuites d'aigua i prendre mesures per reduir el consum en cas de ser necessari amb l'objectiu de contribuir a la sostenibilitat ambiental i estalviar diners en la factura de l'aigua.

Paraules clau - Teleconsum de l'aigua - Sensors - Dades - Bases de dades - Eficiència Hídrica - Dashboard - Sostenibilitat - Estalvi d'aigua

Abstract - The project focuses primarily on analyzing and evaluating water consumption in the buildings of a fictional municipality created by myself, with the aim of providing the local government with a tool to manage water consumption efficiently. To achieve this, it is proposed to monitor real-time water consumption and present the collected information on dashboards or control panels, allowing visualization of water consumption over different time periods. This way, consumption patterns can be identified, potential water leaks can be detected, and measures can be taken to reduce consumption if necessary, with the goal of contributing to environmental sustainability and saving money on the water bill.

Keywords - Teleconsumption of water - Sensors - Data - Databases - Water efficiency - Dashboard - Sustainability - Water savings

1 INTRODUCCIÓ

El consum d'aigua potable en els edificis és un dels factors que influeixen en la sostenibilitat del planeta. La gestió eficient de l'aigua es converteix en una necessitat essencial en una societat on cada vegada és més important tenir un consum responsable i conscient dels recursos naturals.

En aquest context, el teleconsum de l'aigua potable es presenta com una solució tecnològica que permet als usuaris gestionar el consum de manera eficient, a través del monitoratge en temps real del consum i la presentació de les dades recopilades en dashboards. En aquest Treball de Fi de Grau, es tractarà el teleconsum de l'aigua potable en els diferents edificis d'una ciutat fictícia amb aproximadament uns 1.500 edificis i una població de 8.000 persones aproximadament, amb l'objectiu d'avaluar la seva

eficàcia com a eina per millorar l'eficiència hídrica en els edificis i contribuir a la sostenibilitat ambiental. En la part pràctica del treball, es generaran les dades d'aquesta ciutat, per a posteriorment poder-hi fer un tractament i ànalisi d'aquestes i un cop tractades, es visualitzaran tots els dashboards amb la informació obtinguda. Un cop construïts tots els dashboards necessaris, ja tindrem l'estructura per a poder crear una pàgina web en un futur, la qual cosa permetrà als usuaris i a l'ajuntament, accedir a la informació en temps real i de forma còmode, des del seu ordinador o dispositiu mòbil. Amb aquesta iniciativa, s'espera fomentar l'ús d'aquesta tecnologia i promoure un consum més sostenible i conscient de l'aigua potable en les llars.

La realització d'aquest treball és una motivació per a mi degut a que es tracta d'una problemàtica actual i només pel fet de poder aportar el meu granet de sorra per a ajudar a la causa, ja em donaria per satisfet.

• E-mail: peresahuquill@gmail.com
• Treball tutoritzat per: Santiago Rivas.
• Curs: 2022/2023

1.1 Objectius

1.1.1 Objectiu general

L'objectiu principal del treball és promoure la sensibilització sobre l'ús responsable de l'aigua entre la població, i així contribuir a la sostenibilitat i a la reducció de l'impacte ambiental a través de la gestió eficient de l'aigua.

1.1.2 Objectius específics

- Realitzar un estudi de la situació actual del consum d'aigua a Catalunya per identificar les necessitats i problemes en relació amb la gestió de l'aigua.
- Esmentar varíes campanyes de conscienciació per a la població sobre l'ús responsable de l'aigua, amb la finalitat de promoure hàbits més sostenibles.
- Col·laborar amb altres institucions o empreses per a la realització d'estudis i la implementació de projectes que ajudin a la millora de la gestió de l'aigua i la seva sostenibilitat.
- Dissenyar i implementar una base de dades per a l'emmagatzemant i ànalisi de les dades obtingudes.
- Desenvolupar dashboards personalitzats que mostrin informació rellevant sobre el consum d'aigua, com ara els patrons de consum, el rendiment dels sistemes de subministrament d'aigua, la detecció de fuites, entre altres indicadors.
- Proporcionar una eina visual i fàcil d'utilitzar que permeti als gestors de l'aigua prendre decisions informades i eficaces per millorar l'eficiència i sostenibilitat de l'ús de l'aigua als municipis.
- Desenvolupar la base de les dades, per a posteriorment poder generar una pàgina web intuitiva i accessible que presenti les dades de consum d'aigua de manera clara i visualment atractiva perquè els usuaris puguin comprendre millor el seu consum i fer ajustos en els seus hàbits d'ús d'aigua per reduir el seu impacte ambiental.

1.2 Metodologia i planificació de treball

El treball constarà de dues parts principals: el marc teòric i el marc pràctic.

En el marc teòric s'exposarà el treball. M'he centrat en la recopilació d'informació a nivell de Catalunya i Barcelona, ja que és on hi ha més informació disponible sobre el tema del consum d'aigua. Això em permetrà obtenir una base d'informació i coneixements sòlida i fiable per a realitzar l'estudi del consum d'aigua a nivell municipal. Aquesta informació no varia massa en quant a la tendència entre Barcelona o un municipi català qualsevol, ja que la tendència de consum d'aigua té patrons semblants [20]. L'única cosa que s'ha de tenir en compte és que el volum de dades i de consum no és el mateix.

Inicialment, tenia previst dur a terme l'avaluació del consum d'aigua als edificis municipals de Terrassa i als habitatges del Prat de Llobregat. No obstant això, després de rebre les dades necessàries per a l'ànalisi, em vaig adonar que no eren òptimes per dur a terme un estudi exhaustiu. Davant aquesta limitació, he decidit adoptar una nova estratègia.

En lloc d'utilitzar dades existents, he decidit generar jo mateix les dades necessàries per a la investigació. Usant el llenguatge de programació Python [2] per simular un municipi fictici format per aproximadament 1.500 edificis, que és una mostra representativa dels diferents tipus d'edificis que es troben en una ciutat típica.

La generació de dades s'ha realitzat seguint un criteri rigorós, on s'hi ha tingut en compte les característiques i els patrons de consum d'aigua esperats en diferents tipus d'edificis, com ara habitatges unifamiliars, apartaments, oficines i establiments comercials. A més, he tingut en compte factors com les hores punta en quant a consum d'aigua, segons l'edifici del qual es tracta.

Un cop generades les dades simulades, he realitzat l'ànalisi del consum d'aigua en aquests 1.500 edificis ficticis. S'han usat diverses tècniques i eines per extreure informació rellevant de les dades, com ara el càcul de consum mitjà, la identificació de patrons de consum i la detecció de possibles ineficiències o fuites d'aigua.

Per molt que no hagi pogut avaluar directament els municipis de Terrassa ni el Prat de Llobregat, aquesta metodologia de generació de dades m'ha permès obtenir resultats significatius i representatius dels possibles escenaris i reptes associats al consum d'aigua en una ciutat. Aquests resultats proporcionen una base sòlida per comprendre els patrons de consum, identificar àrees de millora i proposar mesures eficients per a la gestió de l'aigua en contextos urbans.

El marc pràctic del meu treball implica la recopilació, processament i ànalisi de les dades del consum d'aigua, amb la seva explicació des de l'inici del procés fins al final. Un cop tractades les dades, els dashboards seran generats amb el programari Kibana [18], que permet la visualització i l'ànalisi de dades en temps real de manera eficient i atractiva. La intenció final és incloure aquests dashboards a una pàgina web feta amb el programa WordPress, així serà una informació de fàcil accés i accessible per a tothom. Mai he realitzat un treball d'aquestes magnituds, pel que serà un repte i una motivació poder arribar fins al final i observar el resultat, del qual n'estic segur que serà molt satisfactori.

Pel que fa a la planificació, he seguit una pauta la qual es troba a l'annex del treball en la figura 5, que s'ha establert segons el temps disponible, i el temps que m'ha comportat realitzar cada apartat. A l'abril és quan vaig iniciar el treball, recopilant varíes fonts d'informació prèvies per a començar a dur a terme el treball. Durant el mes de maig he realitzat l'estudi del consum d'aigua i la part teòrica, mentre que a finals de maig i principis de juny he desenvolupat la part pràctica del projecte. Per al dia 26 de juny està marcada la data d'entrega final.

2 MARC TEÒRIC

2.1 Sequera a les ciutats

El tema de la sequera, és avui en dia una de les principals preocupacions en les ciutats de tot el món, ja que el canvi climàtic i el creixement demogràfic estan agreujant el problema. A continuació podem observar una línia temporal del volum d'aigua disponible per a ser potabilitzada a Catalunya amb dades de l'Institut Nacional d'estadística (INE) [21], i com bé podem observar, aquesta corba decreix durant els darrers anys.

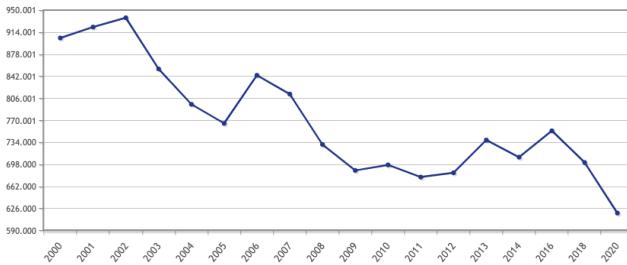


Fig. 1: Volum d'aigua disponible per a potabilitzar anual a Catalunya

La falta d'aigua a les ciutats veiem que és una realitat, i aquesta podria tenir grans conseqüències com podrien ser:

- Escassetat en el subministrament per al consum humà.
- Impacte important en els sectors de l'agricultura i la indústria.
- L'aparició de problemes socioeconòmics i demogràfics.

La sequera no és un problema que només afecta les ciutats, sinó que també afecta directament a la qualitat de vida de la població, i és que estan sorgint noves restriccions com podrien ser talls d'aigua a les dutxes públiques, implementació d'horaris de reg de jardins i la suspensió d'activitats a l'aire lliure per a evitar el consum d'aigua.

Un dels sectors que s'ha vist més afectat per aquestes restriccions esmentades, és el sector de l'agricultura. Aquests, s'han vist obligats a reduir dràsticament la seva producció degut a que els han restringit les hores de reg dels cultius, la qual cosa deriva a unes pèrdues econòmiques significatives, i una producció final més baixa [5].

Quan parlem d'estrés hídric ens referim a la situació en què la demanda d'aigua excedeix a la quantitat disponible i es produeix quan la disponibilitat d'aigua no és suficient per a satisfer les necessitats de la població, l'agricultura i la indústria. L'organització "World Resources Institute" estima que de cara al 2040, Espanya estarà entre els 33 països amb major estrès hídric [11], per darrere de països com Qatar. Sense anar més lluny, segons l'Organització de les Nacions Unides (ONU), de cara a l'any 2025, s'estima que dos terços de la població viuran en zones d'estrés hídric, i moltes altres ciutats com és el cas de Galende, municipi de Castilla y Leon ja hi han posat fil a l'agulla i ja s'han anunciat sancions econòmiques a aquells habitants que facin un mal ús de l'aigua. [23]

2.2 Consum d'aigua

L'aigua és un recurs abundant en la Terra, però lamentablement per al consum humà només en va destinat un 1% d'aquesta. Segons dades de l'OMS [24], només un 57% de la població mundial tenia accés a aigua potable canalitzada. Aquest és un factor directament relacionat amb la sequera, a poc a poc veiem que els recursos d'aigua s'estan esgotant, per la qual cosa el repte dels governs és fer que la població estigui el més conscienciada possible del problema en què ens trobem. Gràcies a l'OpenData de Barcelona [12] he pogut obtenir dades de partir de les quals he creat una línia temporal per a veure i analitzar la tendència del consum d'aigua des del 2012 fins al 2020.

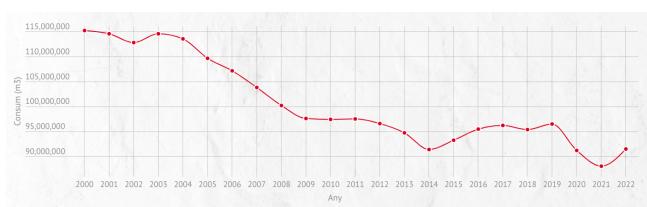


Fig. 2: Consum d'aigua anual a Barcelona

S'han obtingut dades de Barcelona perquè és la capital de Catalunya, i a partir de la qual podem extrapolar les dades als altres municipis. Com podem observar, es pot veure una davallada notòria en quant al consum anual durant els darrers anys. L'any 2000, hi havia un consum aproximat de 115.000 hm³, mentre que el 2020, que és l'últim registre del qual s'han assolit les dades, ens vam trobar amb un consum d'aproximadament 90.000 hm³, la qual cosa confirma l'estat d'alerta en el que ens trobem davant el consum d'aigua. L'aigua es reparteix per a 3 diferents usos:

- Consum domèstic: La major part d'aigua consumida és per a fer-ne un ús domèstic i satisfer les necessitats de casa.
- Consum comercial i industrial: Una part va destinada a les indústries per a la producció.
- Consum municipal: Una petita part d'aquesta aigua, va destinada a serveis de l'ajuntament, ja sigui per a neteja viària, regar les zones verdes de les ciutats o bé per als edificis municipals.

A continuació podrem observar un petit gràfic de la distribució de l'aigua, amb dades extretes de l'Ajuntament de Barcelona [13].

Estan sorgint noves restriccions les quals exigeixen que l'aigua potable només es reservi per beure i no per altres usos municipals. Actualment, podem diferenciar entre dos usos diferents de l'aigua:

- Usos consumptius: L'ús consumptiu és aquell que es destina a les necessitats domèstiques, industrials, de reg i per a l'agricultura. Les fonts d'aigua són recursos hídrics convencionals com les aigües superficials, rius, llacs i de les aigües subterrànies.

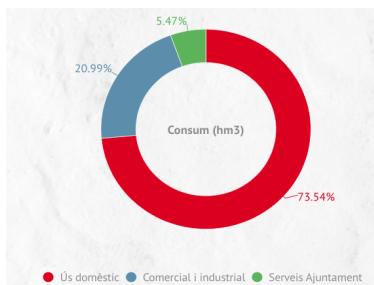


Fig. 3: Úsos de l'aigua a Barcelona

- Usos no consumptius: Els usos no consumptius, suposa la utilització de l'aigua, però no el seu consum i no implica extreure l'aigua del medi on es troba. El seu ús va destinat per a usos energètics i usos lúdics.

2.3 D'on prové l'aigua de Barcelona?

Els recursos hídrics de Barcelona i la seva Àrea Metropolitana, principalment provenen dels rius Ter i Llobregat. Aquesta, un cop passen pel procés de potabilització, es destina a la xarxa, per a fer-ne un ús a les llars. Secundàriament, també hi trobem les aigües subterrànies i aigües pluvials que, aquesta, és la que es destina principalment a l'ús municipal, encara que el seu volum és menor al mencionat anteriorment. Quan parlem d'aigües subterrànies, parlem dels aquífers del Delta de Llobregat i el Delta del Besòs. Pel que fa a les aigües pluvials, s'utilitzen per a la neteja de carrers, reg per a zones verdes entre molts altres usos. Es distribueixen més de 700 m³ anuals d'aigua freàtica. Aquestes aigües són gestionades per l'Ajuntament de Barcelona per a tots els usos municipals i l'aigua regenerada que gestiona l'Àrea Metropolitana de Barcelona [14].

L'aigua de la pluja, es recull i s'emmagatzema en 30 dipòsits d'aigua repartits per la ciutat per a poder disposar-la quan sigui necessari. Com ja he comentat anteriorment, actualment a Catalunya estem patint una situació important de sequera, i per això és important tenir en compte la situació dels rius i embassaments d'aigua, per a poder fer una prediccio de si aquests ens podran proporcionar aigua en un futur no molt llunyà i poder administrar l'aigua de la millor manera possible. Segons l'Agència Catalana de l'Aigua, a dia 16 d'abril de 2023, a Catalunya es disposa només del 26% de les reserves d'aigua. [15]

2.4 Teleconsum d'aigua potable

2.4.1 Definició de teleconsum d'aigua potable

El teleconsum consisteix a orientar a la població sobre el seu comportament diari amb relació al consum d'aigua, i poder saber a temps real el consum actual en les llars. És un tema d'interès perquè té repercussió en la salut humana, desenvolupament socioeconòmic i la sostenibilitat ambiental, però a part de l'interès personal de la població, hi ha moltes empreses, governs i organitzacions que busquen comprendre millor els patrons de consum d'aigua i desenvolupar estratègies per millorar la gestió de recursos hídrics.

2.4.2 Modelització del teleconsum

Per a dur a terme la modelització del teleconsum d'aigua, s'ha de complir i realitzar tot un seguit de passos que veurem a continuació:

1. **Recopilació de dades:** El primer pas per modelitzar el teleconsum d'aigua potable en els edificis, és recopilar les dades sobre el consum d'aigua. Aquestes dades poden venir de comptadors d'aigua o de tecnologies de telemetria que permeten la recopilació de dades de forma remota.
2. **Anàlisi exploratòria de les dades:** Una vegada recopilades les dades, es realitza una ànalisi exploratòria per comprendre millor les característiques del conjunt de dades. Això inclou l'ànalisi de la distribució del consum d'aigua i la identificació de patrons o tendències en les dades.
3. **Selecció de les variables:** El següent pas és seleccionar les variables que es volen utilitzar per modelitzar el teleconsum. Això pot incloure variables com la superfície de la llar, el nombre de persones que hi freqüenten, el nivell socioeconòmic i altres factors que puguin influir en el consum d'aigua.
4. **Selecció de models:** Després de seleccionar les variables, es pot procedir a la selecció de models per a modelar el teleconsum d'aigua potable. Això pot incloure gràfics, models de regressió lineal o no lineal i altres models que s'adequin a les característiques de les dades.
5. **Avaluació del model:** Un cop seleccionat un model, es realitza una àvaluació per mesurar la seva precisió i eficàcia en la predicció del consum d'aigua.
6. **Utilització del model:** Finalment, el model es pot utilitzar per predir el consum d'aigua en una àrea determinada. Això pot ajudar les empreses o organitzacions que gestionen els serveis d'aigua a optimitzar la distribució d'aigua i millorar la gestió dels recursos hídrics.

2.5 Integració de l'IoT i de la IA?

La tecnologia de l'Internet of Things és una eina important per a la gestió de l'aigua que ajuda a les empreses del sector i als governs a ser més eficients i sostenibles en el seu ús de l'aigua. La idea d'Internet of Water busca connectar totes les parts implicades en el subministrament d'aigua (fonts d'aigua, plantes de tractament i sistemes de gestió de l'aigua industrial, instal·lacions de distribució, empreses de serveis públics, etc.) per dotar d'informació útil a les persones responsables de la presa de decisions i facilitar en part la seva feina [8]. IoT és una tecnologia clau per optimitzar la cadena de subministrament d'aigua a través de processos i aplicacions de dades, que ens proporcionen varíes facilitats com:

- Transparència en tots els processos de la cadena de subministrament d'aigua. Gràcies a l'IoT, tenim dades de tota la cadena de subministrament d'aigua.

- Resposta immediata davant de possibles problemes com podria ser la contaminació de l'aigua i reduir al màxim aquests danys.
- Automatització, per a millorar l'eficiència, reduir els costos i millorar la qualitat en el servei.
- Cost optimitzat, al tenir una automatització, això permet reduir costos a llarg termini.
- La sostenibilitat. És el punt més important i en el que s'enfoca el treball. Gràcies al IoT, no només millorem l'eficiència i reduïm costos, sinó que també ajuda a complir els objectius ambientals com la petjada de carboni, la contaminació i el tema a tractar, la preservació de l'aigua.

Gràcies a aquesta integració de l'Internet de les coses al monitoratge de l'aigua, es poden analitzar, processar i tractar aquestes dades que generen els diferents sensors [17]. És aquí, on entra la nova tecnologia de la qual últimament tothom n'està parlant, la intel·ligència artificial. El potencial que té de processar dades crues del sistema, analitza tot el que hi succeeix dins d'aquest, mentre que gràcies a l'aprenentatge automàtic li permet estar millorant contínuament la seva comprensió de les dades i les respostes que ha de donar envers un problema. Amb la intel·ligència artificial es poden crear infraestructures d'aigua, supervisades per una tecnologia que ho veu tot, no es cansa i s'adapta a qualsevol situació que se li posi per davant. Gràcies a aquesta tecnologia, és possible analitzar fluxos d'aigua en temps real, enviant alertes i apagant els sistemes corresponents en cas de detectar-hi qualsevol fuita o anomalia. Amb això, es redueix a gran escala el malbaratament d'aigua, ja que la IA tarda molt poc temps a detectar qualsevol incongruència en el sistema, comparat amb el que tardaria qualsevol enginyer o enginyera a trobar el problema [19].

A partir d'aquestes dades que s'estan generant gràcies al IoT, de tot el procés de subministrament d'aigua, es poden generar models predictius i estratègies sobre com optimitzar l'ús de l'aigua en el futur i poder tendir cap a un futur molt més sostenible i amigable amb el medi ambient.

2.5.1 Tecnologies necessàries per al monitoratge d'aigua

Actualment, al mercat podem trobar diferents dispositius IoT que permeten dur a terme el mesurament i monitoratge del consum d'aigua. Per a començar hi tenim els sensors. Aquests s'encarreguen de mesurar la qualitat de l'aigua, la seva composició química un cop tractada, etc. També poden mesurar el grau de plenitud d'un dipòsit d'aigua, o la pressió en les canonades de distribució. Aquests sensors proveeixen dades, que serviran per a fer l'anàlisi de com està funcionant el subministrament. Al mercat hi podem trobar varis que s'adaptin a les tasques que volem realitzar, com per exemple:

- Sensor capacitiu M.D. Micro Detectors C18P/BP-1A [4]
- Transmissor de nivell ultrasònic SUP-MP [6]

- IoT Multi-Parameter Water Quality Analyzer [7]

El segon dispositiu que és necessari, són els comptadors intel·ligents. Aquests permeten mesurar el consum d'aigua en temps real, ajudar a corregir el consum excessiu i segons els patrons, podem fer algoritmes predictius per al consum futur. Actualment, podem trobar diferents opcions de comptadors intel·ligents al mercat:

- Sensus iPERL [1]
- flowIQ® 2200 [3]

Finalment, gràcies a algoritmes de machine learning, molts sistemes de distribució són automàtics i permeten regular i controlar el subministrament d'aigua, de manera que són el més eficients possibles. Tots aquests dispositius són un exemple de que aquest cas d'ús es pot realitzar actualment, i és que ja hi ha dispositius que ho permeten.

2.6 Incentius i programes governamentals

Tot i que la tecnologia per al monitoratge de l'aigua és una eina essencial per a la gestió sostenible i eficient de l'aigua, no n'hi ha prou. És important també involucrar als ciutadans i incentivar-los a adoptar comportaments més responsables davant el consum d'aigua, ja que, al cap i a la fi, són ells els encarregats de poder reduir aquest consum, i és important que en siguin conscients de la situació en què ens trobem. Barcelona ha implementat programes de conscienciació sobre el consum d'aigua, així com incentius per als ciutadans que adoptin pràctiques més sostenibles.

3 MARC PRÀCTIC

3.1 Explicació

L'objectiu del marc pràctic del Treball Final de Grau és generar la infraestructura de les dades per a que posteriorment aquestes puguin ser representades i visualitzades en una pàgina web. Aquesta pàgina queda fora de l'abast d'aquest TFG. Aquest treball és la base per a que en un futur, amb dades reals es pugui crear. Aquesta, estarà orientada a l'Ajuntament de la ciutat, amb l'objectiu final de proporcionar als responsables de la presa de decisions de l'Ajuntament una eina eficaç i accessible per planificar i gestionar els recursos hídrics de la ciutat de manera més efectiva.

Per a l'obtenció de les dades de consum d'aigua de tots els habitatges de la ciutat, he dut a terme l'estudi sobre els consums dels diferents edificis de la ciutat, realitzant una anàlisi exhaustiva tenint en compte diversos factors importants com poden ser, els horaris laborals, els hàbits de consum específics de cada tipus d'edifici i altres variables rellevants que s'han de tenir en compte a l'hora de generar les dades.

Utilitzant el llenguatge de programació Python, he modelat i ajustat uns valors aleatoris perquè es correlacionin amb els consums reals que s'espera observar en els edificis de la ciutat fictícia. He tractat de fer-ho el més real possible, i per això les dades obtingudes finalment estan estructurades com si vinguessin des del mateix comptador d'aigua, és a dir són dades crues que s'hauran de tractar i processar.

Per a dur a terme aquest processament, usaré Jupyter Notebook, en aquest tractament aplicaré una ontologia per a compartir la comprensió comuna entre persones o sistemes automatitzats i ajudarà a establir una comprensió comuna de les dades de consum d'aigua, definir els conceptes rellevants, les relacions entre ells, i crear una estructura de dades coherent que permeti la fàcil manipulació i anàlisi de les dades.

Un cop processades les dades, he creat diversos dashboards usant el programa Kibana. Aquests dashboards són la peça clau de la futura pàgina web, ja que mostraran de manera visual i interactiva la informació aconseguida. Els dashboards inclouen gràfics i taules dinàmiques que permeten veure el comportament del consum d'aigua en diferents contextos (per exemple, per zones geogràfiques, períodes de temps, etc.).

3.2 Ciutat escollida

Com he esmentat anteriorment m'he decidit per a generar les dades d'una ciutat fictícia. D'aquesta manera, puc tenir les dades que a mi m'interessen per a la realització del projecte. La ciutat que he generat, serà una ciutat situada a la comarca del Baix Camp, i serà una ciutat semblant a Reus, respecte a la situació geogràfica, dimensions i població. Com ja he comentat anteriorment, consistirà en una ciutat d'uns 1.500 edificis aproximadament, els quals disposaran de:

- 1 Ajuntament
- 2 Hospitals
- 2 Comissaries de Mossos d'esquadra
- 4 Escoles
- 339 Oficines
- 1207 Habitatges
- 19 Indústries

En total la ciutat disposarà de 1.574 edificis.

3.3 Obtenció, neteja i anàlisi de dades

3.3.1 Obtenció de les dades

Després de parlar amb diversos ajuntaments i que no em poguessin proporcionar dades interessants per a l'execució del meu projecte, vaig decidir introduir-me personalment a estudiar els consums de cada tipus d'edifici tenint en compte varíes variables com podrien ser; caps de setmana, patrons de consum a l'horari laborable, etc. És a dir que, estudiant el consum a cada tipus d'edifici, he generat totes les dades amb les quals he treballat durant el treball.

- Consum d'aigua als habitatges: Consultant varis informes i estudis[16], [22], he assignat un consum mitjà de 100 litres d'aigua per persona. Amb una mitjana de 4 persones per habitatge.

Les dades s'han generat seguint aquestes referències de consum.

- Consum d'aigua a les escoles: A les escoles, s'ha estimat un consum d'aigua d'uns 15 litres per estudiant, pel que si aproximem que hi ha uns 500 alumnes en tota l'escola, això ens genera valors de consum aproximats als 7.500 litres d'aigua al dia [9]. Distribuïts durant les hores lectives que van de 08:00 - 18:00 h aproximadament.
- Consum d'aigua a l'ajuntament: A l'ajuntament s'ha aproximat un valor d'uns 350 litres d'aigua a dia, distribuïts entre els horaris laborals, que van de 09:00 - 14:00 h i de 16:00 - 20:00 h.
- Consum d'aigua a l'hospital: Als hospitals com bé ja sabem, és un dels llocs on s'hi consumeix més aigua. És un edifici on constantment s'hi estan fent neteges, cures i tractaments als pacients, pel que això s'ha vist reflectit en els valors de consum d'aigua. Aquests valors poden variar entre els 150 - 1.325 litres / pacient al dia [26]. Tenint en compte que parlem d'hospitals mitjans i amb un total de 50 persones aproximadament, hem establert un consum d'uns 15.000 litres d'aigua al dia, distribuïts durant totes les 24 hores del dia.
- Consum d'aigua a les oficines i comissaries de policia: He evaluat els dos conjunts d'edificis per igual, ja que considero que els valors mitjans de consum són molt aproximats, obviament, tenint en compte la grandària de les oficines i tenint present que pot variar el nombre de persones. He assignat un consum d'uns 50 litres / dia aproximadament [25].
- Consum d'aigua a les indústries: A les indústries, tenim un dels casos més difícils d'avaluar el consum, i és que depenen del tipus d'indústria podrà tenir un consum molt elevat o més baix. Gràcies a contactes que tinc que treballen en indústries de diversos sectors diferents, he pogut aproximar els valors de consum.

3.3.2 Neteja

Un cop ja hem obtingut les dades i les hem processat per a disposar-les de la manera que més ens interessa, és el moment de fer la neteja, i posteriorment fer-ne una anàlisi d'aquestes.

Primer farem la neteja per a posteriorment poder analitzar-les, ja que si fem primer l'anàlisi i després la neteja, estarem analitzant possibles valors NaN (Valors nuls) o possibles valors duplicats, la qual cosa pot alterar significativament el nostre estudi.

La neteja és important per a eliminar dades inadequades o irrelevants, tractar els valors nuls i valors duplicats, detectar i corregir errors i anomalies del conjunt de dades i finalment ens ajuda a identificar aquells valors que no són redundants i no ens interessa tenir al nostre dataframe final.

El primer que farem és, veure si s'ha generat algun valor que sigui nul. Això pot ser degut a que els comptadors que enregistren les dades, hagin patit algun problema i no s'hagi enregistrat cap valor. Per a solucionar aquest problema i que afecti el menys possible al nostre conjunt de dades, hi ha varíes maneres de solucionar-ho, n'explicaré les més destacades:

- Es poden esborrar els registres, és a dir que esborrem tota la fila sencera, ja que si no tenim el valor de consum, aquell registre ja no ens interessa, i hi deixem un espai entre els registres adjacents, la qual cosa no és la millor opció, perquè si posteriorment vull crear línies temporals, això suposaria que hi haguessin espais en buit.
- Fer-ne ús de l'estadística. Podem fer ús de l'estadística descriptiva com la mitjana, la mediana i la moda per a substituir els valors nuls.
- Finalment, la interpolació lineal és un procés que ens podria ser de gran utilitat. Aquest tracta en estimar els valors faltants, a partir de les dades ja existents. Per a substituir els valors faltants del meu conjunt de dades, aplicaré el mètode 'ffill', que el què fa és copiar l'últim valor vàlid conegit i s'utilitza per omplir els valors faltants fins que es trobi un altre valor vàlid.

El següent pas i últim, serà comprovar que no hi hagi valors duplicats. Un cop ens assegurem que no hi tenim valors faltants ni duplicats, ja podem començar amb l'anàlisi més profunda de les dades.

3.3.3 Anàlisi

Una bona anàlisi profunda de les dades és una part essencial per a la realització del treball. Aquest procés permet entre d'altres coses conèixer i entendre les dades i identificar anomalies o valors atípics.

Un cop coneixem el conjunt de dades, s'ha d'aprofundir, identificar i avaluar les columnes en les que s'hi poden treure la major informació rellevant i que ens interessi per al nostre estudi.

El nostre conjunt de dades disposa de 7 tipus d'edificis, on a cadascun d'ells se li farà una evaluació del consum d'aigua individual, ja que seria poc significatiu realitzar l'estudi en general, perquè cada tipus d'edifici té uns patrons i un valor de consum totalment diferents de la resta.

Seguidament evaluarem la distribució dels valors de consum obtinguts:

Com bé podem observar a la figura 12, disposem de 7 diagrames de caixa, cadascun dels quals representa els valors de consum per a cada subcategoria dels edificis.

Primerament, per a l'ajuntament veiem que el 50% de les dades estan bastant dispersades, però entren dins d'una normalitat on no hi ha valors outliers.

Seguidament, a les escoles podem trobar una cosa semblant a les dades de l'ajuntament en quant a la dispersió. Tenim el 50% de les dades per valors molt pròxims a 0, on l'escola està tancada per tant no hi ha consum d'aigua. Per altra banda, hi trobem també una alta dispersió en quant als valors de consum, incloent-hi varis outliers que poden ser deguts a una possible fuga d'aigua o alguna font d'aigua oberta per error.

Per a les comissaries de policia, podem observar que el 50% inicial de les dades es troben aproximades al 0, mentre que l'altre 50% de les dades estan bastant dispersades, però entren dins d'una normalitat on no hi ha valors outliers semblant a les dades de l'ajuntament.

Per als hospitals disposem dels valors de consum més alts, on observem que el 50% de les dades es troba entre el 0 - 2.000 litres de consum d'aigua, mentre que l'altre 50% restant, trobem que té una major dispersió, amb valors de fins a més de 7.000 litres d'aigua.

Si continuem amb el diagrama de caixa dels habitatges, trobem que el primer 50% de consum es troba entre els 0 - 30 litres de consum i no hi ha una gran dispersió respecte a l'altre 45% dels valors de consum. Mentre que trobem un 5% de les dades que són outliers i tenen uns valors molt poc comuns i més alterats del normal.

Per a les oficines, el 50% inicial de les dades es troben aproximades als 0 litres de consum, mentre que l'altre 50% de les dades estan bastant dispersades, i hi trobem certs valors els quals ens fan més difícil la lectura del gràfic ja que al ser valors outliers i estar tan dispersats complica poder realitzar correctament la lectura dels valors més comuns.

Finalment, tenim el diagrama de caixa de les indústries on trobem el primer 50% de les dades entre valors de 0 - 20 litres aproximadament, mentre que l'altre 50% de les dades tenen una gran dispersió, arribant fins al consum de 350 litres d'aigua.

L'anàlisi, ens ha permès identificar les tendències de consum per a cada tipus d'edifici i avaluar el comportament de la població en aquests. En algun cas inclòs, hem pogut identificar consums excessius en certs edificis, els quals s'han traduït en alertes a la visualització de les dades tal com veurem a continuació. Gràcies a aquest procés d'anàlisi, hem pogut validar i confirmar que aquest és un projecte que es pot escalar a un cas real, i és capaç de solucionar problemàtiques reals dels municipis.

3.4 Disseny Dashboards

Un cop realitzats tots els passos previs, és hora de visualitzar les dades de manera que l'ajuntament pugui treure les conclusions necessàries. La visualització de les dades, es farà mitjançant dashboards interactius, que permetran una presentació molt visual de les dades, facilitarà l'anàlisi i la comprensió de la informació i es podrà adaptar a les necessitats de l'ajuntament. A l'utilitzar gràfiques, els dashboards permetran comprendre i analitzar les dades d'una manera entenedora i intuïtiva.

He utilitzat el programari Kibana, que és una eina molt útil per a aquest tipus de visualització, però sí que és cert que m'he trobat amb certes limitacions, sobretot amb la quantitat de colors disponibles per a generar els gràfics. Per a crear els dashboards, directament es poden estirar les dades a través d'una API des de la Base de Dades per a tenir dades a temps real, o en el cas del meu treball, penjar un arxiu en format csv amb el conjunt de dades que es vol visualitzar, i a partir d'aquestes treballar les dades.

El meu dashboard s'ha dividit en 2 parts principals. En la primera part, hi podrem observar uns dashboards generals avaluant el consum d'aigua a tota la ciutat. Hi trobarem 3 KPI's:

- La suma total del consum de tots els edificis.
- El nombre total d'edificis de la ciutat.

- EL nombre total d'alarms que hi ha degut a l'excés de consum d'aigua, de manera que es pugui veure de manera intuitiva si tenim alguna incidència o no.

A part dels indicadors, disposem d'uns gràfics complementaris que descriuran les nostres dades.

- Distribució de consum per dia de la setmana: És un treemap que ens ordena de major a menor el consum d'aigua segons els dies de la setmana.
- Distribució de consum per hores del dia: És un treemap que ens descriu les hores del dia en que hi ha major consum. Ens ho ordena de major a menor consum.
- Consum total per edificis: És un treemap que ens mostra el consum total de cada sector.
- Distribució de consums dels edificis per dia: És semblant a l'anterior, però en aquest cas tracta d'un gràfic de barres el qual ens mostra el consum total dels tipus d'edificis, per a cada dia del mes.
- Evolució del consum d'aigua: Ens mostra una sèrie temporal, amb l'evolució del consum amb una freqüència d'hora per hora.

En la segona part, hi tenim cada tipus d'edifici avaluat per individual. Primerament com en la primera part, hi trobem uns KPI's:

- Consum total per tipus d'edifici.
- Nombre total d'edificis.
- Consum mitjà al dia.

A part dels indicadors, també disposem de gràfics, els quals la informació que ens proporcionen és:

- Consum total per dia de la setmana: És un treemap que ens ordena de major a menor el consum d'aigua segons els dies de la setmana.
- Distribució de consum per hores del dia: És un treemap que ens descriu les hores del dia en que hi ha major consum. Ens ho ordena de major a menor consum.
- Evolució del consum mitjà: Ens mostra una sèrie temporal, amb l'evolució del consum mitjà amb una freqüència de dues hores. En cas d'haver-hi un comportament anormal, automàticament es mostra una alarma per a poder avaluar què hi està passant.

3.5 Avaluació dels resultats

Com a part final del treball, toca avaluar els resultats. Personalment, visualitzar i avaluar els resultats finals de tot el treball, és la part que més m'agrada, ja que per fi veus els resultats i com ha quedat finalment tot el procés que has estat duent a terme. La visualització dels dashboards m'ha permès veure, analitzar i avaluar el comportament de tota la població en general referent al consum d'aigua, de manera que amb una sola ullada la persona encarregada, pot saber

si la ciutat està tenint un comportament bo o dolent. Com he comentat anteriorment, s'ha generat un dashboard més general de tota la ciutat, i després un dashboard individual per a cada tipus d'habitatge, que és on ens permetrà veure si hi ha hagut fuites d'aigua i les seves corresponents alertes en cada tipus d'edifici.

A continuació procediré a comentar el dashboard general. Els dashboards individuals de cada tipus d'edifici, els podem trobar a l'annex, pel que és important que un cop haguem observat el general, els hi fem una ullada a les figures 13, 14, 15, 16, 17, 18, 19.

Com bé podem observar en la figura 4, hi tenim uns KPI'S que ens mostren el consum d'aigua de tota la ciutat durant el mes de març, el nombre total d'alarms que han saltat durant aquest període de temps i el nombre total d'edificis de la ciutat. Per a saber de la magnitud d'aigua de la qual estem parlant, aniria bé poder agafar una petita referència. De mitjana en una piscina olímpica hi caben uns 2.500.000 litres d'aigua aproximadament [10], per tant el consum de la ciutat podria omplir en total unes 9,5 piscines olímpiques. Gràcies a les alarmes, s'hi ha pogut detectar en total 13 casos en els quals, el consum d'aigua ha sigut desproporcionat i fora de la normalitat, ja sigui per a una possible fuita d'aigua, o que algú s'hagi deixat una aixeta oberta. Si anem als dashboards de l'annex, podem visualitzar d'una manera més visual i senzilla de quin dia i hora s'ha detectat aquest consum excessiu, i a quin edifici li pertoca. Un cop parlat dels KPI'S, veiem que a la part dreta del dashboard, hi podem trobar 3 treemaps els quals cadascun d'ells ens proporcionen informació de com s'ha distribuït aquest consum.

En el primer d'ells, hi podem veure com s'ha distribuït el consum segons els dies de la setmana, els resultats són els següents. Dimecres (16,73%), dijous (16,72%), divendres (16,61%), dimarts (13,37%), dilluns (13,33%), diumenge (11,67%), dissabte (11,59%). Aquesta distribució ens mostra el percentatge que s'ha consumit cada dia de la setmana, i podem concloure que els dies que menys consum d'aigua hi ha són els caps de setmana, la qual cosa això pot coincidir en el fet que part de la població pot tenir segones residències fora de la ciutat, per la qual cosa el consum els caps de setmana és menor. El gràfic inferior és molt semblant a l'anterior, i aquest ens mostra les 5 hores amb major consum durant el dia. Primerament d'aquest segon gràfic hi podem destacar que les hores on s'hi mostra el major consum, és a les 07:00 h i a les 06:00 h del matí. Aquestes hores coincideixen amb l'hora que s'acostuma a llevar la gent i els nens per anar a treballar o anar al col·legi, i suposo que serà degut a la quantitat d'aigua que es gasta a la dutxa. Les següents hores amb major consum són de 16:00 h - 18:00 h de la tarda, que és quan la majoria de la població torna de treballar. Per tant, puc concloure que hi ha una correlació bastant alta entre l'horari laborable de la població i el consum d'aigua.

Finalment, l'últim gràfic a comentar és el que està situat a la part inferior dreta. Aquest ens mostra la distribució del consum d'aigua segons el tipus d'habitatge. Trobem un tipus d'edifici que hi predomina, que són els habitatges. En total els 1.207 habitatges consumeixen un 77,57% del total d'aigua, d'aquí prové l'alta correlació entre el consum d'aigua i l'horari laborable. El segon edifici amb

major consum són els 4 hospitals que suposen un 11,97%, seguidament de les 339 oficines amb un consum total del 4,89%, seguit de les 19 indústries amb un 3,45% del total. En els edificis restants, el consum total és molt baix comparat amb els ja comentats.



Fig. 4: Dashboard general

4 CONCLUSIONS

A través de l'anàlisi i l'avaluació del consum d'aigua en els edificis d'un municipi fictici, s'ha dut a terme una investigació amb l'objectiu de proporcionar a l'ajuntament una eina eficient per gestionar aquest recurs vital.

L'objectiu principal del treball era monitorar el consum d'aigua en temps real i presentar la informació recopilada en dashboards o panells de control, amb la intenció de visualitzar el consum d'aigua, identificar patrons de consum, detectar possibles fuites i prendre mesures per reduir el consum si fos necessari. Aquest enfocament tenia com a finalitat contribuir a la sostenibilitat ambiental.

Inicialment, la idea era buscar una ciutat on poder realitzar el meu estudi per a poder tractar un cas real. Vaig contactar amb diversos ajuntaments, però a causa de la confidencialitat de les dades, cap d'ells va poder proporcionar-me dades interessants per a dur a terme el treball. Això em va portar a pensar en alguna alternativa per a poder seguir mantenint la idea inicial d'avaluar el consum d'aigua en alguna ciutat, i vaig arribar a la conclusió que, aplicant els coneixements de programació obtinguts en el grau, es podia fer una simulació d'aquestes dades, de manera que això és el que vaig fer. Per tant, al no trobar dades reals, no puc considerar-ho com un fracàs, ja que, al cap i a la fi vaig aconseguir dur-ho a terme. Un cop obtingudes les dades, a través de les dades visualitzades en dashboards, s'han identificat diferents patrons de consum en els diferents edificis. Aquesta informació ha estat vital per comprendre millor els hàbits i comportaments dels residents pel que fa a l'ús de l'aigua i s'han pogut identificar també, les hores punta de consum.

En segon lloc, la implementació d'aquest sistema de monitoratge ha permès detectar fuites d'aigua en alguns edificis. Aquesta capacitat de supervisar de manera contínua el consum d'aigua pot resultar molt beneficiosa per a l'ajuntament en qüestió, ja que permet identificar a temps aquestes fuites d'aigua i actuar de manera instantània sobre elles. Això contribueix a la conservació de l'aigua i evita el malbaratament d'aigua en un moment de sequera tan important com l'actual.

En tercer lloc, els resultats d'aquest treball confirmen que gràcies a la gestió de les dades, és possible proporcionar un

sistema de control per avaluar la situació en què es troba la ciutat referent al tema.

Finalment, per concloure el treball, vull esmentar que es pot entendre aquest treball com un equilibri entre el zoom-out (la part més teòrica del projecte) i el zoom-in (la part pràctica) del concepte. A més, en aquesta darrera secció també vull comentar diferents direccions sobre les quals es podrà seguir desenvolupant el treball futur d'aquest tema. Aquestes direccions són:

- Distingir les diferents categories de consum d'aigua: domèstic, agricultura o comercial.
- Tenir en compte les fuites d'aigua i com controlar-les
- Els orígens de l'aigua

5 AGRAÏMENTS

Principalment, agrair a la meva família, en especial als meus pares i al meu germà, per a recolzar-me incondicionalment i per a la motivació que m'han aportat en els moments en els quals em donava per vençut. Per confiar en mi, i per valorar els meus resultats acadèmics. A tots els meus amics, especialment a la família que m'ha donat aquesta curta etapa, ja que ens hem animat i motivat mútuament per a seguir amb la realització del treball. Al tutor Santiago Rivas, per a tot el suport, coneixement i ganys que m'ha proporcionat per a seguir endavant amb el projecte. I finalment vull estendre el meu agraïment a l'empresa Nexus Geographics, especialment a l'equip de Recitty, per totes les facilitats i suport proporcionats per a dur a terme aquest treball. La seva col·laboració ha estat essencial per a l'obtenció de les dades i les eines necessàries per a la investigació.

REFERENCES

- [1] Contador de agua de estado sólido — contadores de agua iperltm (international). <https://sensus.com/emea/es/products/iperl-international-water-meters/>.
- [2] El tutorial de python. <https://docs.python.org/es/3/tutorial/>.
- [3] flowiq® 2200 — contador de agua para áreas residenciales — kamstrup. <https://www.kamstrup.com/es-es/soluciones-de-medicion-de-agua/contadores-de-agua-inteligentes/meters/flowiq-2200>.
- [4] Sensor capacitivo m.d. micro detectors c18p/bp-1a. automation24. https://www.automation24.es/sensor-capacitivo-m-d-micro-detectors-c18p-bp-1a?previewPriceListId=1&gclid=CjwKCAjwp6CkBhBEiwAlQVyxRhXQ-ge71vTJ3q1E6yHCUox0vVOqbBZMNDBD5-KQHqkF69XcikhzRoCdOoQAvD_BwE.

- [5] Si continuen les restriccions de aigua mancaran els productes de proximitat. <https://jarc.cat/si-continuen-les-restriccions-daigua-mancaran-aliments-de-proximitat/>.
- [6] Sup-mp ultrasonic level transmitter. <https://www.supmeaauto.com/ultrasonic-level-meter/mp-a-ultrasonic-level-meter>.
- [7] Wholesale iot multiparameter water quality monitor solution for wastewater ground water manufacturer and supplier. <https://www.boquinstruments.com/mpg-6099-multi-parameter-analyzer-product/>.
- [8] Internet of water coalition. <https://internetofwater.org/>, 2019.
- [9] Escuelas públicas gastan agua en exceso). <https://www.elsoldelbajio.com.mx/local/exceden-escuelas-consumo-de-agua-alumno-tecnologico-universidad-conagua-4963032.html>, 2020.
- [10] Cantidad de agua en una piscina olímpica. <https://piscinazo.top/piscina-olimpica/cantidad-de-agua-en-una-piscina-olimpica/>, 2022.
- [11] España, el 28º país más amenazado por escasez de agua del planeta, según wri. <https://www.femeninorural.com/espana-el-28o-pais-mas-amenazado-por-escasez-de-agua-del-planeta-segun-wri/>, 2023.
- [12] D'AIGÜES DE BARCELONA. SA, S. G. Evolución mensual del consumo de agua. 1995 - 2023. https://ajuntament.barcelona.cat/estadistica/angles/Estadistiques_per_temes/Economia/Consum_comerc_i_preus/Consum/coev04.htm, 2023.
- [13] DE BARCELONA, A. El consum d'aigua a barcelona. chrome-extension: //efaidnbmnnibpcajpcgkclefindmkaj/ <https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/122956/3/Consum%20Aigua%202020.BCNROC.pdf>, 2020.
- [14] DE BARCELONA, A. M. Ciclo y recursos hídricos. [https://www.amb.cat/s/es/web/ecologia/aigua/cicle-aigua/cicle-i-recursos-hidrics.html/](https://www.amb.cat/s/es/web/ecologia/aigua/cicle-aigua/cicle-i-recursos-hidrics.html).
- [15] DE L'AIGUA, A. C. Estado de las reservas de agua en los embalses. <https://aca.gencat.cat/es/laigua/estat-del-medi-hidric/recursos-disponibles/estat-de-les-reserves-daigua-als-embassaments/>, 2023.
- [16] DE ÁREAS NATURALES PROTEGIDAS, C. N. ¿sabes cuánta agua consumes? <https://www.gob.mx/conanp/articulos/sabes-cuanta-agua-consumes>, 2019.
- [17] DIGITEUM. Ventajas clave de los sistemas de gestión del agua iot (internet de las cosas). <https://www.digiteum.com/smart-water-management-iot/>, 2021.
- [18] ELASTICSEARCH. Kibana. <https://www.elastic.co/es/kibana/>.
- [19] FUTURE ENERGY SUMMIT, W. The power of data: how artificial intelligence is transforming water). <https://www.worldfutureenergysummit.com/en-gb/future-insights-blog/the-power-of-data-how-artificial-intelligence-is-transforming-water.html>, 2023.
- [20] GENCAT. Consum d'aigua per comarques a catalunya. <https://aca.gencat.cat/ca/laigua/consulta-de-dades/dades-obertes/visualitzacio-interactiva-dades/Consum-aigua-comarques-catalunya/index.html>, 2023.
- [21] INE. Volumen de agua disponible (potabilizada y no potabilizada) por comunidades y ciudades autónomas, tipo de indicador y periodo. <https://www.ine.es/jaxi/Datos.htm?txp=53446#!tabs-grafico>.
- [22] MUÑOZ, R. P. Una persona necesita 100 litros de agua al día: Oms. <http://www.gaceta.udg.mx/una-persona-necesita-100-litros-de-agua-al-dia-oms/>, 2022.
- [23] PASCUAL, M. Estrés hídrico, el reto de la escasez de agua de las ciudades. <https://www.newtral.es/estres-hidrico-escasez-agua/20220808/>, 2022.
- [24] PLANETA, M. En què gastem l'aigua els catalans. <https://monplaneta.cat/reportatges/en-que-gastem-laigua-els-catalans-88/>, 2023.
- [25] PLAZA, A. El consumo de agua en la jornada laboral mejora el rendimiento y evita la deshidratación). <https://alicanteplaza.es/Lameningitisfrenadayolvidadaenelltimoaogracias/2021>.
- [26] WE, J. Hospitales y clínicas: grandes consumidores de agua). <https://justwe.org/hospitales-y-clinicas-grandes-consumidores-de-agua/>, 2020.

A ANNEX

A.1 Diagrama de Gantt



Fig. 5: Diagrama de Gantt

A.2 Codi generació de les dades

A continuació deixaré adjunt, el document del codi que he realitzat per a crear de les dades. [Generació de les Dades.tex](#)

A.3 Neteja i anàlisi de les Dades

A.3.1 Codi neteja i anàlisi de les dades

A continuació deixaré adjunt, el document del codi que he realitzat per a dur a terme l'anàlisi i la neteja de les dades. [Neteja i Anàlisi de les Dades.tex](#)

A.3.2 Neteja

Un cop ja hem obtingut les dades i les hem processat per a disposar-les de la manera que més ens interessa, és el moment de fer la neteja, i posteriorment fer-ne una anàlisi d'aquestes. Primer farem la neteja per a posteriorment poder analitzar-les, ja que si fem primer l'anàlisi i després la neteja, estarem analitzant possibles valors NaN (Valors faltants) o possibles valors duplicats, la qual cosa pot alterar significativament el nostre estudi. La neteja és important per a eliminar dades inadequades o irrelevants, tractar els valors faltants i valors duplicats, detectar i corregir errors i anomalies del conjunt de dades i finalment ens ajuda a identificar aquells valors que no són redundants i no ens interessa tenir al nostre dataframe final.

El primer que farem és, veure si s'ha generat algun valor que sigui nul. Per a identificar-los, aplicarem la següent línia de codi:

Com podem observar a la figura 6, a la columna del Consum (l), hi trobem que hi ha 1.885 valors, que no s'han enregistrat correctament. Això pot ser degut a que els comptadors que enregistren les dades, hagin patit algun problema i no s'hagi enregistrat cap valor. Per a solucionar aquest problema i que afecti el menys possible al nostre conjunt de dades, hi ha varies maneres de solucionar-ho, n'explicaré les més destacables.

- Es poden esborrar els registres, és a dir que esborrem tota la fila sencera ja que si no tenim el valor de

```
print("En el dataset hi ha ", df_final.shape[0], " registres amb ",
      df_final.shape[1], " variables.\nDe les quals hi ha missings:\n")
print(df_final.isnull().sum(), "\n")

Num_polica      0
Date_observed   0
Categoria       0
Subcategoria    0
Consum (l)      1885
Dia_setmana     0
Mes             0
dtype: int64
```

Fig. 6: Identificació de valor nuls

consum, aquell registre ja no ens interessa, i hi deixem un espai entre els registres adjacents, la qual cosa no és la millor opció, ja que si posteriorment vull crear línies temporals, això suposaria que hi haguessin espais en buit.

- Fer-ne ús de l'estadística. Podem fer ús de l'estadística descriptiva com la mitjana, la mediana i la moda per a substituir els valors faltants.
- Finalment, la interpolació lineal és un procés que ens podria ser de gran utilitat. Aquest tracta en estimar els valors faltants, a partir de les dades ja existents. Per a substituir els valors faltants del meu conjunt de dades, aplicaré el mètode 'ffill', que el què fa és copiar l'últim valor vàlid conegut i s'utilitza per omplir els valors faltants fins que es trobi un altre valor vàlid.

```
df_final.interpolate(method='ffill', inplace=True)

print("En el dataset hi ha ", df_final.shape[0], " registres amb ",
      df_final.shape[1], " variables.\nDe les quals hi ha missings:\n")
print(df_final.isnull().sum(), "\n")

Num_polica      0
Date_observed   0
Categoria       0
Subcategoria    0
Consum (l)      0
Dia_setmana     0
Mes             0
dtype: int64
```

Fig. 7: Substitució dels valors nuls

Com podem observar a la següent figura 7, hem substituït els valors NaN, i quan fem la comprovació ens surt que ja no n'hi ha cap dins el conjunt de dades.

El següent pas i últim, serà comprovar que no hi hagi valors duplicats amb la següent línia de codi 8.

```
has_duplicates = df_final.duplicated().any()
print(has_duplicates)

False
```

Fig. 8: Identificació de valors duplicats

Un cop ens assegurem de que no hi tenim valors faltants ni duplicats, ja podem començar amb l'anàlisi més profund de les dades.

A.3.3 Anàlisi

Una bona anàlisi profunda de les dades és una part essencial per a la realització del treball. Aquest procés permet entre d'altres coes conèixer i entendre les dades i identificar anomalies o valors atípics. Per a començar amb l'anàlisi, començarem imprimint per pantalla una descripció del conjunt de dades. Això ens servirà per posar-nos en context i saber el volum de dades amb el que treballem.

	Num_poliça	Date_observed	Categoría	Subcategoria	Consum (l)	Dia_setmana	Mes
count	1171056	1171056	1171056	1171056	1.169171e+06	1171056	1171056
unique	1574	NaN	2	7	NaN	7	1
top	0001	NaN	Edifici Privat	Vivenda	NaN	Dimecres	Març
freq	744	NaN	1164360	898008	NaN	188880	1171056
mean	NaN	2023-03-16 11:30:00	NaN	NaN	2.030204e+01	NaN	NaN
min	NaN	2023-03-01 00:00:00	NaN	NaN	-1.280385e+03	NaN	NaN
25%	NaN	2023-03-08 17:45:00	NaN	NaN	2.830134e+00	NaN	NaN
50%	NaN	2023-03-16 11:30:00	NaN	NaN	1.522641e+01	NaN	NaN
75%	NaN	2023-03-24 05:15:00	NaN	NaN	2.409191e+01	NaN	NaN
max	NaN	2023-03-31 23:00:00	NaN	NaN	7.549186e+03	NaN	NaN
std	NaN	NaN	NaN	NaN	8.939636e+01	NaN	NaN

Fig. 9: Descripció del conjunt de dades

Com bé podem observar a la figura 9, disposem d'un dataset, on hi tenim 1.171.056 registres i 7 variables que són:

- **Num_pòlissa:** És el número identificador del contracte que ens relaciona l'edifici amb el contracte de la companyia d'aigües.
- **Date_observed:** Valor de la data en format datetime YYYY-mm-dd HH:mm:ss. Observem que tenim registres de cada hora, des del 01-03-2023 a les 00:00:00, fins al 31-03-2023 a les 23:00:00.
- **Categoría:** Ens defineix si l'edifici és públic o privat.
- **Subcategoria:** Ens diu de quin tipus d'edifici es tracta.
- **Consum (l):** És el valor de consum que registra al final de cada hora. És a dir, a les 01:00:00, ens dirà el consum d'aigua que s'ha enregistrat des de les 00:00:00 fins les 01:00:00. El valor que ens proporciona és en litres / hora.
- **Dia_setmana:** És el dia de la setmana en que s'ha enregistrat.
- **Mes:** El mes en que s'ha enregistrat el valor. Serà sempre el mes de març.

En quant al tipus de variables, veiem que:

- Variables de text: Num_pòlissa, Categoría, Subcategoria, Dia_setmana i Mes.
- Variable tipus data: Date_observed.
- Variable tipus flotant: Consum (l)

Un cop evaluat el conjunt de dades per sobre, és moment d'aprofundir més, i avaluar la columna Subcategoria, ja que és una columna prou descriptiva d'on podem treure informació rellevant que ens pot ser interessant.

Com podem observar, aquesta disposa de 7 tipus d'edificis, on de cadascun d'ells se li farà una valuació del consum

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1171056 entries, 0 to 1171055
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Num_poliça  1171056 non-null  int64  
 1   Date_observed 1171056 non-null  datetime64[ns]
 2   Categoría    1171056 non-null  object  
 3   Subcategoria 1171056 non-null  object  
 4   Consum (l)   1169171 non-null  float64 
 5   Dia_setmana  1171056 non-null  object  
 6   Mes          1171056 non-null  object  
dtypes: datetime64[ns](1), float64(1), int64(1), object(4)
memory usage: 62.5+ MB
```

Fig. 10: Descripció 2 del conjunt de dades

Tipus de edificis: ['Ajuntament' 'Hospital' 'Policia' 'Centre Educatiu' 'Oficines' 'Vivenda' 'Industria']

Fig. 11: Tipus d'edificis del dataframe

d'aigua individual, ja que seria poc significatiu realitzar l'estudi en general, perquè cada tipus d'edifici té uns patrons i un valor de consum totalment diferents de la resta. Seguidament avaluarem la distribució dels valors de consum obtinguts:

Com bé podem observar a la figura 12, disposem de 7 diagrames de caixa, cadascun dels quals representa els valors de consum per a cada subcategoria dels edificis.

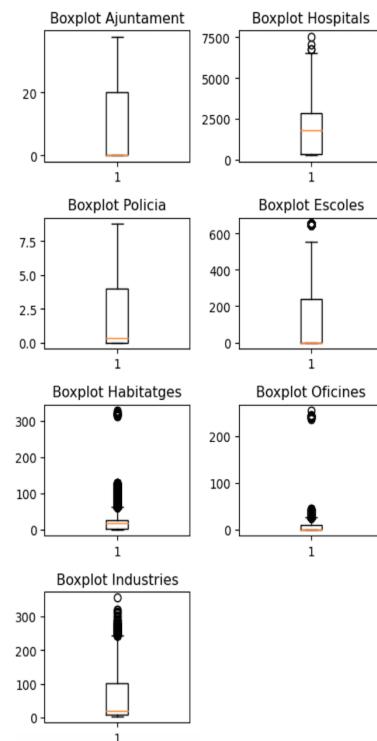


Fig. 12: Boxplots dels edificis

En primer lloc, per a l'ajuntament veiem que el 50% de les dades estan bastant dispersades, però entrem dins d'una normalitat on no hi ha valors outliers.

Seguidament, a les escoles podem trobar alguna cosa semblant a les dades de l'ajuntament en quant a la dispersió. Tenim el 50% de les dades per valors molt propers a 0, on l'escola està tancada per tant no hi ha consum d'aigua. Per altra banda hi trobem també una alta dispersió en quant als

valors de consum, incloent-hi varis outliers que poden ser deguts a una possible fuga d'aigua o alguna font d'aigua oberta per error.

Per a les comissaries de policia, podem observar que el 50% inicial de les dades es troben aproximades al 0, mentre que l'altre 50% de les dades estan bastant dispersades, però entre dins d'una normalitat on no hi ha valors outliers semblant a les dades de l'ajuntament.

Per als hospitals disposem dels valors de consum més alts, on observem que el 50% de les dades es troba entre el 0 - 2.000 litres de consum d'aigua, mentre que l'altre 50% restant, trobem que té una major dispersió, amb valors de fins a més de 7.000 litres d'aigua.

Si seguim amb el diagrama de caixa dels habitatges, trobem que el primer 50% de consum es troba entre els 0 - 30 litres de consum i no hi ha una gran dispersió respecte l'altre 45% dels valors de consum. Mentre que trobem un 5% de les dades que són outliers i tenen uns valors molt poc comuns i més alterats del normal.

Per a les oficines, el 50% inicial de les dades es troben aproximades als 0 litres de consum, mentre que l'altre 50% de les dades estan bastant dispersades, i hi trobem certs valors els quals ens fan més difícil la lectura del gràfic ja que al ser valors outliers i estar tan dispersats complica poder realitzar correctament la lectura dels valors més comuns.

Finalment, tenim el diagrama de caixa de les indústries on trobem el primer 50% de les dades entre valors de 0 - 20 litres aproximadament, mentre que l'altre 50% de les dades tenen una gran dispersió, arribant fins al consum de 350 litres d'aigua.

A.4 Dashboards individuals

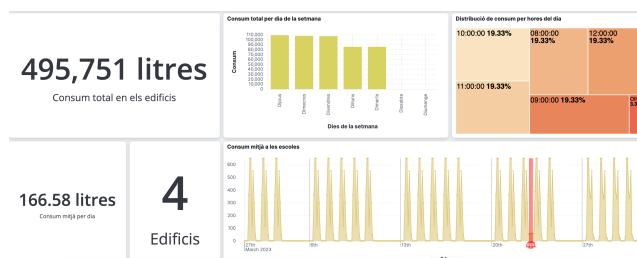


Fig. 13: Dashboard d'escoles

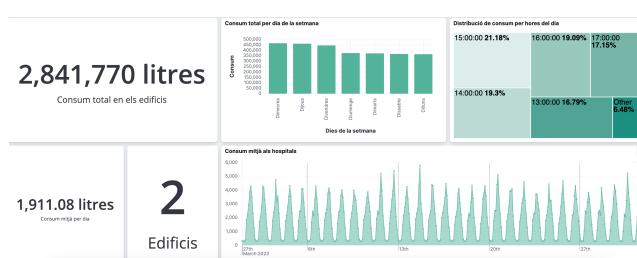


Fig. 14: Dashboard d'hospitals

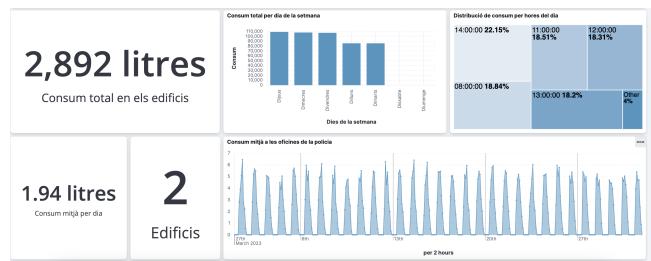


Fig. 15: Dashboard policia

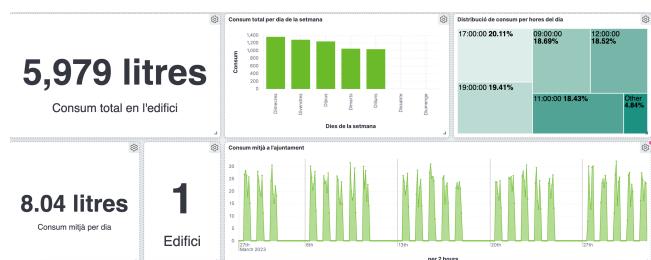


Fig. 16: Dashboard ajuntament

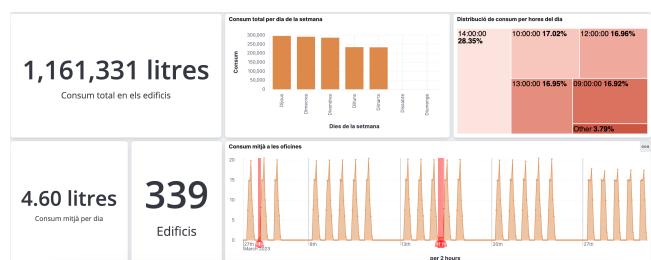


Fig. 17: Dashboard d'oficines

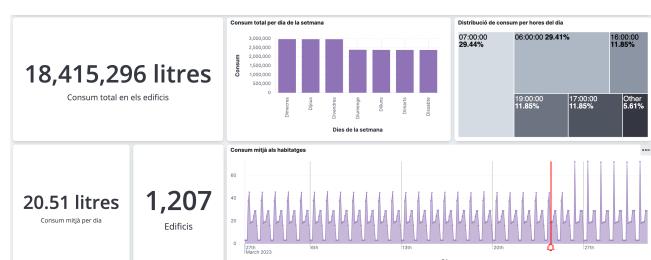


Fig. 18: Dashboard habitatges

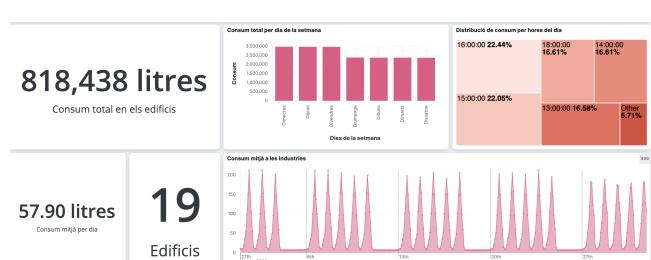


Fig. 19: Dashboard d'indústries