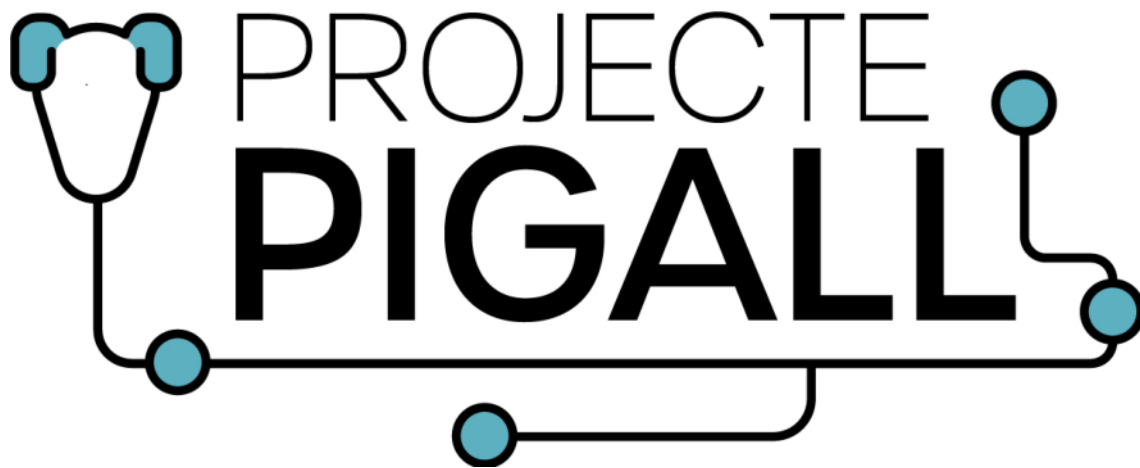


INFORME *SPONSOR*



Es coneix com a Pigall la persona o animal que guia i acompanya a qui necessita ajuda per desplaçar-se. És una assistència que et condueix cap a on vols anar, i té la capacitat de discernir eventuais perills degut a obstacles, cosa que requereix de gossos d'intel·ligència elevada i entrenament avançat.

Pigall és també, com s'anomena aquest projecte pilot d'implementació de IA a l'Àrea de Serveis Socials de l'Ajuntament de Girona, ja que es considera que la funció de la IA ha de ser conduir-nos allà on decidim, facilitant la visió d'allò que tenim ocult, ajudant-nos a superar obstacles i fent més fàcil i planer el camí per arribar-hi.

ÍNDEX

1. EL PROJECTE PIGALL	4
1.1 PRINCIPIS DEL PROJECTE	4
2. L'EINA PIGALL	6
3. METODOLOGIA PIGALL	8
3.1 DETECCIÓ DE NECESSITATS I RECERCA D'EXPERIÈNCIES.....	10
3.1.1 Recerca d'experiències i sistemes d'informació	10
3.1.2 Entrevistes i trobades amb professionals de l'àrea i de l'Ajuntament	10
3.1.3 El contingut de l'eina Pigall	12
3.2 EXPLORACIÓ DE LES DADES.....	15
3.2.1 Coneixement de les dades internes disponibles i comparabilitat amb altres sistemes	15
3.2.2 Anàlisi de les dades internes disponibles i la seva qualitat: SIAGI-Serveis Socials....	17
3.2.3 Procediments de millora de la qualitat de la dada interna	18
3.2.4 Els Ajuts d'Urgència Social com a variable seleccionada per crear una metodologia per elaborar models predictius	20
3.2.5 Dades externes	22
3.3 CONSTRUCCIÓ DE MODELS.....	23
3.3.1 Clústering de perceptors d'ajuts d'urgència social (AUS)	23
3.3.2 Canvi de tendència al Servei Porta d'Entrada	25
3.3.3 Predictor de risc de pèrdua d'habitatge	26
3.3.4 Risc de vulnerabilitat en persones grans.....	26
3.3.5 Evolució del pressupost.....	26
3.3.6 Resum de models i correspondència amb les funcionalitats de l'eina Pigall.....	27
3.4 PRODUCCIÓ DE L'EINA	27
3.4.1 Quadre de Comandament.....	29
3.4.2 Quadre d'Avisos	38
3.4.3 Escenaris de Futur	42
3.4.4 Eines de Consulta	45
3.5 TESTEIG I FORMACIÓ.....	47
3.5.1 Testeig	47
3.5.2 Formació.....	48
4. REPLICABILITAT DEL PROJECTE	49
5. CANVIS PRODUI TS DURANT EL PROJECTE.....	50
6. ELEMENTS SATISFACTORIS, MILLORABLES I RECOMANACIONS PER AL DESPLEGAMENT.....	51

7. IMPACTES	53
7. 1 IMPACTES INTERNS	53
7.2 IMPACTES EXTERNS I DIVULGACIÓ	55
ANNEXOS.....	57
ANNEX 1	57
ANNEX 2	62
ANNEX 3	78
ANNEX 4	110
ANNEX 5	115

PREÀMBUL

L'adopció de la Intel·ligència Artificial ha d'alliberar al personal de Serveis Socials de les tasques mecàniques i permetre'ls dedicar més temps a treballar els vincles, les relacions, el treball socioeducatiu i comunitari.

Amb aquesta premissa i a partir diferents necessitats i oportunitats detectades a l'Àrea de Serveis Socials de l'Ajuntament de Girona neix el projecte Pigall.

Com a necessitat, a l'hora de planificar i dissenyar estratègies d'intervenció, es detecta la necessitat de conèixer de forma més precisa la realitat i les dinàmiques socials de la ciutat de Girona. També concorre la voluntat de poder fer polítiques proactives per evitar actual de forma reactiva a les situacions de risc i àmbits competencials dels serveis socials.

Com a oportunitat, l'Àrea de Serveis Socials genera i disposa d'una gran quantitat de dades a través de la seva eina de registre d'intervenció. Fins a l'inici del projecte, aquestes dades únicament s'emmagatzemaven per fer consultes puntuals i seguiments, i no s'empraven per extreure'n coneixement ni fer anàlisi de dades. Per tant, l'explotació d'aquestes dades pot possibilitar anàlisis robustos que poden ajudar en la feina que es desenvolupa a l'àrea.

Aquestes constatacions van portar a la ideació d'un projecte que permetés crear una eina capaç de fer anàlisis i prediccions sobre la realitat social de la ciutat, treballant no només sobre la gran quantitat de dades de l'Àrea de Serveis Socials, sinó també combinant aquestes dades amb altres fonts d'informació. A partir de converses preliminars amb investigadors de la Universitat de Girona, sorgeix la proposta d'aplicar els últims avenços en intel·ligència artificial a l'àmbit dels serveis socials i crear així una eina d'anàlisi i consulta que faciliti les tasques als tècnics de l'àrea: el Projecte Pigall.

Amb aquest objectiu, l'any 2022 l'Ajuntament de Girona sol·licita una subvenció al Departament de Drets Socials de la Generalitat de Catalunya, en el marc del Pla de recuperació, transformació i resiliència, finançat pel fons Next Generation EU dins la línia 3, "Projectes tecnològics que contribueixin a la millora i transformació dels models de gestió dels serveis socials" per dur a terme el Projecte Pigall.

1. EL PROJECTE PIGALL

El projecte Pigall té com a principal objectiu implementar analítica avançada a partir de les dades que Serveis Socials i l'Ajuntament generen per així poder oferir una optimització i una millora de l'eficiència en la gestió dels recursos.

Consisteix en el desenvolupament d'una eina basada en Intel·ligència Artificial per entendre les dinàmiques socials de la ciutat de Girona i anticipar possibles escenaris de futur, per poder ser proactius en l'abordatge de les situacions socials de l'àmbit competencial de serveis socials.

El projecte Pigall és alhora un procés científic-tècnic de disseny i un producte tecnològic final, és la metodologia i l'eina.

Es tracta de desenvolupar un producte tecnològic i alhora sistematitzar un mètode de treball, per tant en aquest document, s'expliquen les característiques i funcionalitats del producte final, l'eina Pigall, i alhora el procés metodològic desenvolupat, la metodologia Pigall.

1.1 PRINCIPIS DEL PROJECTE

Col·laboració científica

El projecte Pigall es construeix gràcies a una aliança estratègica entre l'Ajuntament de Girona, el grup de recerca eXiT de l'Institut d'Informàtica i Aplicacions (iiiA) de la Universitat de Girona, i l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya.

L'Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC), que pren forma de Càtedra a la Universitat de Girona, s'insereix en l'Estratègia d'Intel·ligència Artificial (IA) impulsada per la Generalitat de Catalunya amb el nom de d'Estratègia Catalunya IA. L'estratègia inclou un eix sobre "Ètica i Societat", que és on es situa l'OEIAC, per tal de promoure el desenvolupament d'una IA ètica, que respecti la legalitat vigent, que sigui compatible amb les nostres normes socials i culturals, i que estigui centrada en les persones.

El grup eXiT, Enginyeria de Control i Sistemes Intel·ligents, és un grup de recerca interdisciplinari de l'Institut d'Informàtica i Aplicacions (IiA) de la Universitat de Girona (UdG), amb una sòlida experiència en la transformació digital en les àmbits de la salut i les ciutats intel·ligents. Les línies de recerca d'aquest grup es centren la aplicació d'Intel·ligència artificial en els àmbits de la salut (e-health) i les ciutats i xarxes energètiques intel·ligents (smart cities i smart grids).

El grup eXiT està acreditat amb el segell TECNIO. Aquest és el segell que atorga la Generalitat de Catalunya, a través d'ACCIÓ, per identificar i donar visibilitat als desenvolupadors de tecnologia del sistema R+D+I de Catalunya que presentin capacitats tecnològiques diferencials i amb capacitat de transferir-les per enfortir el teixit empresarial i fer atractiu internacionalment el ecosistema català d'innovació.

El treball conjunt entre l'acadèmia, els tècnics socials municipals i els tecnòlegs contractats és un dels punts forts del desenvolupament de l'eina Pigall, que ha permès conèixer els últims avenços en Ciència de Dades i Aprenentatge Automàtic més adequats per a les aplicacions proposades.

A més, aquesta col·laboració també s'ha concretat en l'assessorament en els usos ètics i socials de la IA, en particular, tenint en compte els següents aspectes:

- El codi de bones pràctiques d'usos de la IA arreu del món i la legalitat vigent (especialment en el context Europeu).
- L'impacte de la implementació d'aquesta tecnologia en la provisió de serveis socials.
- Promovent una metodologia de treball que combinés una explotació quantitativa de les dades disponibles (més pròpia de l'àmbit tecnològic) amb anàlisis i discussions de caire qualitatiu (més propis de les ciències socials).

Durant tot el procés de desenvolupament del projecte, s'ha mantingut una reunió mensual de l'equip Pigall amb els assessors de l'Ajuntament de Girona. Aquestes reunions han servit per supervisar els avenços realitzats en cada fase del projecte, revisar resultats obtinguts i proposar solucions tecnològiques per a l'anàlisi i tractament de les dades i pel desenvolupament de les funcionalitats de l'eina Pigall. També han servit per validar els resultats obtinguts i comprovar que aquests eren coherents i adients, tant a nivell de qualitat tecnològic, com des de la perspectiva social.

S'ha tingut present comprovar que tant en les funcionalitats proposades, com en el procés de reflexió i elaboració de les solucions implementades s'estaven seguint els principis ètics que són constitutius del projecte per crear una eina d'IA responsable.

Validació i contrast amb professionals de serveis socials

El pilar fonamental sobre el qual es construeix el projecte Pigall és la col·laboració i el diàleg constant entre els tècnics de l'àmbit social i els tecnòlegs que han desenvolupat el projecte amb la resta de membres de l'Àrea de Servei a les Persones (en endavant Serveis Socials o Àrea de Serveis Socials), així com amb la resta d'àrees de l'Ajuntament de Girona.

Aquest diàleg constant entre els diferents actors del projecte ha permès identificar necessitats, veure quines solucions són tècnicament factibles i assegurar que el desenvolupament tècnic del projecte avanci sempre en la direcció correcta.

Aposta per la innovació des de l'Administració pública

A més, es dissenya i es manté des dels recursos propis de l'Ajuntament, sense dependències d'operadors provats. Això s'ha fet a partir de **la contractació directa de l'equip tècnic Pigall** format per 2 enginyers de dades, un matemàtic amb coneixements d'IA i una tècnica social. A més s'ha comptat amb la direcció d'una professional pròpia de l'Ajuntament de Girona i suport administratiu.

Aquesta aposta, suposa que no hi ha ànim de lucre ni transferència de dades. Alhora el projecte és immune a eventualitats i discontinuïtats de la contractació pública. Es pot garantir una IA responsable amb models transparents¹.

¹ La metodologia Pigall incorporarà un informe de transparència algorítmica que properament es publicarà al Portal de Transparència de l'Ajuntament de Girona.

2. L'EINA PIGALL

L'eina Pigall, com a producte, facilita la presa de decisions tècniques per a una millora de la planificació de la intervenció social.

L'eina Pigall es crea amb les següent finalitats:



- **Anàlisi:** proporciona informació de la realitat de la ciutat per identificar problemàtiques.
- **Planificació:** assisteix en el disseny d'estratègies i planificació/assignació dels recursos disponibles.
- **Predicció:** permet anticipar-se a escenaris de risc futurs.
- **Automatització:** realitza automàticament algunes tasques mecàniques realitzades pels tècnics de Serveis Socials i que no requereixen prendre decisions.

En concret es tracta d'una **infraestructura tecnològica que es divideix en el visor web i el servidor on es gestiona tot el processament de les dades, la creació d'API's i sistema d'avisos.**

Servidor

Dins del servidor s'implementa tot el codi que ajuda a actualitzar la informació que posteriorment es visualitza en el visor web.

La finalitat de tenir un servidor que permeti treballar les dades és que es pugui generar nova informació a partir de les dades recollides i permeti també l'optimització de tasques que poden consumir temps.

A més, es pretén desenvolupar models d'aprenentatge automàtic que seran carregats en aquest servidor i utilitzats a partir d'APIs.

Visor Web

El visor web és la part visual del projecte, i que està a disposició dels usuaris. La plataforma que s'utilitzarà és Oracle APEX, ja que és el programari que s'utilitza des dels Serveis Sistemes i Tecnologies de la Informació de l'Ajuntament i d'aquesta manera es garanteix el manteniment i sostenibilitat.

D'acord amb les finalitats previstes, l'eina té 4 funcionalitats bàsiques:

- Quadre de comandament; amb visor de variables i correlació de dades internes i externes.
- Escenaris de futur: Projecció i predicció de situacions socials per anticipar les polítiques socials i la intervenció. Simulacions d'escenaris.
- Quadre d'avisos: sistema d'alertes en relació a llindars crítics respecte demandes socials.
- Eines de consulta; per facilitar gestions i tasques.

El conjunt de dades disponibles es mostren en el visor directament, d'altres amb informes

automàtics, i d'altres geolocalitzades a la ciutat, mostrant mapes de calor respecte els territoris o entorns més afectats per una situació social o variable determinada.

Bases de l'eina

- És una eina d'ANÀLISI i planificació.
- Utilitza la base de dades interna de SERVEIS SOCIALS i la posa en relació a dades externes estandarditzades (IDESCAT, SEURETAT SOCIAL, SEPE/SOC, INE,...).
- No pren decisions automatitzades sobre la vida de les persones. Ajuda a la presa de decisions humanes dels professionals i responsables polítics.
- S'utilitza de manera interna a l'Ajuntament de Girona. Hi accedeixen només professionals autoritzats amb codis d'accés.
- No és consultable ni està a disposició de la ciutadania en general.
- No transfereix dades a tercers ni entitats privades.
- Vol mostrar escenaris i dades agregades però necessita arribar a la màxima granul·lilitat justament per agregar: per barris /sectors, i geolocalitzar necessitats.

Els usuaris/es són professionals de serveis socials responsables de serveis i per tant amb capacitat de presa de decisions i organització; directores dels equips bàsica d'atenció social (EBAS), directora de serveis atenció a la gent gran i dependència, directora de l'Oficina Municipal d'habitatge, cap de l'Àrea de serveis socials.

De cara a la sostenibilitat futura de la infraestructura tecnològica s'ha utilitzat la plataforma APEX que és la que s'utilitza en el desenvolupament d'aplicacions informàtiques de l'Ajuntament de Girona. Es resta pendent explorar la compatibilitat amb la futura base de dades WSOCIAL, de la Generalitat de Catalunya.

Tant pel desenvolupament de l'eina, com per a la seva sostenibilitat a llarg termini, s'ha treballat en col·laboració amb altres àrees de l'ajuntament com el Servei de Gestió Documental, Arxiu i Publicacions, la secció d'Estadística, el servei de Sistemes i Tecnologies de la Informació i la Unitat Municipal d'Anàlisi Territorial (UMAT).

A l'apartat 4 "Producció" es desenvolupen amb més detall les característiques i funcionalitats.

3. METODOLOGIA PIGALL

La metodologia Pigall és el conjunt de processos desenvolupats per a l'elaboració i disseny de l'eina. **La metodologia utilitzada és on recau el valor de la replicabilitat del projecte, ja sigui en altres àmbits de la política pública, com en altres municipis amb la mateixa finalitat orientada a serveis socials.**

Aquest procés inclou el diàleg permanent entre professionals especialitzats i tecnòlegs respecte l'explicabilitat i la utilitat de les aplicacions, així com la recerca pel que fa a procediments d'anàlisi i processament de dades per a posar-les al servei de les necessitats previstes.

La metodologia Pigall es pot resumir en els següents passos:

- 1. Detecció de necessitats:** el projecte Pigall neix d'una sèrie de necessitats l'Àrea de Serveis Socials. El primer pas consisteix en definir rigorosament aquestes necessitats, veure quines solucions es poden oferir des del Pigall, i si són tècnicament factibles. En aquesta fase també és important realitzar una recerca bibliogràfica exhaustiva per identificar altres experiències similars que es puguin estudiar, i també per identificar quines tècniques estadístiques, matemàtiques i d'Intel·ligència Artificial es poden aprofitar per al projecte.
- 2. Exploració de dades:** conjuntament amb el pas anterior, també cal fer un anàlisi de les dades que té disponibles Serveis Socials i la seva qualitat, identificant com es poden millorar, processar i explotar per donar resposta a les necessitats identificades.
- 3. Models predictius:** Un cop analitzades i tractades les dades disponibles, es desenvolupen els programes i models predictius, que utilitzen eines estadístiques, matemàtiques i d'aprenentatge automàtic per extreure coneixement d'aquestes dades.
- 4. Producció:** Una vegada superada l'etapa de desenvolupament, es comença a crear el visor web on es mostraran els resultats.
- 5. Testeig i formació:** L'eina es comença a utilitzar per les professionals per tenir una opinió sobre les funcionalitats i l'estètica del visor i així poder millorar l'experiència d'usuari. Per últim, també es fa un acompanyament i una formació sobre les potencialitats de l'eina.

Pel que fa a aquestes fases i en relació al calendari d'execució del projecte Pigall, les fases exposades corresponen al següent cronograma:

FASES	DESCRIPCIÓ	2024	2025			
		JUL - DES	1r trim	2n trim	3r trim	4rt trim
1a Fase: Detecció de necessitats i recerca						
Recerca d'experiències prèvies i sistemes d'informació	Recerca sobre antecedents i experiències d'èxit en la implementació de IA i Bigdata en els serveis socials					
Treball de camp amb professionals socials de l'Àrea	Treball d'anàlisi i formulació de preguntes importants a les quals l'eina ha de donar resposta; entrevistes, focus grup					
2a Fase: Exploració de dades						

Anàlisi de les dades disponibles i la seva qualitat	Anàlisi de la qualitat de les dades. Harmonització de dades i anàlisi descriptiu					
Tasques de "millora" de les dades disponibles	Aplicació de mecanismes de millora de la qualitat de les dades. Valors nuls, regles de coneixement, etc,...					
Recull d'indicadors i variables disponibles	Recull de variables més adequades per respondre les preguntes formulades					
Centralització de la informació	Sol·licitud de dades a altres administracions i integració de dades					
3a Fase: Models predictius						
Construcció de models	Anàlisi i comprovació de validesa de les dades corresponents a les preguntes amb caràcter prospectiu. Construcció de models en base a les preguntes amb caràcter prospectiu					
Elaboració de l'esquema generador de preguntes	Selecció de variables internes i externes i sistema generador de preguntes i funcionalitats de l'eina. Validació amb professionals.					
Elaboració de l'AIPD i accés al Padró Municipal d'habitants	D'acord amb les variables seleccionades i les funcionalitats previstes, elaboració de l'anàlisi d'impacte en la protecció de dades i accés al PMH.					
4a Fase: Producció						
Disseny de l'eina i visor	Elaboració de prototips visuals que responguin a les funcionalitats previstes. Validació amb professionals.					
Preparació del Backend	Servidor amb màquina Linux, Gitlab, Airflow, FastAPI, Docker i PostgreSQL instal·lats					
Disseny de l'arquitectura relacional	Emmagatzematge de les dades generades en els processos ETL a PostgreSQL					
Disseny dels processos ETL	Disseny dels scripts ETL amb l'objectiu de disposar de dades actualitzades					
Disseny del Frontend	Disseny de les diferents pàgines de l'eina					
Desenvolupament ETL i API	Desenvolupament dels scripts ETL per nodrir PostgreSQL i modificar i/o ampliar informació de les dades originals en base a regles de coneixement Desenvolupament API per models i altres peticions					
Maquetació Frontend	Construcció del Frontend					
Models en producció i monitorització	Models en producció amb FastAPI y monitorització del rendiment					
5a Fase: Testeig i formació						
Testeig	Realització de test global					
Elaboració del Manual usuari	Redacció final del manual d'ús de l'Eina prèvia a la formació dels professional					
Elaboració documentació	Documentació (la documentació es realitzarà en totes les etapes)					
Formació	Formació als diferents serveis i equips de l'àrea de Serveis Socials					

A continuació s'exposa el detall de les tasques realitzades en cada fase del projecte, d'acord amb la metodologia emprada, que es correspon als següent apartats.

3.1 DETECCIÓ DE NECESSITATS I RECERCA D'EXPERIÈNCIES

3.1.1 Recerca d'experiències i sistemes d'informació

A l'inici del projecte, es realitza una recerca bibliogràfica sobre antecedents, experiències i casos d'èxit en projectes d'Intel·ligència Artificial i Ciència de Dades en l'àmbit de Serveis Socials, amb l'objectiu de conèixer quines eines i aplicacions s'han desenvolupat i quines problemàtiques han hagut d'afrontar en el processos

Per conèixer directament algunes d'aquestes experiències i projectes similars, des de l'equip Pigall es van realitzar diverses trobades:

- Amb els Serveis d'Innovació i Nous Projectes de l'Institut Municipal de Serveis Socials de Barcelona per conèixer les línies de treball i els diferents projectes i productes que han desenvolupat i implementat als serveis socials de la ciutat de Barcelona.
- Amb la Diputació de Barcelona per conèixer línies de treball i projectes que han dut a terme.
- Amb la responsable del Pla de Transformació Digital dels Serveis Socials Bàsics del Departament de Drets Socials. Concretament aquesta trobada serveix per conèixer la plataforma WSocial, quins canvis i noves prestacions s'estan desenvolupant.
- Amb la cap de Serveis Socials del Consell Comarcal de la Selva, ja que participa en la prova pilot de la implementació del programa wSocial del departament de Drets Socials de la Generalitat de Catalunya.

De les diferents trobades se n'extreuen reflexions i conclusions respecte encerts i dificultats de les diferents tecnologies que s'estan desenvolupant, això com les potencialitats de la pròpia proposta.

3.1.2 Entrevistes i trobades amb professionals de l'àrea i de l'Ajuntament

Per a desenvolupar una eina que s'adapti a les necessitats de l'àrea, es segueix un procés metodològic de recerca-acció. Aquest procés s'ha portat a terme mitjançant un treball amb elements de co-creació i participació al llarg del procés d'elaboració del producte final a tres nivells:

a) A nivell intern de l'equip Pigall, entre els professionals socials i els tecnòlegs:
Treballar pel mateix objectiu suposa un aprenentatge per ambdues parts, tant la social com la tecnològica. Acostar els dos llenguatges i construir una visió compartida, entendre i prioritzar les necessitats de cada part, és una de les tasques principals per tal d'aconseguir l'èxit del projecte.

b) A nivell d'Àrea de Serveis Socials:
Durant la realització d'aquest treball prospectiu, i seguint la metodologia d'investigació-acció, es recullen les demandes de l'equip directiu de l'àrea, de manera que aquesta tasca inicial

consisteix en el plantejament i disseny de l'eina que donarà resposta a aquestes necessitats aconseguint un consens intern.

L'equip directiu de l'àrea de Serveis Socials manifesta la necessitat d'una eina de suport per optimitzar la gestió dels Serveis Socials i el disseny de les polítiques socials. S'identifiquen tres funcionalitats necessàries:

- Una eina de diagnòstic, que permeti visualitzar la informació rellevant per a una gestió més eficient dels recursos i per prendre decisions a partir de dades.
- Una eina per a la planificació i d'anàlisi prospectiu que permeti dimensionar els recursos disponibles i els necessaris.
- Una eina de predicció, que realitzi projeccions de possibles escenaris d'intervenció social en el futur.

En aquesta primera fase de treball del projecte Pigall, també es realitzen trobades amb les tècniques responsables dels diferents serveis propis, per conèixer els processos de treball i les bases de dades vinculades a cada un d'aquest:

- Servei Porta d'Entrada (SPE): Servei de primera atenció telefònica a les persones que sol·liciten ser ateses pels Serveis Socials.
- Centre de Distribució d'Aliments (CDA); servei co-gestionat amb entitats del tercer sector que distribueix dels recursos mitjançant cistelles d'aliments i targetes moneder.
- Equips Bàsics d'Atenció Social (EBAS): equips territorialitzats en 7 sectors, que duen a terme l'atenció social bàsica².
- Servei d'Atenció a la Gent Gran i Dependències (SAGGID): Atén les persones amb dependència, dissenya els plans individuals d'atenció (PIA) i gestiona els serveis d'ajuts a domicili (SAD), teleassistència, càterin i menjadors per a gent gran.
- Oficina Municipal Habitatge (OMH). Amb funcions de gestió dels processos de desnonament i emergència habitacional.

Al mateix temps es realitzen sessions de treball grupal amb professionals dels diferents equips esmentats per identificar "les preguntes" que tenen més interès i utilitat per la seva feina.

Per incorporar la mirada territorialitzada es realitzen 2 tallers, un amb les direccions d'SBAS, i un altre amb les direccions de centres cívics, amb el següent plantejament:

A nivell de barri, quina informació es necessita per elaborar un diagnòstic que sigui útil per la tasca dels serveis socials i comunitaris?

Quines dades són les més rellevants per la teva tasca a nivell de barri? I quina informació a "temps real" és la més important pel treball comunitari?

Per al pronòstic, quina informació necessitaríem per fer seguiment de la seva evolució o quins temes ens interessa fer seguiment de com evolucionaran en els propers anys?

Els resultats es poden dividir en agrupacions de necessitats detectades a les que podria donar resposta el projecte:

² Els EBAS de la ciutat de Girona s'organitzen en 7 sectors territorials: St. Narcís, Sta. Eugènia i Can Gibert del Pla, Palau, Pont Major, Ter-Taialà, Girona Est i Barri Vell.

- Disposar de dades sociodemogràfiques i socioeconòmiques de la població actualitzades de cada barri de la ciutat.
- Poder visionar quantes Ajudes d'Urgència Social (AUS) s'atorguen al barri, la distribució, mesurar la carrega assistencial d'un sector, conèixer si hi ha nuclis familiars que concentren els recursos i preveure si cal una redistribució dels recursos pels següents anys.
- Evolució de la situació de l'habitatge i com això repercutirà a les famílies més vulnerables, en els desnonaments i en les ocupacions.
- Interessen dades sobre la situació dels joves en els barris, sobretot dels que no acaben els estudis, així com de quines alternatives existeixen.
- Interessen dades de salut de la població, sobretot de salut mental, suïcidis i de salut sexual i reproductiva, sobretot pel què fa al jovent.
- Hi ha interès per conèixer els canvis en la població de la tercera edat i dades sobre envelliment, vulnerabilitat i soledat.
- Hi ha necessitat de disposar d'una eina que permeti millorar la gestió dels recursos i la necessitats de dimensionar els serveis atenent l'evolució demogràfica.
- Hi ha necessitat d'unificar la informació respecte als procediments de desnonament, ja que aquesta informació està molt dispersa, la qual cosa dificulta la intervenció i l'acompanyament social en els processos de desnonament.

A partir d'aquesta tasca de participació i interlocució prèvia per identificar quines necessitats tenen els diferents serveis de l'Àrea de Serveis Socials, es comença a definir els continguts de l'eina Pigall.

3.1.3 El contingut de l'eina Pigall

La idea principal és que des de Serveis Socials es puguin observar i comprendre les dinàmiques territorials de la ciutat a diferents escales espacials i temporals. Així doncs, es pretén conèixer la informació actual per entendre el present; analitzar les dades històriques per veure possibles dinàmiques estacionals, cícliques o accidentals; i fer prediccions a futur per ser proactius en la presa de decisions.

a) Nivells d'anàlisi:

Es vol poder arribar a 3 nivells d'anàlisi

Persona / Cas

El cas és la unitat bàsica d'anàlisi de serveis socials. Un cas pot ser a nivell individual o de nucli familiar, i pot presentar una única problemàtica o vàries.

Barri (sector EBAS)

El barri o sector de serveis socials és el segon nivell d'anàlisi de l'escala i ens permet observar les dinàmiques que es donen dins d'un territori determinat per a cada EBAS i també posar-ho en relació a la resta de la ciutat. Presenta una distribució espacial i temporal dels indicadors.



Ciutat

La ciutat és la unitat d'anàlisi més gran i les dades que ens ofereix són per conèixer la realitat contextual i possibles tendències i comportaments d'indicadors diversos.

b) Àmbits d'interès

El ventall d'àmbits d'intervenció dels serveis socials i el seu marc competencial és molt ampli, tant pel que fa a col·lectius a atendre, com pel que fa a situacions a abordar; infància en risc, famílies, dones, situacions de manca d'ingressos i empobriment, situacions de violència, gent gran i envelliment, dependència, manca d'habitatge, acció comunitària, etc.

Després del procés de consultes realitzat amb les persones referents dels serveis, i en relació a les dades disponibles els àmbits d'interès sobre els quals treballar es concreten en els següents:

- Habitatge: evolució i anàlisi de dades en relació al risc de pèrdua d'habitatge
- Envel·liment: gent gran i vulnerabilitat.
- Socioeconòmic: evolució dels d'ajuts econòmics puntuals i característiques dels seus beneficiaris.
- Volum d'atenció als serveis: Increment de sol·licituds d'atenció.

Aquesta selecció deixa fora molts dels àmbits esmentats, però obeeix a un doble criteri; d'una banda, la disponibilitat de dades enregistrades respecte l'àmbit a nivell de completesa i històric, i de l'altra, la rellevància social de la temàtica en el moment actual.

c) Infraestructura tecnològica

Per poder assolir els objectius del projecte, es proposa la creació d'una infraestructura tecnològica dividida en:

- Un visor web, que servirà als usuaris per interactuar amb l'eina i poder visualitzar informació.
- El servidor, on es processaran totes les dades i es construïran els models d'aprenentatge automàtic.

El visor web tindrà 3 pàgines principals, que responen a tres funcionalitats possibles:

1- Quadre de comandament

Presenta la informació de manera que permet la visualització de les dades en diferents formats.

Imatge del prototip inicial:



2- El quadre d'avísos

Es tracta d'un sistema d'alertes que s'activaran en el moment en què les variables presentessin canvis destacables.

Imatge del prototip inicial:



3- Escenaris de futur

Els escenaris de futur consisteixen en models matemàtics i/o d'aprenentatge automàtic que fan prediccions a curt i llarg termini per preveure una situació concreta sigui d'interès

Imatge del prototip inicial:



3.2 EXPLORACIÓ DE LES DADES

3.2.1 Coneixement de les dades internes disponibles i comparabilitat amb altres sistemes

Pel que fa a les eines de registre els Serveis Socials municipals per al registre de la intervenció social, les dues principals són el SIAGI-Serveis Socials i l'eina de registre de la Matriu d'Autosuficiència (SSM-CAT).

El projecte s'inicia amb un treball d'anàlisi de les dades disponibles. Des de la UdG s'encarrega a l'equip Pigall analitzar les dades recollides amb SSM-CAT i harmonitzar-les amb les dades de SIAGI-Serveis Socials per valorar la validesa científica per a la construcció de nous indicadors.

SIAGI- Serveis Socials³

El SIAGI – Serveis Socials és l'eina de registre d'intervenció dels serveis socials de l'Ajuntament de Girona. Creat l'any 2009 pels serveis informàtics municipals per donar resposta a la necessitat de registrar la intervenció dels Serveis socials en les seves diferents competències. Amb el temps s'ha anat ampliant i modificant per adaptar-se a les necessitats dels diferents serveis d'atenció social de l'Àrea de Serveis Socials de l'ajuntament.

També és l'eina de gestió de les prestacions pròpies, com són els Ajuts d'Urgència Social (AUS)⁴ o d'algunes de les prestacions de suport a la gent gran i persones dependents.

Les funcions principals que facilita el SIAGI són:

- Registre:
 - Registre de les dades de les persones ateses.
 - Registre de les intervencions.
 - Registre d'activitats: agenda de les persones ateses i dels usuaris.
- Gestió:
 - Actuacions: informes emesos sobre la persona atesa
 - Gestió d'ajuts econòmics d'urgència
 - Gestió de recursos d'atenció a domicili

El registre de les dades identificatives de les persones ateses i dels nuclis familiars recull dades dels diferents àmbits de la seva vida i la valoració professional del risc d'exclusió en cadascun. Els àmbits amb els quals es recullen les dades a SIAGI-Serveis Socials són:

1. Situació administrativa de residència
2. Àmbit econòmic
3. Àmbit laboral
4. Àmbit habitatge
5. Àmbit educatiu-formatiu
6. Àmbit relacional
7. Àmbit de salut

³ L'Acrònim SIAGI respon a les sigles de *Sistema Informació Ajuntament de Girona*, i té una aplicació específica per a la gestió i intervenció de serveis socials. Per això la denominació SIAGI-Serveis Socials. Veure ANNEX 1 Glossari.

⁴ Veure ANNEX 1 Glossari.

8. Situació relacional

El *registre d'intervencions* recull el contingut de totes les entrevistes, acompanyaments i visites a domicili que s'han mantingut amb la persona atesa i les reunions de cas on s'ha parlat de la seva situació.

El *registre d'activitats* és l'agenda de cada persona atesa per cada servei de l'àrea de Serveis Socials.

SSM-CAT

La Matriu d'autosuficiència SSM-Cat, impulsada pel Departament de Drets Socials i Inclusió de la Generalitat de Catalunya i utilitzada per totes les àrees bàsiques de serveis socials de Catalunya, és un eina de calibratge especialment dissenyada per ajudar els professionals del serveis socials bàsics a identificar persones amb necessitats d'atenció social complexes. L'SSM-CAT es basa en la descripció de la situació de la persona en un moment concret vers 13 àmbits o dominis: finances, treball i educació, ús del temps, habitatge, relacions domèstiques, salut mental, salut física, ús de substàncies, Activitats bàsiques de la vida diària, Activitats Instrumentals de la vida diària, xarxa social, participació comunitària i ordre legal.

SSM-CAT s'implementa el 2024 als Serveis Socials de Girona, pel què hi ha poques dades registrades que puguin ser tractats fàcilment. Inicialment amb la col·laboració de tècniques d'EBAS, s'ha estudiat la correlació entre els indicadors de SSM-CAT i les dades recollides a SIAGI i les conclusions que se n'han extret són les següents:

- La Matriu SSM-Cat fa una fotografia instantània de les capacitats d'una persona en un moment determinat i en canvi al SIAGI es recullen les dades de cas partint de la teoria d'inclusió/exclusió, pel què s'estan valorant qüestions diferents.
- Tampoc hi ha una equivalència total entre la majoria de categories de l'SSM-Cat i SIAGI.

Tot i que SSM-CAT té moltes potencialitats per l'anàlisi, la seva implementació als Serveis Socials és recent, per la qual cosa no es disposa d'un volum suficient de dades per nodrir l'eina Pigall. No obstant, no es descarta valorar les seves potencialitats quan l'eina estigui més evolucionada.

L'SSM-CAT és encara un reservori de dades amb poc volum i incomplet, fet pel qual s'ha descartat aquesta via de treball i s'ha continuat l'anàlisi amb les dades del sistema de registre de SIAGI-Serveis Socials, ja que:

- Permet obtenir un nombre important de dades a nivell numèric, que permetrà l'aplicació de models predictius.
- Permet disposar d'un històric en l'atenció, la qual cosa permet entendre les dades a nivell evolutiu.
- Permet disposar de dades dels diferents àmbits assenyalats d'interès.

Padró municipal d'habitants (PMH)

L'altre font principal de dades municipals disponibles procedeixen del **Padró Municipal**. L'accés a aquestes dades representa un dels obstacles més importants s'han hagut de superar per avançar amb el desenvolupament de l'eina Pigall. S'ha treballat amb l'Encarregada de Protecció de Dades municipal minimitzar l'impacte en protecció de dades de l'eina.

Per tal de prendre les mesures necessàries per respectar els principis de confidencialitat, integritat i disponibilitat definits a la llei vigent de protecció de dades, es realitza la corresponent AIPD (Anàlisi d'Impacte en la Protecció de Dades).

3.2.2 Anàlisi de les dades internes disponibles i la seva qualitat: SIAGI-Serveis Socials

L'anàlisi preliminar permet una contextualització analítica de la situació actual de les dades generades per Serveis Socials. Per fer-ho, l'anàlisi s'ha dividit en dues parts: anàlisi sistemàtica de les taules i anàlisi detallada dels valors recollits a SIAGI per grups. Per tal de realitzar aquesta anàlisi s'ha utilitzat en gran part Qlick Sense, per així agilitzar el procés analític de les dades.

Això permet comprendre l'arquitectura de les dades i ofereix una visió general de la distribució dels registres, i també avaluar-ne aspectes com l'exactitud, la integritat, l'actualització i la confiabilitat. L'anàlisi preliminar també permet identificar possibles problemes: els nuls i duplicats. Tota aquesta tasca es recull en l'*Informe tècnic de les dades de Serveis Socials 2024* (Annex 2) que conclou que es pot seguir amb el tractament de dades.

S'analitzen les dades disponibles de cada àmbit d'inclusió que es registren al SIAGI: àmbit econòmic, residencial, relacional, educatiu, de salut i laboral.

Les dades que consten al SIAGI corresponen a:

- Usuaris⁵ totals: **78.474** (Persones ateses, persones d'interès⁶ i nucli⁷)
- Usuaris atesos: **54.623** (Excloent persones d'interès i nucli)
- Dels usuaris atesos: **2.507** són *possibles duplicats* i **42** són tipus test o errors
- Total usuaris "nets": **52.074**

S'ha analitzat detalladament tots els camps recollits a SIAGI, centrant-se en l'exactitud, la integritat, l'actualització i la confiabilitat de les dades, per així poder valorar la qualitat d'aquestes.

Distribució de dades: Aproximadament el 80% de les dades recollides són categòriques, *booleanes*, identificadors i dates, cosa que redueix els errors. Els camps de text i numèrics poden presentar més variabilitat, amb registres manuals o automàtics.

Exactitud: La majoria de camps són exactes, però hi ha problemes amb usuaris duplicats i alguns camps de tipus test que poden afectar la precisió global.

Integritat: Alguns grans conjunts de dades, com els Serveis d'atenció social i els Centres Oberts, presenten un volum de valors nuls significatiu. La informació complementària i els Centres Oberts mostren una alta proporció de dades mancants, fet que indica problemes amb la integritat. S'ha de tenir en compte, però, que de tots els casos registrats trobarem completa la

⁵ Un usuari atès és aquell que té una activitat o un expedient registrat al SIAGI

⁶ Persona d'interès és aquella persona que està relacionada amb la persona atesa, ja sigui per ser familiar o que manté una relació amb ella.

⁷ Nucli és la unitat de convivència.

informació general. No obstant, la complementària només estarà complerta en aquells àmbits que interessin per a la posterior intervenció, pel què poden contenir taules o camps buits.

Actualització: La majoria dels camps estan actualitzats fins al 2024, però alguns camps específics no es registren des de dates recents. L'anàlisi de tendències mostra variacions en la freqüència de registre que poden reflectir canvis socials o en les polítiques dels serveis socials.

Confiabilitat: La presència de registres duplicats afecta la precisió i consistència. És essencial eliminar duplicats i establir procediments per evitar futurs errors. La informació ha de ser estructurada tenint en compte les limitacions de les dades per assegurar una anàlisi acurada.

En aquest moment **també es decideix treballar a partir de variables i no en la construcció de nous indicadors**, ja que la via escollida és la construcció d'avisos i escenaris futurs. Les variables seleccionades d'altres fonts s'hauran d'enllaçar i integrar amb la informació que ja disposem per complementar-la i contribuir a la construcció dels avisos i escenaris de futur.

3.2.3 Procediments de millora de la qualitat de la dada interna⁸

El sistema SIAGI presenta un sistema de minimització d'errors de registre a partir de la captació d'informació amb valors preestablerts. Tot i això, es necessita implementar millores en la integritat de les dades. A més, en futures anàlisis serà important estructurar el sistema i hàbits de registre sobre els àmbits de Serveis Socials per reduir els biaixos i obtenir resultats més precisos.

a) Duplicats

L'alt nombre de persones usuàries duplicades es deu a diverses raons, com pot ser error a l'hora d'introduir el nom o cognom que fa que SIAGI no detecti que la persona ja està registrada o que s'introdueixi un número de document identificatiu, com el DNI i la persona ja estava registrada, però amb NIE. Es realitza una tasca d'eliminació de duplicats a partir de la revisió per part d'una tècnica i es desenvolupa com a sistema de millora un detector de duplicats.

Per evitar l'existència de duplicats, es decideix introduir a l'eina Pigall una aplicació que permeti detectar si, en el moment d'introduir les dades d'una persona, aquesta ja consta en el SIAGI. El **detector de duplicats** s'encarrega de detectar possibles usuaris duplicats a la base de dades encara que el nom o la data de naixement no coincideixin exactament. L'objectiu és ajudar a identificar si una persona que s'està introduint al sistema ja podria existir amb una lleugera variació en el nom o amb algun error en les dades.

El que es fa primer és normalitzar els textos dels noms. Això vol dir que els passa a minúscules, elimina accents i espais sobrants, de manera que diferents formes d'escriure el mateix nom quedin comparables. Després compara els noms de forma flexible utilitzant un sistema que permet tolerar errors d'ortografia, canvis en l'ordre de les paraules o cognoms que falten o sobren. Això permet detectar que dos noms són probablement de la mateixa persona encara que no siguin exactament iguals.

⁸ Veure Annex 2: Informe tècnic de les dades de Serveis Socials 2024

També revisa les dates de naixement i intenta corregir alguns errors habituals, com anys mal introduïts abans de 1900 que en realitat corresponen a anys del segle XX.

Amb aquesta informació calcula una puntuació de semblança entre la persona que es busca i els usuaris existents. Aquesta puntuació dona més importància a la coincidència dels noms (70%) i una part menor a la coincidència de la data de naixement (30%).

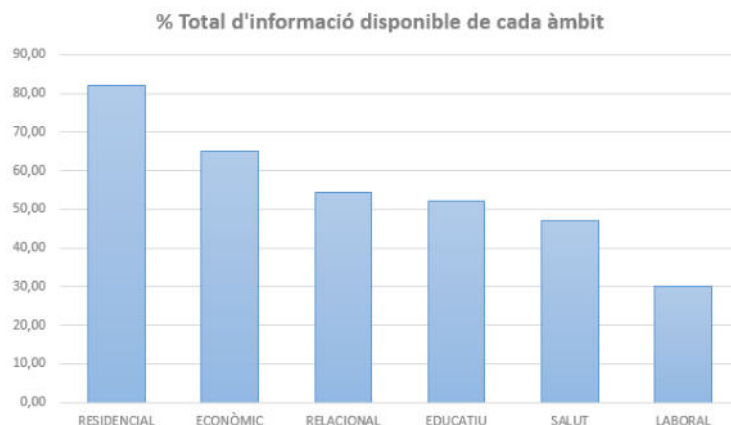
A continuació es consulta la base de dades d'usuaris i selecciona els registres que podrien semblar-se al nom introduït, descartant alguns casos coneguts de proves o registres incorrectes. Per a cada possible coincidència calcula la puntuació de semblança i es queda només amb aquelles que superen un determinat nivell de similitud.

Finalment retorna fins a cinc possibles coincidències amb el seu identificador, nom, cognoms, data de naixement i el nivell de similitud.

b) Anàlisi de la No resposta⁹

El treball continua amb la identificació i anàlisi de la **No resposta per a cada camp**, per tipus d'ajuts i l'evolució per anys. L'equip social identifica en quins camps sí s'espera resposta i ens quins no és rellevant i s'exposen els arguments per a la No resposta.

Pel que fa a la informació disponibles per a cada àmbit, en la següent taula es presenta el total d'informació disponible a cada àmbit previ al tractament de la No Resposta:



Es segueix amb l'anàlisi de la qualitat en termes estadístics i de la necessitat de la dada en termes d'interès social. Pel què fa a l'interès social, la recollida de dades es fa dels àmbits en què es dona informació rellevant per entendre la problemàtica de la persona usuària o nucli de convivència i permet configurar una fotografia de la situació per decidir com abordar la seva millora. En quant a la necessitat de la dada en termes d'interès social, aquesta té a veure amb les problemàtiques associades a una situació que han de quedar registrades.

Com podem veure a la gràfica, l'àmbit que més informació recull és el residencial. No obstant, el recull de dades en molts camps es fa de manera automàtica. És a dir, si s'omple un registre i no s'emplenen els camps, aquests seran registrats com a "NO", atès que la majoria de camps

⁹ Veure l'Annex 3: *Anàlisi de les dades dels àmbits d'actuació de serveis socials: pre-processament de les dades.*

són booleans (SI/NO). Per tant, hem d'entendre que podem obtenir resultats esbiaixats. Així doncs, l'àmbit econòmic passa a ser el més completat i per tant, amb el que treballarem primer.

c) Reducció dels valors Nuls: Aplicació de regles de coneixement

La finalitat del SIAGI no és estadística, sinó que és principalment una eina de registre d'intervenció. Això implica que no sempre consten totes les dades possibles de cada persona usuària de Serveis Socials, sinó que es registren aquelles que el/la professional social considera necessàries per a la intervenció.

En la fase de processament de dades s'apliquen estratègies de tractament dels valors nuls a partir de la integració d'altres Fonts de dades i es treballa la **No resposta** mitjançant **regles de coneixement expert**, dissenyades a partir de l'experiència dels tècnics socials i tenint en compte dues premisses:

1. La dada que no consta al SIAGI (No resposta) es pot inferir a partir d'altres dades que sí que consten o la combinació d'altres dades?
2. La dada té històric? Si és que sí, es pot inferir la dada que falta a través de l'evolució de l'històric?

La reducció dels valors nuls a través de regles de coneixement s'ha fet sobre aquests àmbits de registre de dades del SIAGI-Serveis Socials i que podeu veure amb més detall a l'Annex 3.

3.2.4 Els Ajuts d'Urgència Social com a variable seleccionada per crear una metodologia per elaborar models predictius

La Cartera de Serveis Socials de la Generalitat de Catalunya, defineix els Ajuts d'Urgència Social (AUS) com a "ajut destinat a atendre i resoldre determinades situacions de necessitat puntuals, urgents i bàsiques de subsistència en què es troben les persones que no disposen de recursos econòmics suficients per afrontar-les i no estan en condicions d'aconseguir-los o rebre'ls d'altres fonts".

Els Serveis Socials Bàsics (SBAS) són el servei que gestiona aquesta prestació. En el cas de l'Ajuntament de Girona, els AUS es tramiten d'ofici i no de part. És a dir, la ciutadania no realitza una sol·licitud prèvia d'aquesta prestació. Els AUS estan concebuts com a una eina que els/les professionals utilitzen com a recurs per atendre situacions de necessitat social. Les persones ateses exposen la seva situació i el/la professional informa i proposa el recurs més adequat per atendre aquesta situació. Al no ser un acte administratiu que s'iniciï a instància de part no existeix la sol·licitud pròpiament dita, la qual cosa no ens permet dimensionar quin és el volum de sol·licituds d'AUS, ja que els AUS analitzats són els efectivament aprovats.

Existeixen dos modalitats d'ajuts:

- Ajut puntual: es tracta d'una prestació única, no periòdica, per abordar una determinada situació de necessitat. Una persona pot tenir més d'un ajut puntual en la seva trajectòria d'atenció als SBAS i més d'un en el transcurs d'un any, existint un límit màxim orientatiu del màxim d'import en AUS que pot rebre una unitat familiar durant els 12 mesos anteriors a l'ajut puntual.

- Ajut de Pla de Treball: és un ajut amb una periodicitat puntual, no superior a 4 pagaments, per desenvolupar el pla de treball que el/la tècnica acorda amb la família atesa.

La informació continguda en el registre d'AUS del SIAGI i que ha estat objecte d'anàlisi, es registra en variables categòriques i numèriques.

Amb aquest procediment s'ha posat de manifest que l'àmbit econòmic dins el sistema de registre del SIAGI-Serveis Socials és el que té més capacitat d'explicació pels motius següents:

- Per la seva completesa:

- És condició indispensable perquè un ajut sigui aprovat per la cap d'Atenció Primària, que contingui totes les dades necessàries quan s'emet l'informe proposta.
- Els ajuts contenen una àmplia informació de dades socioeconòmiques de les persones/famílies destinatàries.

- Per la seva traçabilitat:

- Permet tenir una visió longitudinal en el temps. En aquests cas, s'ha decidit observar el període comprès entre el 2015-2025. Aquesta decisió es deu sobretot a que aquest període engloba dos fets d'especial rellevància: els efectes de la crisi econòmica que es va iniciar el 2010 i que va deixar sentir els seus efectes especialment en la crisi de l'habitatge en els anys posteriors i la pandèmia.

- Per la seva capacitat explicativa:

- Permet tenir una visió de la dinàmica socioeconòmica de la ciutat, especialment de les persones que es troben en situació de vulnerabilitat econòmica.
- Permet observar la dinàmica global de la ciutat i també disposar d'una visió de cada un dels barris-sectors de la ciutat.

Els Ajuts d'Urgència Social és per tant, la variable seleccionada per iniciar una metodologia per elaborar els primers models d'anàlisi i models predictius. En fases posteriors del projecte s'opta també per treballar el risc de pèrdua d'habitatge, la demanda al servei d'intervenció primària, i la vulnerabilitat en persones grans.

Criteris de selecció de les variables internes

Més completesa
Més continuïtat (2014-202...)
Més volum de dades
Preocupacions principals dels serveis:
Residencial, econòmic, dependència

2. VARIABLES INTERNES / Com afecta a les necessitats socials de les persones

Canvia la demanda de tipus d'AUS

Canvis en situació/ règim de tinença de l'habitatge

Canvis en el tipus de Demandes a SBAS / Activitats

Canvis en els PIA (Demandes)

Canvis en emergència habitacional

3.2.5 Dades externes

Tal i com s'ha explicat, una de les principals finalitats de l'eina Pigall és que des de Serveis Socials es puguin observar i comprendre les dinàmiques territorials de la ciutat a diferents escales espacials i temporals. Per aconseguir aquest objectiu, l'eina ha de poder posar en consonància les dades internes que generen els Serveis Socials i l'Ajuntament de Girona amb altres fonts de dades externes.

Per dades externes entenem aquelles que no són generades pels Serveis Socials i que la font pot ser el mateix Ajuntament (e.g. el Padró Municipal) o altres d'altres administracions o entitats privades i que tenen basses de dades accessibles a través d'API.

Identificació de dades possibles

En paral·lel a les taules del SIAGI de Serveis Socials que recullen variables internes, s'ha elaborat un document per treballar amb la UMAT els indicadors i variables d'altres àrees de l'ajuntament o externes i que cal estudiar-ne la qualitat i interoperabilitat.

Es va recollir i classificar un conjunt de 150 variables i indicadors que poden aportar informació que ens interessa. Aquest és un document viu que s'ha anat treballant i depurant mentre ha avançat el disseny i concreció de l'eina Pigall.



Selecció de dades disponibles

S'han seguit uns criteris a l'hora de prioritzar unes variables per sobre les altres. Aquests criteris han estat la continuïtat, la freqüència i la granul·litat de la dada.

- De la continuïtat s'ha valorat que tingui un històric d'almenys 5 anys.
- De la freqüència, que s'actualitzi com a mínim una vegada a l'any.

Criteris de selecció de les variables externes

Capacitat explicativa
Continuïtat en el temps
Granul·litat (municipi/secció censal)
Més interessants per l'anàlisi
Comparabilitat

- De la granul·litat, que el nivell de desagregació sigui el més petit possible o proper a la unitat censal.

Seguint aquestes premisses, el llistat s'ha simplificat a **25 variables externes**.

Cal destacar que, en el moment de desenvolupar els models que requereixen de la interoperabilitat de dades externes, en la majoria de casos ens hem trobat que moltes BBDD d'administracions públiques encara no estan disponibles. Aquest fet, que ha limitat en gran mesura el producte final, ens ha portat a qüestionar-nos la conveniència de promoure i

subvencionar el desenvolupament de projectes d'implementació de tecnologies d'IA, si les administracions i les dades encara no estan preparades.

Finalment, de tots els indicadors i dades proposats que reuneixin els requisits anteriorment exposats de completesa, granulitat i freqüència, que a més disposin d'API per a la seva connexió, només s'ha pogut disposar dels següents:

- Padró acumulat (Font: Ajuntament de Girona)
- Preu per m2 mitjà de lloguer (Font: Idealista)
- IPC 2014 a actualitat = (INE)
- IPC 2003 al 2013 = (EUSTAT)
- Risc Pobresa 2009 a actualitat = (INE)
- Risc Pobresa 2005 al 2008 = (DATOSMACRO)
- Cotitzacions Socials 2000 al 2022= (Observatori Girona)
- Impost Renda Patrimoni 2005 al 2022 = (Observatori Girona)
- Atur 2005 al 2024 = (IDESCAT)

3.3 CONSTRUCCIÓ DE MODELS

Com s'ha dit, Pigall es una eina que busca fer ús de tecnologies d'Intel·ligència Artificial modernes per tal d'explotar la gran quantitat de dades dels serveis socials. Això ha de permetre fer anàlitzes de la situació actual de la ciutat, i també fer prediccions a futur. Amb aquest propòsit, es desenvolupen diversos programes per al projecte que utilitzen eines estadístiques, matemàtiques i d'aprenentatge automàtic.

El principi d'explicabilitat és element clau tant de l'eina Pigall com de qualsevol projecte d'Intel·ligència Artificial en àmbits que afecten a les persones, i en particular en l'àmbit social. El **principi d'explicabilitat** en la intel·ligència artificial (IA) és la capacitat d'un sistema per fer comprensibles els seus processos (tant el raonament de les prediccions que fa, com l'impacte que puguin tenir). Es tracta d'un element clau de la IA ètica i responsable ¹⁰.

A continuació s'explica breument la metodologia seguida per desenvolupar aquests programes, això com les claus explicatives i interpretatives de les diferents eines.

3.3.1 Clústering de perceptors d'ajuts d'urgència social (AUS)

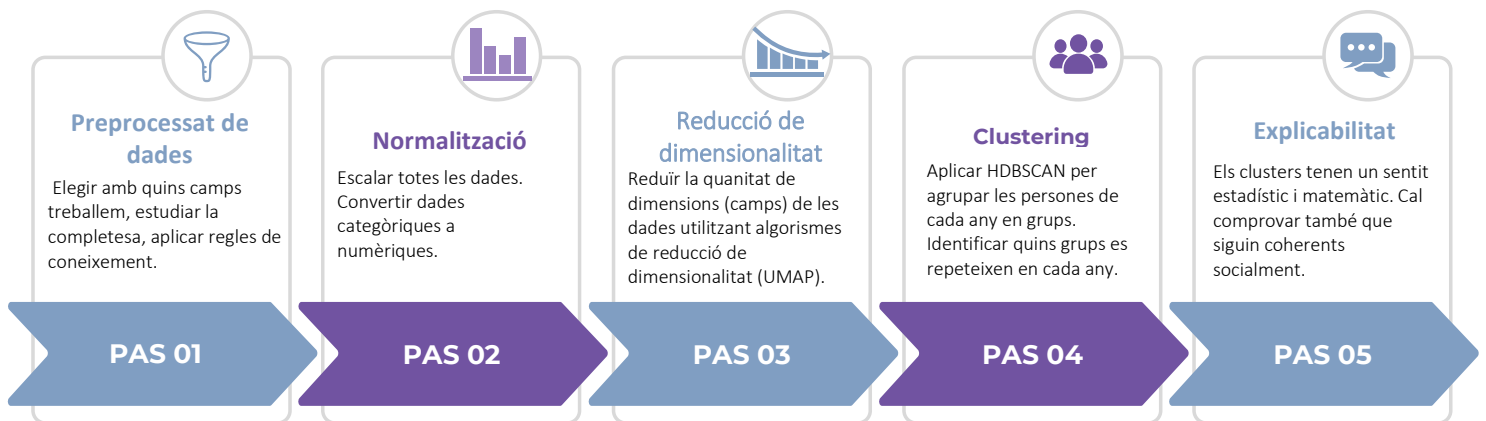
En l'àmbit de l'aprenentatge automàtic, s'anomena clústering a la tasca d'agrupar un conjunt d'objectes en diferents grups segons les seves propietats, de manera que tots els elements d'un mateix grup comparteixin alguna característica estadísticament significativa. Exemples clàssics de problemes de clústering inclouen agrupar espècies de flors segons les seves característiques filogenètiques, agrupar usuaris de plataformes on-line segons els seus patrons de comportament, etc.

En el cas de la metodologia Pigall, es decideix clusteritzar a les persones perceptores d'ajudes econòmiques per cada any. Això ens permet:

¹⁰ Explicabilitat: *Capacitat d'un sistema d'IA per presentar els seus processos, decisions i resultats de manera comprensible per als humans, revelant el "per què" i el "com" dels seus outputs de forma significativa.* Segons la "Guia d'explicabilitat de models d'IA" del CTTI. Veure Annex 3

- Veure quins grups de persones amb característiques compartides han percebut ajudes puntuals a cadascun dels anys entre 2015 i 2025.
- Estudiar si hi ha hagut alguna variació en el nombre de grups entre dos anys consecutius.
- Estudiar quins grups s'han mantingut al llarg dels anys, i quins han patit variacions (han aparegut en un any concret, s'han fusionat amb altres grups o s'han fragmentat, etc).

Per desenvolupar els programes de clústering s'ha seguit la metodologia habitual que s'utilitza per treballar aquest tipus de procediments¹¹:



La informació continguda en el registre per a la prestació d'AUS de SIAGI i que ha estat objecte d'anàlisi, es registra en variables categòriques i numèriques .

Pel que fa a les variables categòriques distingim entre variables descriptives i variables valoratives. Les primeres són identificatives de les característiques de les persones que componen la unitat familiar o descriptives de la seva situació socioeconòmica, i poden comprovar-se empíricament. Per exemple, el nombre de fills o el gènere de la persona. Les segones mostren la conclusió diagnòstica que ha fet el/la professional a través de la informació que ha obtingut durant la intervenció social. Per exemple, té "ingressos insuficients per cobrir les necessitats bàsiques" o "la xarxa ajuda a cobrir part de les necessitats bàsiques" són variables descriptives i "vulnerabilitat o risc d'exclusió econòmica", és una variable valorativa.

En relació a aquesta variable agafada com a model, es fa una tasca d'aplicació d'algorismes de clustering i classificació. Aquest anàlisi ha permès identificar comportaments, grups i la seva evolució al llarg del temps.

El dataset es correspon a les ajudes puntuals concedides entre els anys 2014 i 2025. S'ha construït un algorisme d'aprenentatge automàtic que ajuda a agrupar les dades en clústers. Una vegada fet això, es crea un nou data set on apareixeran les dades originals però etiquetades segons els clústers trobats. Amb això, s'entrena un algorisme de classificació bastat en xarxes neuronals per tal que, donades les dades d'una persona nova, pugui dir a quin clúster pertany.

El clústering ens permet l'anàlisi de les característiques de les persones que perceben un ajut econòmic d'urgència, la qual cosa ens mostra una "fotografia" de quines són les casuístiques principals de cada període, l'aparició de nous grups o situacions socials.

¹¹ Veure Annex 4 *Clústering de perceptors d'ajudes econòmiques*

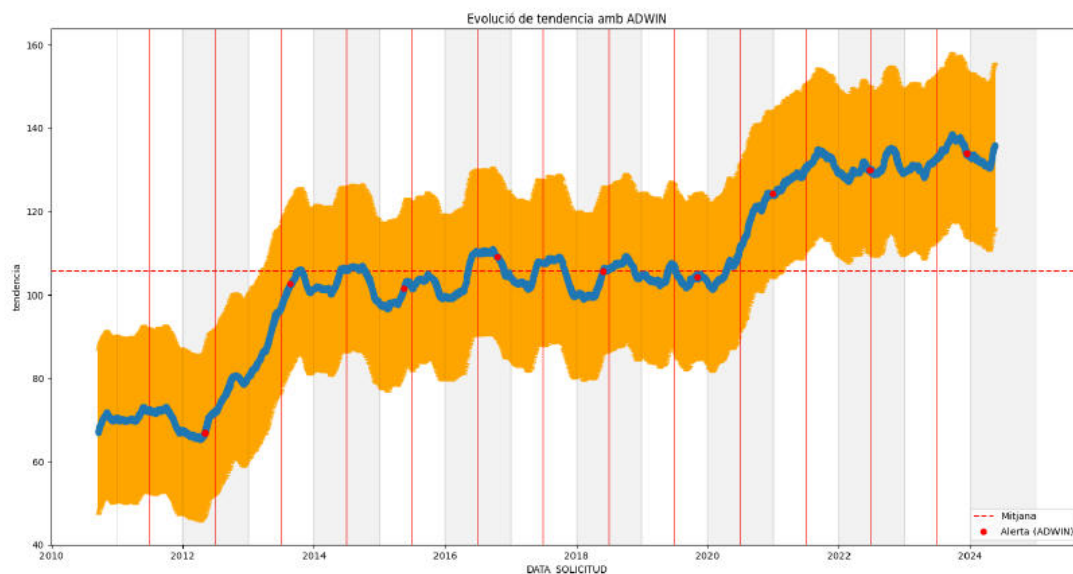
Això pel que fa a la planificació i intervenció social permet analitzar l'evolució de les característiques de les persones perceptores, per tal d'adaptar la intervenció, així com la destinació principal dels ajuts econòmics. Per exemple; es pot observar que en un any determinat la major part de perceptors són homes joves que viuen sols, i l'any següent són dones mares de llar monomarental. Dona informació sobre les transformacions en els col·lectius en situació de precarietat econòmica, així com la destinació principal dels ajuts, per exemple, en anys anteriors podia ser per "subsistència" i el 2025 poden ser majoritàriament per "habitatge".

El procediment de clústering s'aplica en relació a l'eina Pigall a la funcionalitat de quadre d'Avisos. Aquest procés s'executa automàticament una vegada l'any, a principis de gener, clusteritzant les dades de l'any anterior. També es generen una sèrie d'informes automàtics que donen informació de com han evolucionat els diferents grups.

3.3.2 Canvi de tendència al Servei Porta d'Entrada

El Pigall també disposa d'una funcionalitat capaç d'analitzar el volum d'atencions registrades diàriament al Servei de Porta d'entrada i detectar-hi automàticament canvis de tendència. Aquests canvis de tendència poden ser augments o disminucions abruptes del volum d'atencions, així com estabilitzacions.

El Pigall utilitza l'algorisme ADWIN (vegeu Annex 1) per tal de detectar aquests canvis de tendència. A mode d'exemple, la següent imatge mostra el volum d'atencions dels últims 16 anys (en taronja), la tendència (en blau) i els canvis de tendència detectats (en vermell). Aquests canvis de tendència poden relacionar-se amb esdeveniments de caràcter contextual que han impactat de manera important a les persones a nivell socioeconòmic i, per tant, han repercutit en l'atenció de Serveis Socials.



En particular, es detecta un increment important l'any 2012 com a resultat de la crisi del 2010, que resultà en un important increment de l'atur i l'endeudament. Posteriorment, el següent canvi de tendència important es manifesta a l'any 2020 a conseqüència de la pandèmia de la Covid-19 i el confinament, que va portar a moltes persones a una situació de manca d'ingresos.

En aquest sentit, la gràfica ens demostra que les deteccions de l'algorisme són correctes. I per tant, se li atribueix validesa per aplicar-lo a l'eina Pigall en la detecció de canvis de tendència, i concretament en la funcionalitat del Quadre d'Avisos.

3.3.3 Predictor de risc de pèrdua d'habitatge

Poder anticipar si una persona o família té o no risc de perdre la seva llar té molt d'interès pels tècnics de serveis socials. Per aquest projecte s'ha desenvolupat un classificador basat en l'algorisme Random Forest que, donades les dades econòmiques d'una família, és capaç de predir si està en risc de perdre l'habitatge en un futur proper.

Aquest model es fa servir en dos funcionalitats;

- El primer és per generar avisos; quan es registra una nova concessió d'ajut econòmic (sigui del tipus que sigui) a una persona. En cas que el risc sigui positiu, es genera un avís.
- En segon lloc, aquest model es fa servir també per crear escenaris de futur. A partir d'una sèrie de característiques proposades d'un nucli familiar, el model ens diu si hi ha o no risc de pèrdua d'habitatge

La resposta del predictor es binària (si/no).

A més, s'ha desenvolupat una segona versió d'aquest predictor, que treballa també amb variables externes (IPC, risc de pobresa, etc). Aquesta segona versió es fa servir únicament per la funcionalitat dels escenaris de futur i permet modificar les variables externes per tal de veure com afecta això a les persones que tenim registrades a SIAGI.

3.3.4 Risc de vulnerabilitat en persones grans

Segons la Llei 12/2007 de Serveis Socials, així com la Llei 39/2006, de promoció de l'autonomia personal i atenció a les persones en situació de dependència, es competència de serveis socials bàsics l'atenció a la gent gran en els processos d'envelliment i de dependència. Per a fer-ho, existeixen indicadors objectivables i elements de valoració social que permeten identificar si una persona gran està en risc de vulnerabilitat.

Aquests indicadors són considerats pels tècnics socials a l'atendre una persona nova o a l'actualitzar les seves dades, i serveixen per poder diagnosticar possibles situacions de risc. Tenen a veure principalment amb el seu estat de salut, els ingressos, i els suports de què disposa la persona.

Pel projecte Pigall s'han adaptat aquest conjunt d'indicadors i valoracions socials dels casos atesos de gent gran. Així doncs, s'ha creat un algorisme que permet detectar casos de vulnerabilitat en gent gran.

3.3.5 Evolució del pressupost

Utilitzant l'algorisme de predicció de sèries temporals SARIMAX, l'eina Pigall és capaç d'estudiar l'evolució de la quantitat de diners que s'hauran de concedir en ajudes puntuals durant els propers dos anys. Aquest algorisme predictiu té dues funcions: la primera és, naturalment, proporcionar una estimació de la tendència de les despeses en el futur proper. La segona és, donat un pressupost inicial, predir en quantes setmanes s'acabarà.

3.3.6 Resum de models i correspondència amb les funcionalitats de l'eina Pigall

Nom del Model	Descripció	Utilitat	Localització dins l'eina
Clústering d'AUS	Model que clusteritza els AUS anuals, i intenta identificar quins clústers s'han preservat d'un any a altre. Genera un informe a l'any amb els resultats del clústering.	Planificació, anàlisi.	Quadre d'avisos
Canvi de tendència a porta d'entrada	Genera un avís quan hi ha un canvi de tendència en el volum de la demanda a porta d'entrada.	Planificació, anticipació de risc.	Quadre d'avisos
Predictor de desnonament	-Genera un avís si es detecta una família amb risc de pèrdua d'habitatge. -Permet fer simulacions de sota quines circumstàncies una família pot patir risc de pèrdua d'habitatge.	Anticipació de risc, simulació.	Quadre d'avisos i escenaris de futur.
Predictor de risc de vulnerabilitat en gent gran	-Genera un avís si es detecta una persona gran en risc de vulnerabilitat. -Permet fer simulacions sobre quines circumstàncies fan que una persona gran estigui en risc.	Anticipació de risc, simulació	Quadre d'avisos i escenaris de futur.
Predictor de fi de pressupost	Permet predir en quin moment de l'any s'acabarà el pressupost destinat a AUS.	Anticipació de risc, simulació, planificació.	Escenaris de futur.

3.4 PRODUCCIÓ DE L'EINA

Per poder assolir els objectius d'anàlisi del projecte, es proposa com la creació d'una **infraestructura tecnològica dividida en:**

- **Un visor web**, que serveix als usuaris per interactuar amb l'eina i poder visualitzar informació.
- **El servidor**, on es processen totes les dades i es construeixen els models d'aprenentatge automàtic.

A continuació s'exposa el desenvolupament del visor web, ja que la informació respecte el servidor es pot trobar en el **Manual de la infraestructura Pigall**.

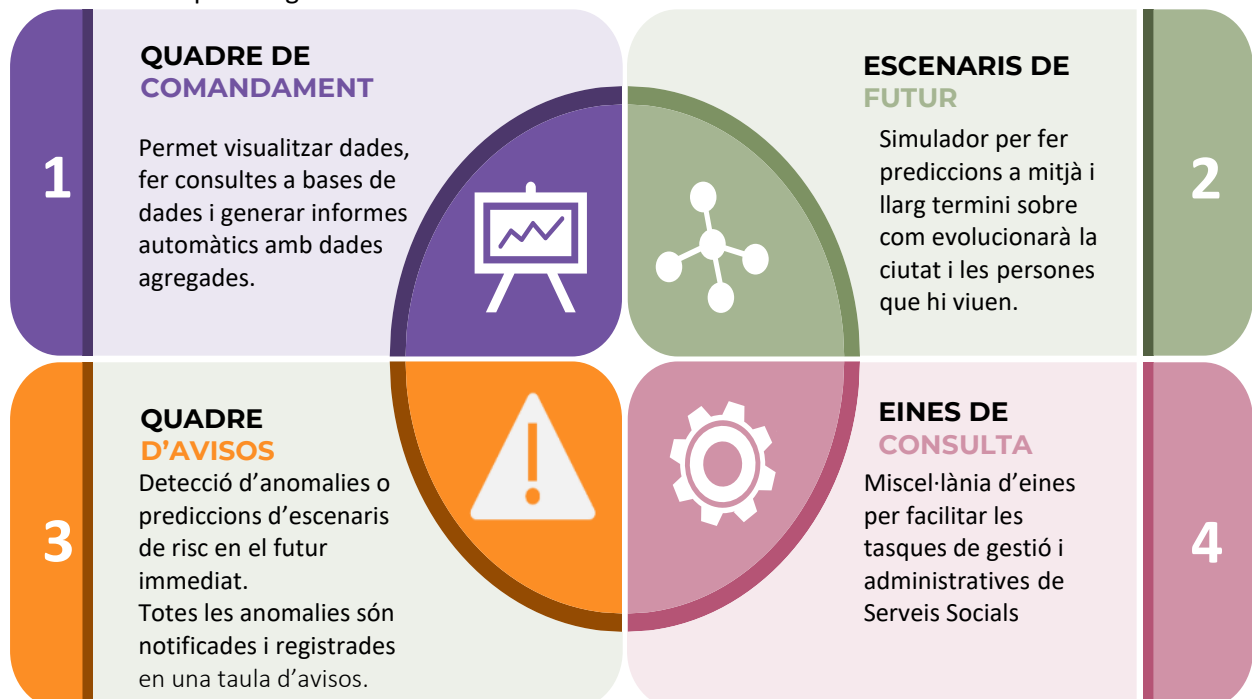
L'ús del visor web també queda explicat en el **Manual d'Usuari Pigall**.



El visor web Pigall té 4 funcionalitats diferenciades:

El *Quadre de Comandament* aporta informació sobre el passat i el present a partir d'unes determinades variables seleccionables; el *Quadre d'avisos* informa del present i del futur immediat, creant alarmes d'afectacions en determinats col·lectius o de canvis de tendències detectats, mentre que els *Escenaris de futur* permeten l'anticipació a possibles situacions si es donen determinades circumstàncies.

L'eina es complementa amb una funcionalitat *d'Eines de consulta*, que estan pensades per facilitar tasques de gestió i administratives.





3.4.1 Quadre de Comandament

Arquitectura tecnològica

El quadre de comandament es nodreix de les dades que es generen diàriament a les bases de dades **PostgreSQL** i **Oracle**, que actuen com a repositoris principals d'informació del sistema. Aquestes dades inclouen informació operativa, històrica i d'estat, i s'actualitzen de manera periòdica mitjançant processos interns d'exploració i sincronització.

La visualització de la informació al quadre de comandament es realitza a partir de peticions generades des del frontend, que sol·liciten dades específiques segons els filtres i criteris seleccionats per l'usuari. Aquestes peticions no accedeixen directament a la base de dades, sinó que es canalitzen mitjançant una capa intermèdia de serveis.

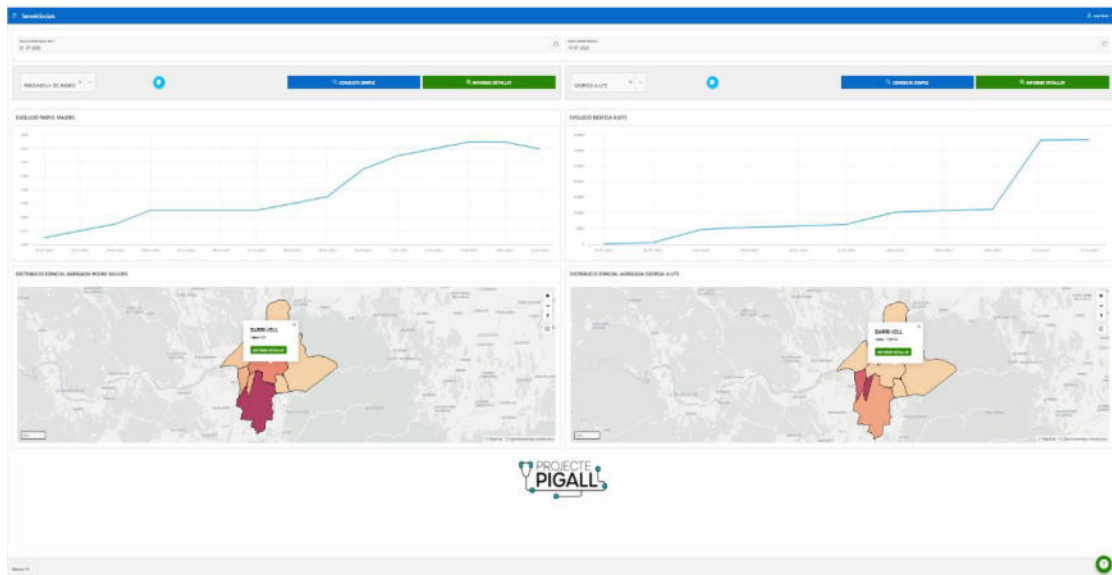
Per gestionar aquesta comunicació, s'utilitza codi **PL/SQL** o *REST Data Sources* d'**Oracle APEX** quan la informació prové d'altres sistemes o serveis externs. En aquest context, es consumeix l'API desenvolupada per l'equip Pigall¹², que actua com a capa d'abstracció i integració entre les fonts de dades i la interfície d'usuari.

Aquesta API permet estructurar, filtrar i transformar la informació abans de retornar-la al frontend, garantint així una resposta eficient i adaptada a les necessitats de visualització. Finalment, les dades rebudes s'utilitzen per alimentar les diferents gràfiques, indicadors i mapes interactius del quadre de comandament, proporcionant una visió actualitzada i coherent de l'estat del sistema.

¹² Consultable al Manual de la Infraestructura Pigall

Què mostra el Quadre de Comandament

Aquesta funcionalitat de l'eina mostra informació de les variables seleccionades, tant internes com externes, en un període temporal determinat. Aquesta informació es retorna, si la dada es pot desasgregar, en la seva distribució espacial per cada un dels 7 sectors de Serveis Socials de Girona, sobre un mapa de la ciutat de Girona. També permet fer comparatius d'aquestes variables, quan el seu comportament pot tenir relació sobre l'evolució d'una altra variable.



El Quadre de Comandament es nodreix de dades internes de l'Ajuntament (SIAGI i Padró Municipal) i permet tenir una aproximació directa a aquestes dades de manera automàtica, la qual cosa simplifica processos que abans requerien diferents eines i operacions. Per altra banda, permet tenir sempre actualitzada la informació a l'instant, ja que l'eina Pigall està connectada directament amb les bases de dades corresponents de diferents fonts.

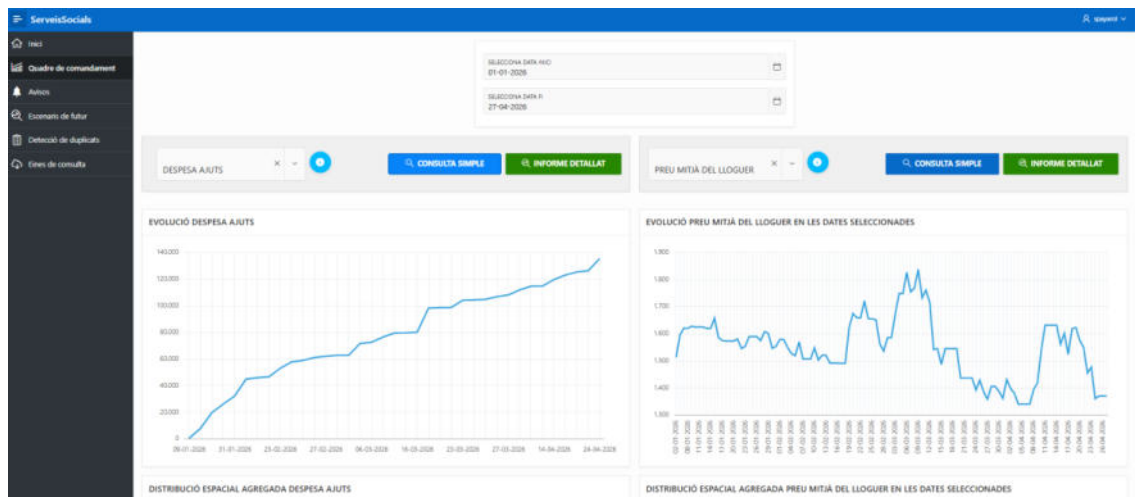
La informació que es mostra es pot veure en gràfiques temporals, tal com es veu a la imatge anterior, i també es pot descarregar informes detallats que amplien la informació mostrada. L'eina genera dos informes, un general sobre la informació de tota la ciutat i un per a cada sector de Serveis Socials.

Així mateix, les dades es poden descarregar en format excel, que conté informació més detallada, quan això és factible per tractar-se de dades internes.

La consulta Simple

La funcionalitat "Consulta Simple" retorna informació visual en format gràfica i mapa a partir de les variables seleccionades en un període temporal. Aquesta informació es desglossa (especialment pel que fa a les dades internes) en els informes detallats.

El quadre de comandament permet fer dues consultes simultànies que es mostren en dues columnes separades. D'aquesta manera es poden comparar visualment, a partir de les gràfiques, dues variables relacionades, com per exemple, l'evolució de la despesa dels ajuts econòmics en relació a l'evolució del preu mitjà del lloguer.



Les consultes es poden actualment sobre aquestes variables:¹³

Despesa d'ajuts econòmics (AUS)

Aquest apartat mostra informació relacionada amb la despesa municipal relativa als ajuts econòmics d'urgència concedits a la ciutadania. La informació s'obté a partir dels sistemes interns de gestió econòmica i de serveis socials de l'ajuntament.

Les dades inclouen els pagaments (en euros) efectuats per l'Ajuntament en exercicis actuals i anteriors, així com els ajuts d'urgència concedits a persones usuàries, tenint en compte l'expedient, el concepte de l'ajut, l'import i la data de concessió.

Aquesta informació permet analitzar com es distribueix la despesa pública destinada a l'atenció social, identificar tipologies d'ajuts, imports concedits i col·lectius beneficiaris, així com fer un seguiment temporal de les actuacions municipals.

A l'informe les dades es presenten de forma agregada i tractada, incorporant informació bàsica de les persones beneficiàries (com edat, zona de residència o situació administrativa) amb l'objectiu de facilitar una anàlisi social i econòmica que ajudi a la planificació i millora dels serveis municipals.

Padró Municipal d'Habitants

Aquestes dades provenen del Padró Municipal d'Habitants, el registre administratiu on consten les persones que resideixen al municipi. El padró és gestionat i actualitzat directament per l'Ajuntament.

La informació recollida reflecteix la població empadronada en una data determinada i permet analitzar l'evolució demogràfica del municipi, així com la distribució de la població per zones, edats o altres variables d'interès.

Persones empadronades de 0 a 17 anys.

¹³ El Quadre de Comandament pot ampliar-se amb més variables, sempre que les dades estiguin disponibles via API.

Aquestes dades provenen del Padró Municipal d'Habitants. La informació recollida reflecteix la població empadronada de persones menors de 18 anys en una data determinada i permet analitzar l'evolució demogràfica del municipi en relació als infants i adolescents.

Persones de 85 o més anys

Aquestes dades provenen del Padró Municipal d'Habitants. La informació recollida reflecteix la població empadronada majors de 84 anys en una data determinada. Aquestes dades permeten tenir una visió del sobreenvelliment de la població a la ciutat.

IPC

Dades extretes de l'API de l'INE.

L'Índex de Preus de Consum (IPC) és un indicador estadístic que mesura l'evolució dels preus de béns i serveis representatius de consum familiar al municipi o regió. Això permet visualitzar com varia el cost de la vida al llarg del temps.

En aquesta aplicació, les dades es descarreguen directament des de l'API del Institut Nacional d'Estadística (INE), concretament de la sèrie de dades 251856, que correspon a l'IPC publicat per l'INE. La consulta recupera els valors més recents indicats per paràmetres.

Preu mitjà del lloguer

Aquestes dades han estat recopilades diàriament des del 14/08/2025 des del portal immobiliari privat Idealista. Representen el preu mitjà de l'oferta de lloguer a Girona per dia.

L'objectiu és permetre visualitzar l'evolució dels preus ofertats al llarg del temps, identificar tendències i facilitar la comprensió del mercat de lloguer a la ciutat.

L'Informe detallat

Aquesta funcionalitat és especialment útil perquè aporta informació desglossada sobre alguna de les variables. Per exemple, els informes detallats sobre despesa d'ajuts econòmics aprovats mostren:

- Un resum de les dades que es mostren.
- Informació global: el total de despesa a la ciutat en ajuts econòmics d'urgència en el període seleccionat.
- Un desglossament de l'import per tipologia d'ajuts on consta: total d'ajuts, valor màxim, valor mínim, la mitjana i la desviació típica.
- Una gràfica del total d'ajuts per cada tipologia, amb l'import i el percentatge que representa sobre el total.
- La distribució dels ajuts per gènere de les persones sol·licitants i tipologia d'ajut.
- La distribució dels ajuts per sector de Serveis Socials de la ciutat.

Ajuts per tipus

Informe generat el dia '28-04-2026'.

De què ens informen aquestes dades?

Informe sobre ajudes econòmiques agrupades per tipus. S'han considerat les ajudes concedides entre les dates 01-01-2026 i 27-04-2026.

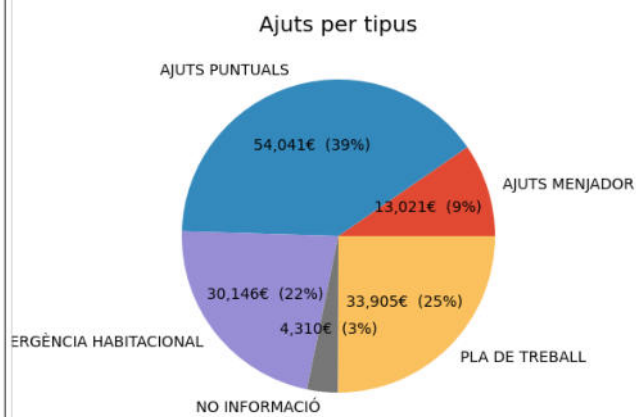
Informació global

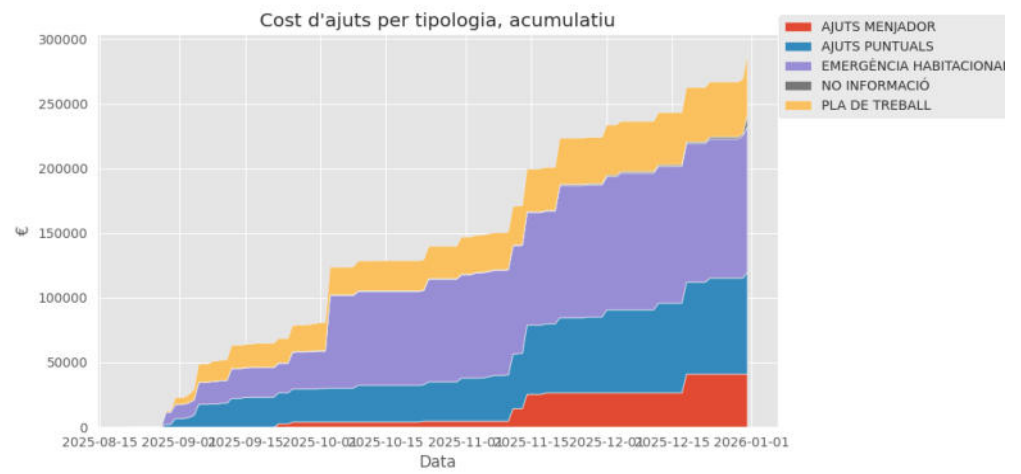
En aquesta secció apareix informació general dels ajuts, segons la seva tipologia, quantitat i cost.

En total ens hem gastat 135425 66€ en ajuts. Tot seguit es desglossa aquesta quantitat segons tipologia d'ajut.

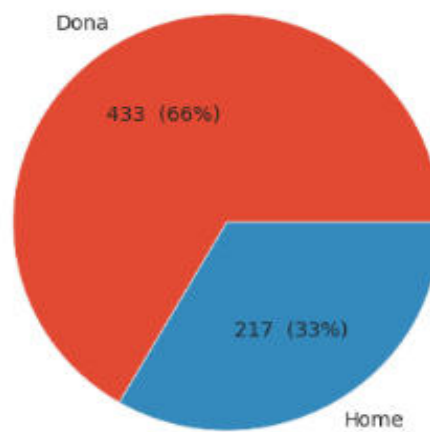
Tipus d'ajut	Total	Màxim	Minim	Mitjana	Desv. Típica
AJUTS MENJADOR	13021.16	1352.2	23.57	245.68	262.52
AJUTS PUNTUALS	54041.57	1633.5	30.0	274.32	178.27
EMERGÈNCIA HABITACIONAL	30146.23	2391.62	0.5	548.11	751.63
NO INFORMACIÓ	4311.0	550.0	13.5	226.89	112.68
PLA DE TREBALL	33905.7	800.0	60.0	287.34	87.6

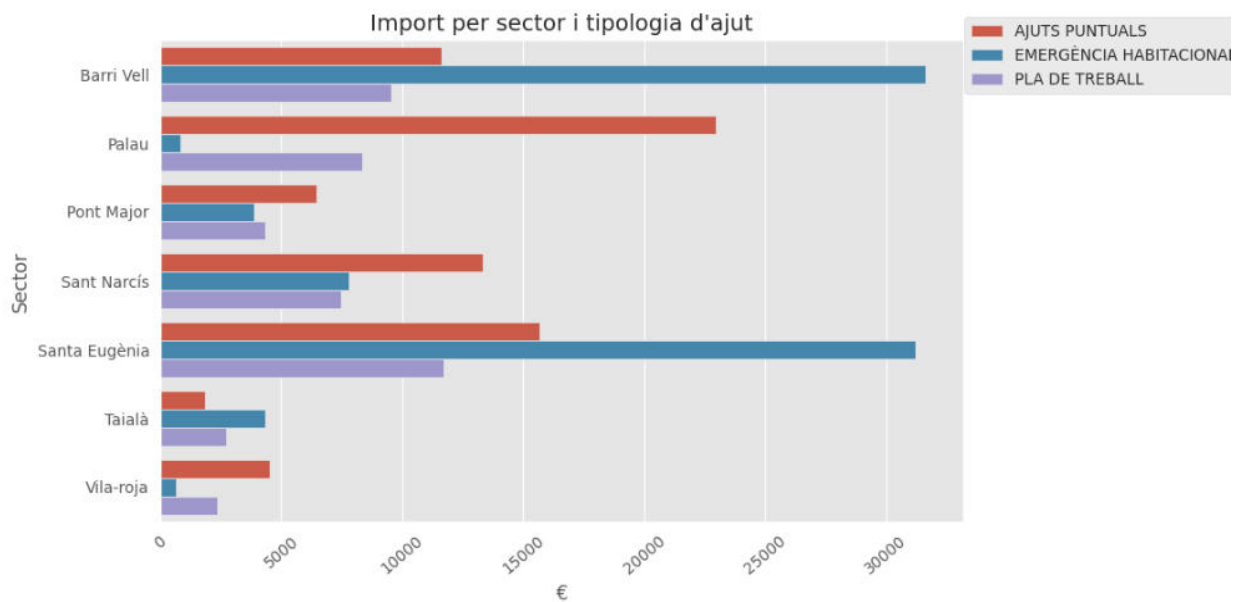
PLA DE TREBALL	33905.7	800.0	60.0	287.34	87.6
----------------	---------	-------	------	--------	------





Quantitat d'ajudes concedides per gènere





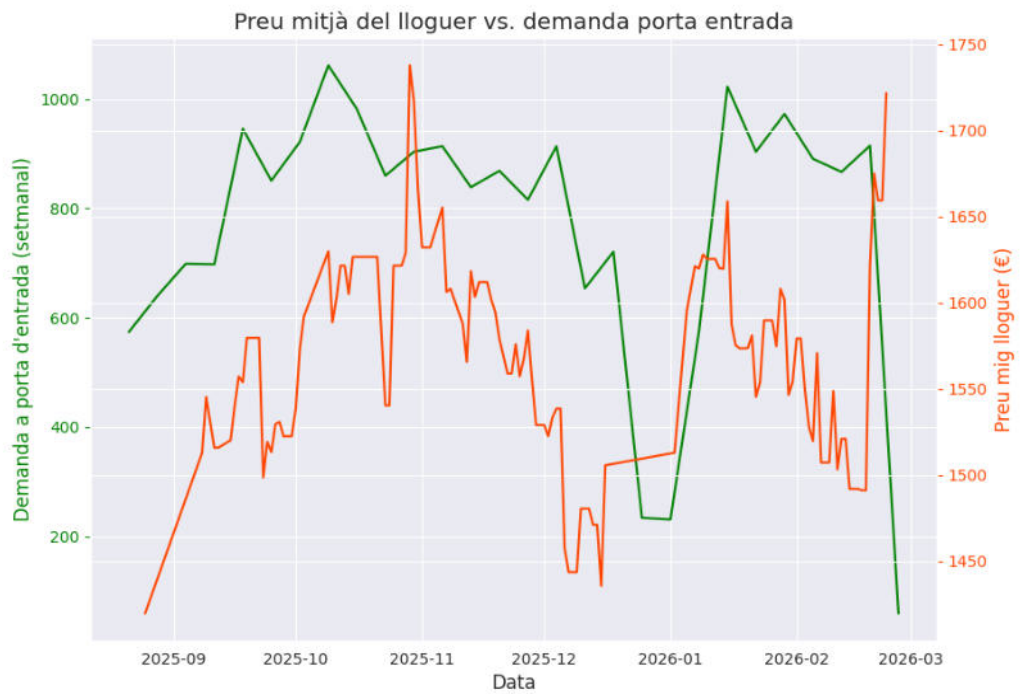
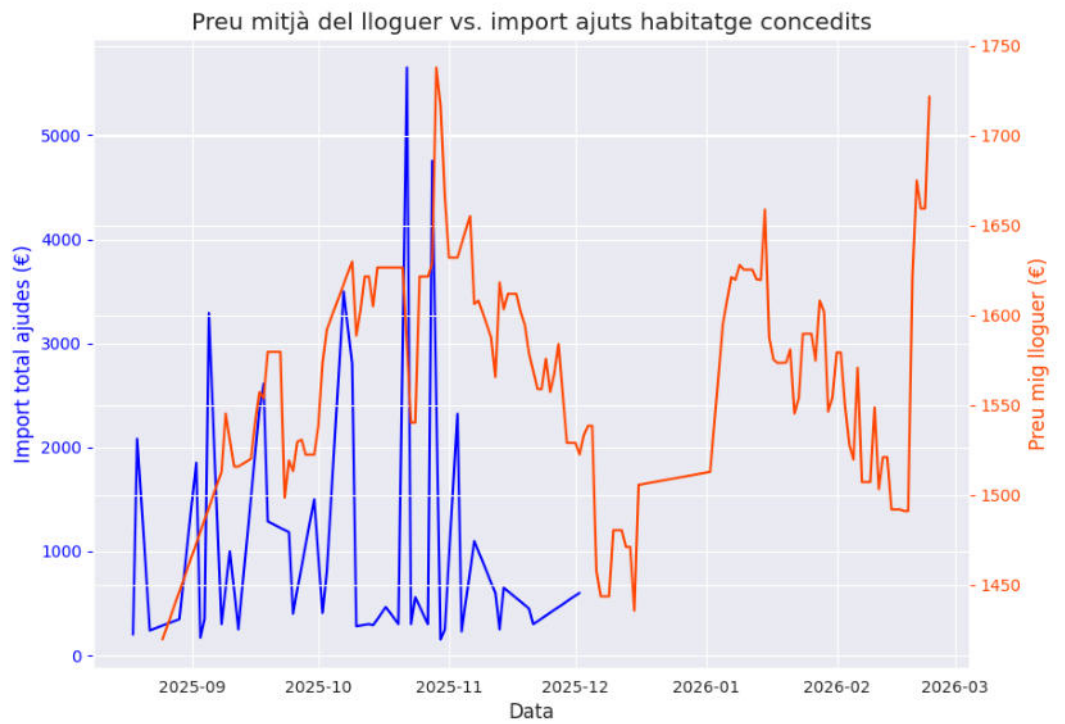
Al mateix temps, els informes detallats sobre el Padró Municipal mostren:

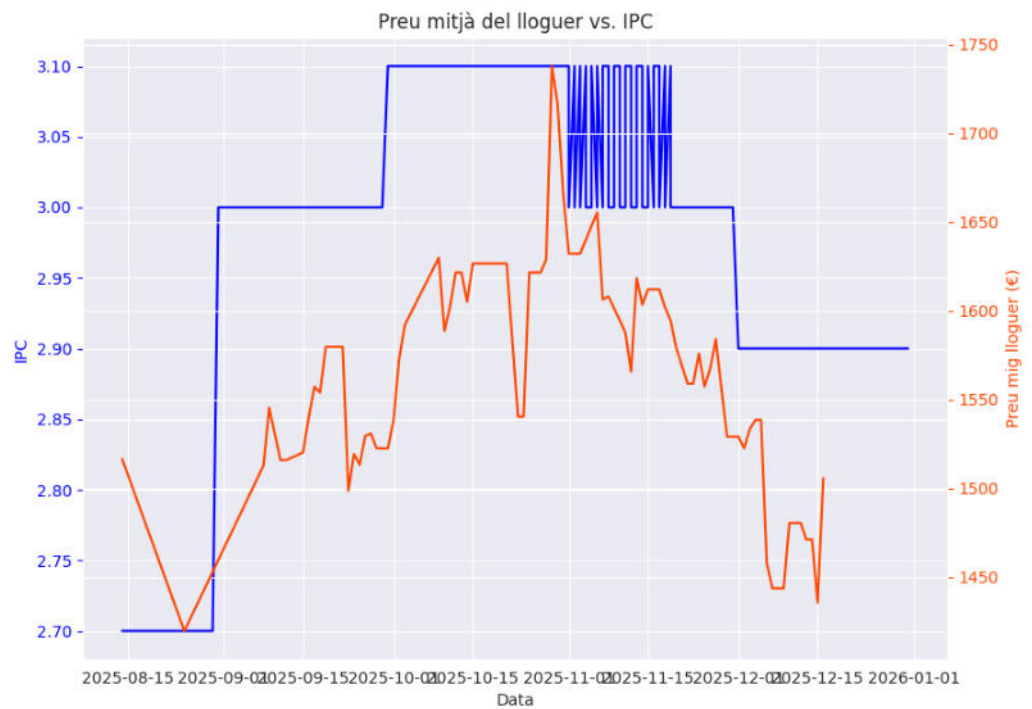
- La informació més rellevant de l'evolució del padró en el període temporal seleccionat,
- numèricament i en gràfica.
- Una gràfica que mostra l'evolució del Padró per cada sector dels Serveis Socials de la ciutat.
- Informació del nivell educatiu general dels ciutadans de Girona.

L'informe, per tant, retorna informació sobre la variable seleccionada i, en alguns casos, de la relació d'aquesta amb altres variables que formen el Quadre de Comandament quan és d'interès la comparació evolutiva d'aquestes variables.

L'informe sobre evolució del lloguer a la ciutat de Girona retorna la informació més rellevant de l'evolució del preu del lloguer de nova oferta a la ciutat de Girona entre les dates seleccionades:

- Informació sobre el preu mitjà i l'import màxim i mínim del preu del lloguer.
- Gràfica de l'evolució del preu del lloguer en les dates seleccionades.
- Gràfica de l'evolució del preu de lloguer per sector de Serveis Socials de la ciutat (quan aquesta dades està disponible)
- Gràfica comparativa de l'evolució del preu del lloguer amb els ajuts econòmics concedits durant el mateix període.
- Gràfica comparativa de l'evolució del preu del lloguer amb l'evolució de la demanda d'atenció als Serveis Socials.
- Gràfica de l'evolució del preu mitjà del lloguer vs IPC.





Descàrrega de dades

Aquesta funcionalitat aporta les dades en format excel. En el cas dels ajuts econòmics, el full de càlcul mostra totes aquelles dades que utilitza el SIAGI relatives als ajuts econòmics, per la qual cosa, aquesta funcionalitat és molt útil tant per estudis concrets sobre la gestió dels ajuts econòmics, com per la realització de memòries o informes específics.

Pel que fa a les dades de Padró, les dades estan anonimitzades per tal de respectar els principis de seguretat en els tractament de dades personals, però es poden agrupar en el nivell més baix possible que no compromet aquest principi. Després de valorar diferents possibilitats, s'arriba a l'acord que l'element màxim de granulitat sigui l'illa de cases, atès que això permet estudiar l'evolució del padró d'acord amb les mesures que conté la AIPD (anàlisi d'impacte de protecció de dades) i ofereix un nivell d'aproximació òptim per estudiar la distribució territorial de les dades.

3.4.2 Quadre d'Avisos

L'aplicació del Quadre d'avisos és un mecanisme que permet informar de col·lectius o situacions de risc detectades a la ciutat, en temps real.

Arquitectura tecnològica

Els avisos es computen amb diferents rutines temporals (diària, setmanal, mensual i anual) mitjançant processos ETL específics que permeten extreure, transformar i carregar la informació procedent de diverses fonts de dades. Aquests processos estan orquestrats i gestionats amb Apache Airflow, que permet planificar, monitorar i controlar l'execució de les tasques de manera automatitzada i fiable.

Els scripts ETL s'encarreguen de consolidar la informació, aplicar les transformacions necessàries (neteja de dades, agregacions, càlcul d'indicadors, validacions, etc.) i generar els resultats estructurats segons les necessitats del sistema. Aquest enfocament garanteix la coherència, la traçabilitat i la qualitat de les dades abans que siguin explotades.

Un cop processada, tota aquesta informació s'emmagatzema en taules específiques de PostgreSQL, dissenyades per optimitzar la consulta i el rendiment en entorns d'explotació analítica a Apex. Aquestes taules actuen com a capa de persistència intermèdia entre els processos de càlcul i la capa de visualització.

Finalment, el frontend consumeix aquestes dades a través de consultes o serveis intermedis, permetent representar els avisos mitjançant gràfiques, indicadors i altres elements visuals del quadre d'avisos. Aquesta arquitectura desacobra els processos de càlcul del sistema de visualització, assegurant escalabilitat, manteniment i eficiència en el tractament de la informació.

Què mostra el Quadre d'avisos

La pantalla d'avisos és un sistema d'alertes vinculades a determinades problemàtiques que s'aborden des dels Serveis Socials. Aquesta funcionalitat permet disposar d'una visió proactiva, ja que l'eina Pigall anticipa situacions que es poden donar i que afecten a les persones ateses als Serveis Socials fins al moment.

Els avisos creats fins al moment es poden dividir en tres tipologies depenent de la informació que ens aporta l'alerta:

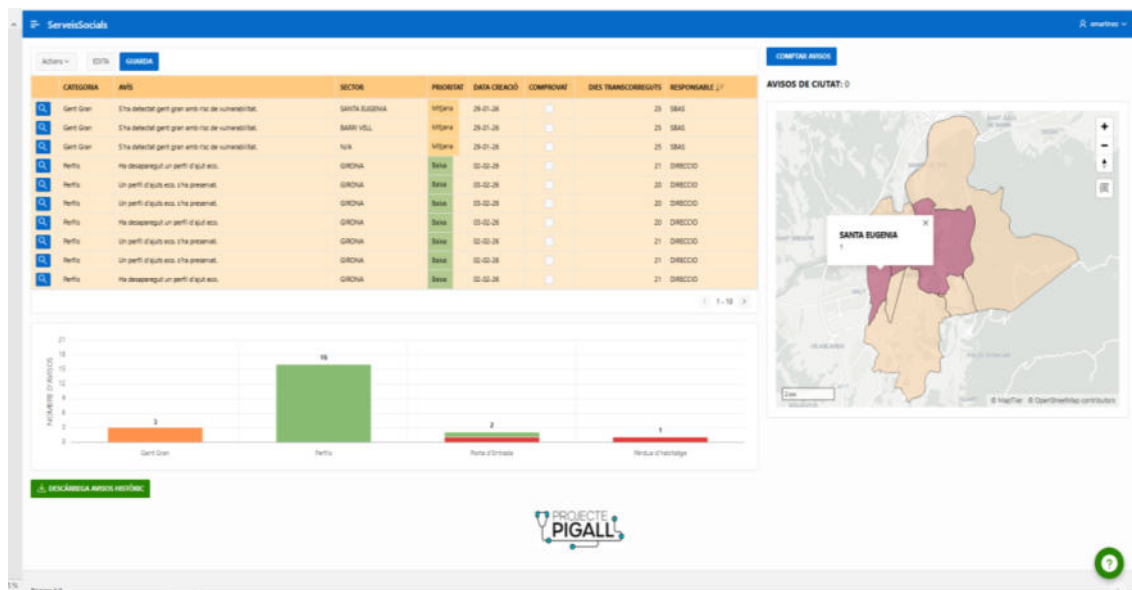
- Avisos de canvi de perfil (Canvi en el perfil de les persones beneficiàries d'ajuts econòmic) derivades del Clústering.
- Canvis de tendència detectats (Augment o disminució detectada en el nombre de sol·licituds d'atenció social rebudes al Servei Porta d'Entrada).
- Alerta sobre situacions de risc que poden afectar les persones ateses.
 - 1) risc de pèrdua d'habitatge
 - 2) vulnerabilitat en gent gran

Els avisos que ens alerten sobre situacions que poden afectar a persones ateses s'han creat a partir de la identificació dels paràmetres que configuren determinades situacions de risc.

Aquesta identificació s'ha creat a partir de l'expertesa dels i de les tècniques implicades (SBAS, Equip d'Atenció a la Gent Gran, Equip d'Emergència Habitacional, etc.).

Els avisos es creen a partir de la revisió periòdica i automatitzada de les dades que configuren els paràmetres. Els avisos dissenyats fins al moment tenen dos temporalitats: anuals, quan alerten de canvis que afecten a un col·lectiu concret (beneficiaris d'ajuts) o diaris, quan els avisos informen de possibles situacions de risc que poden afectar persones ateses.

La funcionalitat dels avisos respecte persones ateses per serveis socials, està preparada per facilitar la comunicació directa amb els tècnics referents de cada cas, ja que cada avís pot generar automàticament un missatge que s'envia automàticament al tècnic referent de cas.



La tipologia d'avisos s'estructura de la següent manera:

- Ajuts urgència social (AUS)

Pigall analitza les característiques perfil de les persones beneficiàries d'ajuts econòmics d'urgència. L'avís sobre perfils ens informa de com s'agrupen els demandants d'ajuts durant l'any i ens alerta sobre els canvis de perfil que s'han donat durant aquest període. Aquesta funcionalitat permet detectar si s'ha donat algun canvi en el perfil de persones demandants durant l'any anterior.

- Vulnerabilitat de gent gran

El sistema detecta quan una persona reuneix unes característiques que, a priori, determinen que pot trobar-se en situació de risc. Les variables que es tenen en compte són:

- Té més de 75 anys o més.
- Viu sola o en parella.
- No consta que tingui cap persona referent.
- No disposa de xarxa de suport identificada.
- Disposa o no de servei de teleassistència.
- Fa entres 3 i 5 mesos que no es disposa d'informació de la persona.

A partir de la combinació d'aquestes variables ens alerta de les persones que poden estar en

situació de vulnerabilitat, quan concorren aspectes com: viuen soles, no tenen xarxa, fa més de 5 mesos sense contacte amb serveis socials, per exemple.

Els avisos s'actualitzen diàriament.

En el model d'informe que genera cada avís es pot arribar a mostrar el codi identificador (ID) de la persona que assenyalava l'avís, per tal de poder contactar amb els professionals de referència o en tot cas, assignar-lo per fer-ne el seguiment. A més, el sistema envia un correu automàtic al professional de referència, notificant l'avís.

Avis de risc de vulnerabilitat en persones grans al sector de Vila-roja

Informe generat el 20-04-2020

ID avís: 12883

Categoria: 000Vulnerabilitat

Què ens mostra aquest avís?

Llistat de persones grans que són susceptibles a ser vulnerables.

Període temporal al que correspon l'avís

Predictió feta el dia 2020-04-16.

Dades significatives de l'avís

S'ha detectat 1 persona que podria trobar-se en situació de vulnerabilitat, ja que compleix els següents criteris:

- Fa entre 3 i 5 mesos que no es disposa d'informació recient de la persona.
- Té 75 anys o més.
- Té contractat el servei de teleassistència.
- Viu sola, o en parella.
- No consta cap persona referent.
- No disposa de cap xarxa de suport identificada.

La següent taula mostra algunes característiques d'aquesta persona:

ID Usuari	Gènere	Sector
16507	D	Vila-roja

- Risc de pèrdua d'habitatge

L'avís de risc de pèrdua d'habitatge mostra totes aquelles famílies que el model ha identificat com a susceptibles de perdre el seu habitatge. Els avisos de pèrdua d'habitatge es creen a partir de la identificació resultant del model entrenat. Aquest model s'entrena a partir de l'històric de persones que han demanat ajuts i que han sigut o no desnonades.

Les característiques de les famílies amb les quals ha sigut entrenat el model té en compte aquestes característiques del registre al SIAGI:

- Tipus d'unitat familiar (persona sola, família nuclear, família nombrosa, família monoparental).
- Gènere de la persona.
- Gènere i franja d'edat de les persones membre.
- Situació administrativa de residència.
- Situació econòmica i tipus de deute.
- Situació laboral.
- Cobertura econòmica.
- Ingressos totals de la unitat.

- Despeses totals de la unitat.
- Import dels ajuts concedits.
- Ajuts concedits i tipus (subsistència, habitatge, altres).
- Règim de tinença de l'habitatge.
- Tipus de xarxa de suport.

L'avís es genera sobre les persones que han sol·licitat un ajut econòmic d'urgència entre els 3 i 5 mesos anteriors i ens alerta si s'ha detectat que, ateses les seves característiques i en base a l'històric de casos similars, pot estar en risc de pèrdua d'habitatge. Aquest avís complementa per tant, el diagnòstic que realitza el tècnic que està atenent el cas, la qual cosa permet anticipar possibles actuacions per prevenir la pèrdua d'habitatge.

En el model d'informe que genera cada avís es pot mostrar el codi identificador de cada nucli familiar en risc per tal de poder contactar amb els professionals de referència, així com informació referent als membres que conviuen en el nucli i el sector en que es troba.

- Demandes d'atenció al Servei Porta d'Entrada

L'avís de demandes d'atenció es crea quan l'algorisme detecta un canvi de tendència en el volum de demandes. Aquest algorisme segueix les bases de l'aprenentatge automàtic en temps real, on es té en consideració la detecció del *drift* per tal de captar una nova realitat en les dades.

Pigall genera un avís quan detecta que hi ha un canvi de tendència significatiu en nombre de demandes al Servei Porta d'Entrada previstes, ja sigui per augment o per disminució. Aquest avís permet adaptar l'organització del servei de Porta d'Entrada a la demanda prevista.

En el model d'informe es pot observar l'evolució de la demanda en els últims mesos fins a l'actualitat, i es mostra el punt exacte corresponent a la data on hi ha hagut el canvi de tendència.

3.4.3 Escenaris de Futur

Arquitectura tecnològica

L'aplicació Escenaris de futur es construeix a partir de formularis interactius que permeten definir i ajustar diferents variables associades a cada casuística. Aquests formularis ofereixen la possibilitat de modificar paràmetres rellevants, com ara perfils, condicions específiques o valors quantitatius amb l'objectiu de simular diferents situacions i analitzar-ne l'impacte potencial.

Un cop l'usuari ha completat el formulari, la informació introduïda es processa i s'envia al backend mitjançant codi **PL/SQL** o bé a través de *REST Data Sources* d'Oracle APEX. En ambdós casos, la comunicació es realitza utilitzant l'API desenvolupada específicament per a aquest propòsit.

Aquesta API actua com a capa d'integració entre el frontend i els models predictius, permetent fer ús dels algorismes i models desenvolupats prèviament per a l'anàlisi i la simulació. A partir de les dades rebudes, el sistema executa els càlculs corresponents i genera una resposta amb els resultats de la predicció.

Finalment, el frontend rep aquesta resposta i la representa de manera visual mitjançant gràfiques, indicadors o altres elements interactius, facilitant així la interpretació dels resultats i la presa de decisions basada en dades. Aquest mecanisme permet oferir una eina flexible i dinàmica per a l'anàlisi prospectiva i l'avaluació d'escenaris futurs.

Què mostren els Escenaris de Futur

Els Escenaris de Futur són una eina que permet preveure com es veurà afectada la població atesa si es donen determinades circumstàncies. També permet simular com evolucionaran aspectes que tenen impacte en la gestió dels Serveis Socials.

Fins el moment, s'han desenvolupat dos simuladors d'escenaris de futur relacionats amb el risc de pèrdua d'habitatge i un sobre l'evolució del pressupost de l'Àrea de Serveis Socials destinat a ajuts d'urgència social:

Simulador de risc de pèrdua d'habitatge per cas

Aquest simulador prediu el risc de pèrdua d'habitatge en un cas particular si es donen certes circumstàncies socioeconòmiques. Per fer la simulació, el model predictiu té en compte l'històric de casos que han tingut un desnonament i el combina amb els casos que no han estat desnonat, tenint en compte totes les variables estadísticament significatives.

La funcionalitat d'aquests predictor és poder detectar, abans que es produeixi, el risc que una família té de perdre el seu habitatge. Per fer la predicció, s'introdueix el perfil de la família en base a aquests paràmetres:

- Informació del nucli familiar:
 - Tipus d'unitat familiar
 - Nombre de persones que formen la unitat familiar
 - Edats dels membres de la unitat familiar
 - Situació administrativa de residència

- Informació econòmica del nucli:

- Tipus de deute
- Tipus de cobertura econòmica
- Tipus de suport
- Règim de tinença de l'habitatge
- Ingressos
- Despeses

- Informació sobre el nombre, tipologia i imports d'ajuts demanats fins al moment

Amb aquestes variables, el model prediu si una unitat de convivència concreta pot patir una situació de pèrdua d'habitatge en el futur proper.

Simulador de risc de pèrdua d'habitatge global

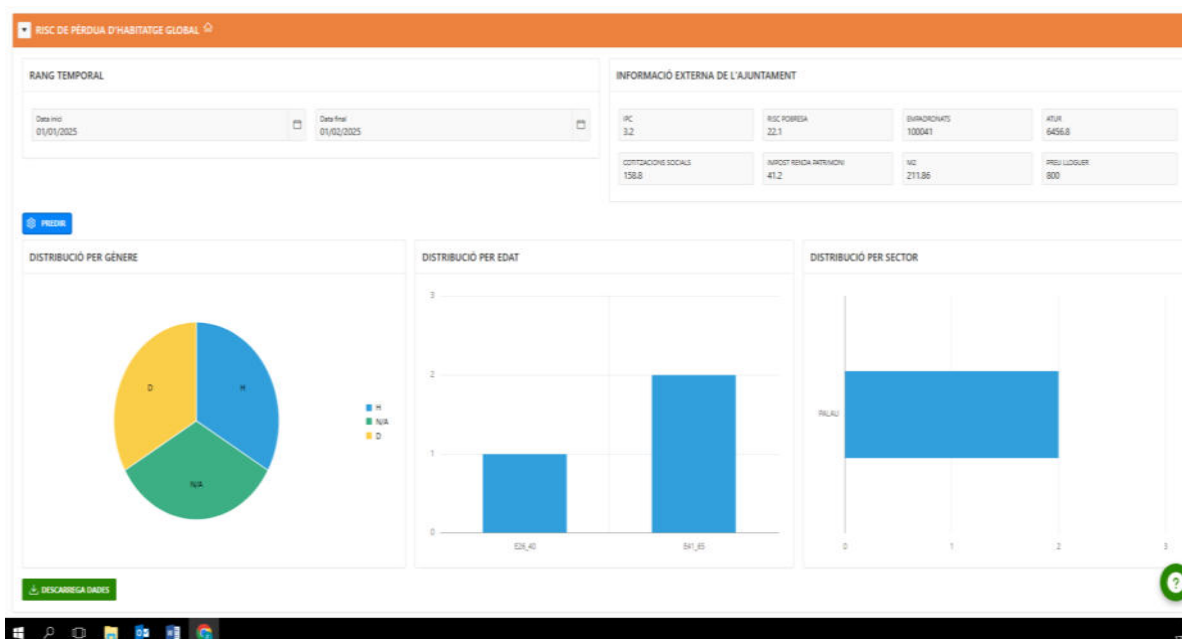
Aquest predictor combina les dades pròpies del SIAGI amb variables externes de context socioeconòmic, de tal manera que prediu quantes famílies poden estar en risc de desnonament si varien les condicions socioeconòmiques en el període temporal seleccionat.

Les variables externes que es tenen en compte per fer la predicció són:

- Índex de preus al consum
- Taxa de risc de pobresa o exclusió social. Taxa AROPE.
- Atur: mitjana anual de l'atur registrat a Girona.
- Índex de cotitzacions socials a Girona (font Observatori Municipal)
- Impost Renda Patrimoni
- Metres quadrats habitatges
- Preu mitjà del lloguer

Aquest escenari de futur, el model retorna la predicció dels casos detectats en risc de pèrdua d'habitatge en dos formats:

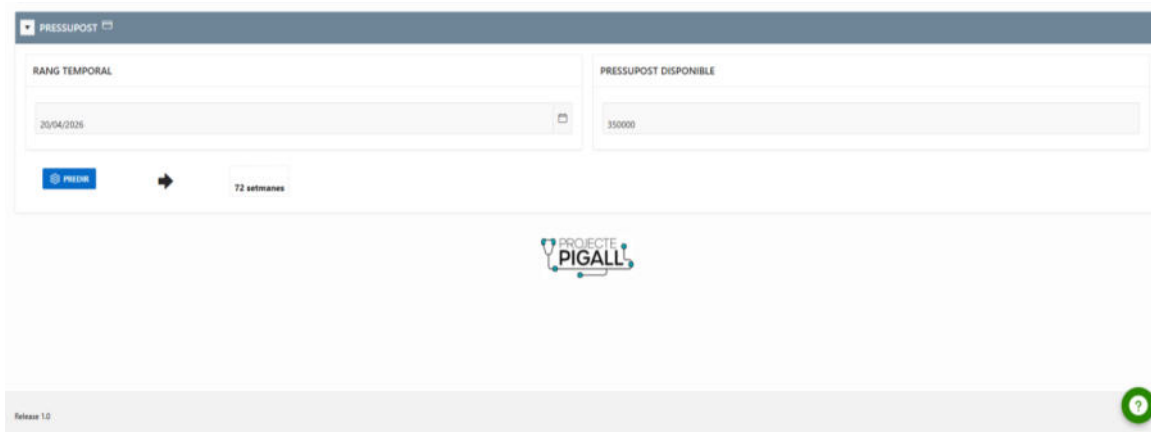
- En gràfiques de distribució per gènere, edat i sector de serveis socials.
- En format excel on consten totes les persones ateses que es poden veure afectades per risc de pèrdua d'habitatge si es donen les condicions socioeconòmiques seleccionades.



Pressupost

La despesa dels Ajuts d'Urgència social concedits es tramita de manera setmanal a l'Ajuntament de Girona, de manera que cada setmana es resolen una quantitat x d'ajuts per al seu corresponent pagament.

Aquest escenari de futur ens informa, a partir de la data seleccionada i el pressupost total disponible per ajuts econòmics d'urgència, de la quantitat de setmanes que es pot seguir atorgant ajuts d'acord amb el ritme de resolucions i despesa dutes a terme fins al moment.



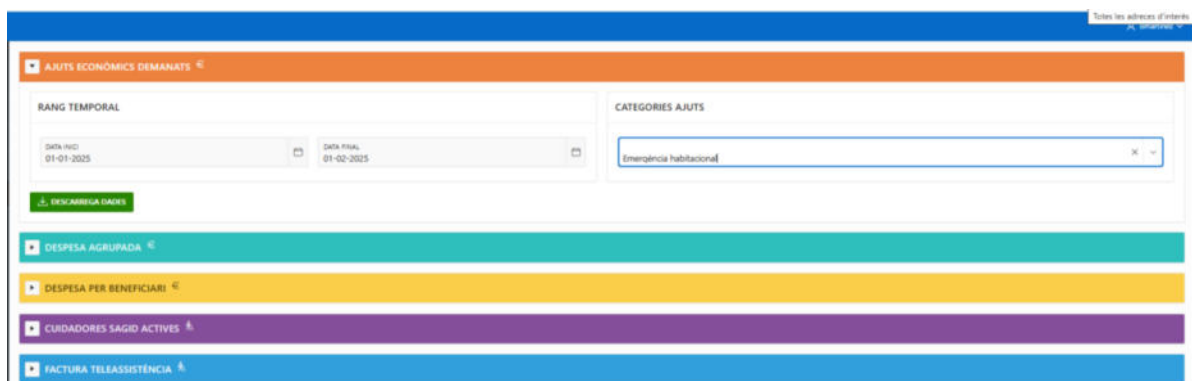
3.4.4 Eines de Consulta

El fet que el projecte Pigall es desenvolupa directament des de l'Ajuntament i els tecnòlegs s'han incorporat a l'equip de l'Àrea de Serveis Socials, permet una adaptació continua de l'eina a les necessitats dels professionals de l'àrea. Aquesta proximitat permet un alt grau d'agilitat per trobar solucions tecnològiques que permeten simplificar processos i tasques que anteriorment requerien una alta inversió de temps i esforç. És des d'aquesta òptica que Pigall incorpora una nova funcionalitat, no prevista a l'inici del projecte, d'eines de consulta, una miscel·lània de funcionalitats per agilitzar processos administratius i de gestió.

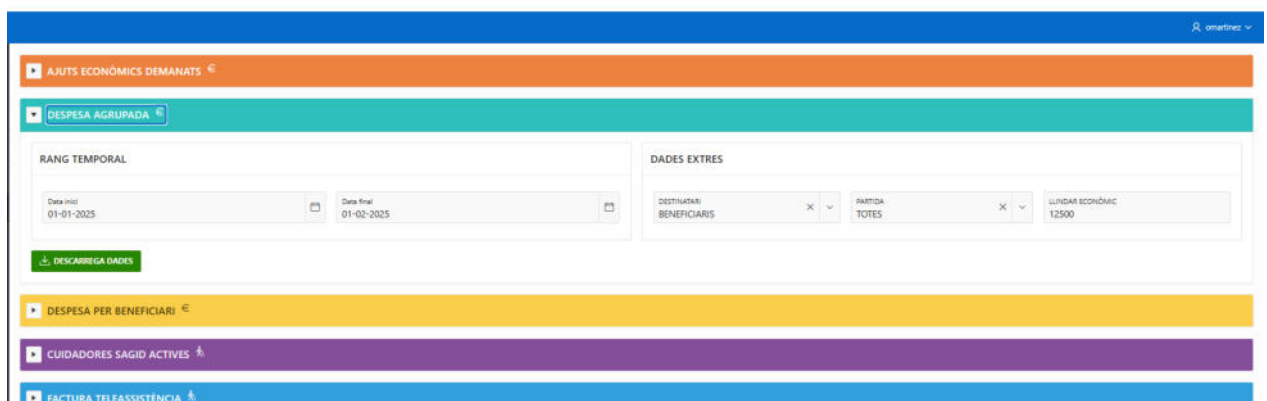
Aquestes funcionalitats són possibles perquè l'eina Pigall té accés directe a les diferents bases de dades que s'utilitzen a Serveis Socials i l'ajuntament, tant per la intervenció, com per la gestió pressupostària.

Actualment, les funcionalitats desenvolupades són les següents:

- **Ajuts econòmics demanats:** ens facilita un llistat de totes les persones que han rebut un ajut econòmic d'urgència en cada una de les seves modalitats i en el període temporal seleccionat.



- **Despesa agrupada:** aquesta funcionalitat permet identificar beneficiaris que han rebut un ajut econòmic d'urgència en un període determinat i per sobre del llindar econòmic seleccionat.



- **Despesa per beneficiari:** aquesta funcionalitat permet conèixer tots els ajuts que ha rebut una persona concreta durant el període temporal seleccionat.

- **Cuidadores SAGID Actives:** aquesta funcionalitat descarrega un llistat de totes les persones que consten al SIAGI com a cuidadores actives d'alguna persona amb dependència o discapacitat. Disposar d'aquests llistats de manera automàtica i actualitzats li és útil al Servei d'Atenció a la Gent Gran i Dependències per realitzar periòdicament accions grupals i de suport adreçades a les persones cuidadores.
- **Factura Teleassistència:** eina de control de la factura mensual que envia l'empresa que realitza el servei de Teleassistència. És útil per comprovar que la informació és correcta i confrontar-la amb la informació disponible al SIAGI sobre la gestió d'aquest servei.

3.5 TESTEIG I FORMACIÓ

Durant tot el procés de desenvolupament del projecte s'ha vetllat perquè els avenços tecnològics proposats estiguin alineats amb les necessitats de l'àrea de Serveis Socials i que els resultats obtinguts fossin coherents des de l'òptica social.

3.5.1 Testeig

En concret, el testeig de l'eina s'ha fet per part de diferents agents i amb diverses mirades:

1) Revisió i validació dels instruments i tècniques utilitzades per part de la Universitat de Girona.

La col·laboració amb la UdG ha estat a tres nivells:

- Assessorament tecnològic: facilitant instruments tecnològics, supervisant el desenvolupament dels models matemàtics, estadístics i d'aprenentatge automàtic.
- Coherència tècnico-social: assessorant de l'ús i possibilitats que tenien les dades necessàries per al funcionament de l'eina.
- Seguiment dels principis ètics en la utilització de la IA: a banda de les recomanacions rebudes durant tot el procés per seguir els principis ètics en l'ús de la tecnologia, l'eina ha estat avaluada seguint el Model PIO de l'OEIAC¹⁴.

2) Revisió i validació de l'eina per part dels tècnics socials de l'equip Pigall.

Els tècnics socials de l'equip Pigall han participat en cada una de les fases del desenvolupament del projecte, aportant la visió social en la metodologia emprada, els conceptes utilitzats i les solucions proposades. Aquesta participació ha permès donar coherència des del punt de vista social a l'eina i, al mateix temps, testar les funcionalitats de l'eina en els següents aspectes:

- Traducció de les necessitats socials detectades a possibles funcionalitats i solucions
- Explicabilitat i interpretació de la informació
- Nomenclatura i llenguatge de l'eina
- Experiència d'usuari: facilitat d'ús, accessibilitat, disseny visual de la interfície.
- Coherència i utilitat dels productes:
 - Quadre de comandament: valoració del funcionament i de la coherència dels informes que emet l'eina en relació a les diferents variables.
 - Avisos: comprovació de la veracitat i oportunitat dels avisos generats per l'eina
 - Escenaris de futur: coherència i utilitat dels resultats obtinguts en cada una de les prediccions.
- Proposta de millora i noves utilitats.

¹⁴ Veure Annex 5 Avaluació ètica i legal del Projecte Pigall a partir del Model PIO

La participació dels tècnics socials de l'equip Pigall en el testeig de l'eina s'ha donat en la fase de construcció d'aquesta, com a pas previ de validació abans d'iniciar la fase de testeig als usuaris finals de l'eina.

3) Revisió i validació de les funcionalitats de l'eina per part d'una directora de Equip Bàsic d'Atenció Social.

Atès que les directores dels Equips Bàsics d'Atenció Social (EBAS) són un dels agents principals als qui va destinada l'eina, un cop realitzada la validació per part dels tècnics socials de l'equip Pigall, s'ha comptat amb el suport d'una directora de SBAS per fer el testeig de l'eina. Aquesta figura aporta, per la seva experiència professional en els serveis bàsics i, en concret, en Servei de Porta d'Entrada, una visió àmplia de les necessitats dels serveis.

La seva visió és bàsica en la globalitat de l'eina i, especialment, en totes les funcionalitats i resultats que tenen a veure amb els ajuts econòmics d'urgència, el risc de pèrdua d'habitatge i els canvis de tendència de la demanda rebuda al Servei Porta d'Entrada.

4) Revisió i validació de les funcionalitats de l'eina per part de la responsable del Servei d'Atenció a Gent Gran i Dependències (SAGGID).

Atès que l'eina incorpora diferents funcionalitats que tenen a veure amb el risc de vulnerabilitat en la gent gran, s'ha comptat amb la participació de la responsable del Servei d'Atenció a la Gent Gran i Dependències a dos nivells: en la conceptualització de les variables utilitzades per definir el risc de vulnerabilitat en la gent gran i en el testeig de les funcionalitats relacionades amb aquesta temàtica.

En testeig realitzat per les dues la directora d'EBAS i del SAGGID s'ha fet al mateix nivell que el realitzat pels tècnics de l'equip Pigall: experiència d'usuari, validació i utilitat dels productes i propostes de millora i de noves utilitats

3.5.2 Formació

La formació en l'ús de l'eina Pigall es fa a dos nivells: eines de suport a l'usuari i formació bàsica sobre el funcionament de l'eina.

Eines de suport:

- Manual d'usuari de Pigall: s'ha elaborat un manual d'usuari per explicar cadascuna de les funcionalitats de l'eina.
- Cada una de les pantalles de l'eina, conté una icona d'informació, que conté una breu explicació sobre les variables emprades.
- Cada funcionalitat conté una icona d'informació sobre el seu funcionament.

Sessions de formació:

S'han realitzat diferents sessions de formació explicant el funcionament de l'eina Pigall amb els caps de secció de l'Àrea de Serveis Socials, i amb les direccions dels EBAS i els serveis especialitzats.

4. REPLICABILITAT DEL PROJECTE

En el procés de treball d'elaboració de l'eina Pigall, s'han fet evidents les seves potencialitats. Constitueix un model metodològic per al desenvolupament d'eines de planificació amb suport d'IA, que pot ser aplicable diversos àmbits de coneixement i de polítiques públiques. En aquest sentit, ha generat important interès, a nivell intern, per part de diferents regidories de l'Ajuntament de Girona, ja que pot ser aplicable a polítiques socials, educatives, ambientals, urbanístiques, econòmiques, etc.

Aquest projecte i en concret l'eina Pigall és referència en el camp de la **innovació** social, i ha estat desenvolupada des de l'administració pública, gràcies al treball col·laboratiu entre professionals del camp social i tecnòlegs contractats directament per l'Administració i assessorament ètic científic per part de la Universitat.

Aquest model de treball ha general interès i és referència a nivell extern per part d'altres administracions, així com els avaluadors dels fons NGEU pe que fa a la possible replicabilitat en altres administracions públiques, obrint una fase de treball de divulgació.

Com s'ha dit, la replicabilitat del projecte recau en la seva metodologia i funcionalitats. Aquestes queden recollides en la *Memòria de la Infraestructura Pigall* així com en el futur Informe de transparència algorítmica que serà publicat al portal de transparència de l'Ajuntament de Girona.

És important destacar que la replicabilitat del projecte serà diferent depenent de si s'amplia a altres àrees de l'Ajuntament de Girona o a una altra administració pública. Tenint en compte això, hem d'entendre que replicar el projecte a altres àrees del nostre Ajuntament significarà que la infraestructura tecnològica ja la tenim construïda i per tant, només serà necessari treballar sobre aquesta. Però si el que es vol és replicar el projecte a una altra administració, s'ha de tenir en compte que s'haurà de construir tota la infraestructura primer, amb la que es podrà treballar en fases posteriors.

5. CANVIS PRODUITS DURANT EL PROJECTE

El projecte Pigall s'ha vist fortament condicionat per la reducció del període d'execució, que ha fet reorientar alguns dels plantejaments inicials per tal d'adaptar els resultats esperats a la disponibilitat real de temps.

Si bé l'inici del projecte esta previst pel gener de 2023, a efectes pràctics, l'inici de l'execució del projecte Pigall es data el mes de juliol de 2024, amb la contractació de l'equip tècnic. Aquest any i mig de retard en l'inici del projecte, suposa la reducció en un 50% del temps previst per a la seva execució, fet que obliga a redefinir objectius, així com resultats previstos.

En data de 29 de juliol de 2022, l'Ajuntament de Girona presenta sol·licitud a la convocatòria de subvencions per la realització del "Projecte Pigall de millora, actualització i implantació de noves tecnologies en els serveis socials". El projecte presentat, esta previst ser desenvolupat en 3 anys, identificant 4 fases de treball.

Projecte Pigall	2023	2024	2025
FASE 1 revisió de conjunts de dades disponibles"			
FASE 2 integració de conjunts de dades a través de l'enllaç de dades administratives en combinació amb la resta de dades disponibles			
FASE 3 harmonització de les dades integrades a través de l'anàlisi i mitigació de biaixos			
FASE4 anàlisi descriptiva dels conjunts de dades i la seva monitorització			

L'any 2023 l'Ajuntament de Girona no disposa de l'equip d'especialistes necessari per elaborar un projecte tècnic d'aquesta magnitud i no pot dur a terme el procés selectiu laboral ordinari per contractar l'equip d'especialistes necessari, entre d'altres motius, pel canvi de govern.

En data 11 de gener 2024 el Departament de Drets Socials publica una resolució d'acceptació de sol·licituds de modificació de calendari d'execució per la qual s'autoritza a les entitats sol·licitants modificar el seu calendari d'execució i, per tant, les anualitats atorgades i a justificar concedides inicialment. En el cas de l'Ajuntament de Girona es manté la dotació i s'aplica als anys 2024 i 2025.

Per Decret de l'Alcaldia 2024017623 de data 01/07/2024, es fa l'aprovació interna del "Projecte Pigall de millora, actualització i implantació de noves tecnologies en els serveis socials", i s'inicia l'execució efectiva un cop fet el procés selectiu específic dels professionals adscrits al projecte Pigall.

En data 20/12/2024 es signa el Contracte de col·laboració científica amb la Universitat de Girona per la realització del projecte Pigall, si bé la col·laboració s'inicia des del 2022.

En conseqüència el nou calendari d'execució obliga a repensar el projecte inicial i a reformular el cronograma d'acord amb el que s'exposa a l'apartat 3. Metodologia Pigall.

6. ELEMENTS SATISFACTORIS, MILLORABLES I RECOMANACIONS PER AL DESPLEGAMENT

La creació de l'eina Pigall ha suposat la introducció d'una nova perspectiva i metodologia de treball a la gestió de serveis socials gràcies a les oportunitats que ofereixen les noves tecnologies i la Intel·ligència Artificial.

Aquesta eina té un seguit de fortaleeses, que la distingeixen d'altres projectes tecnològics portats a terme a l'Administració:

1. Projecte innovador en el camp de l'administració local i, en particular, en els Serveis Socials.
2. Dissenyat, creat i mantingut íntegrament amb recursos propis de l'Administració Pública, per tant fàcil de mantenir i escalar.
3. Segueix estrictament els principis d'IA responsable, gràcies al suport de l'OEIAC. L'eina és incapaç de prendre decisions automàtiques que puguin afectar a la vida, el benestar o la dignitat de les persones.
4. Projecte d'interès públic i fàcilment replicable a altres entitats o àrees de l'administració pública.
5. Sense dependències del sector privat. No hi ha ànim de lucre ni transferència de dades. És també un projecte immune a les eventualitats de la contractació pública, i es pot garantir una IA responsable amb models totalment transparents.
6. Es basa en el diàleg transversal entre professionals socials i tecnòlegs. Aquest diàleg col·laboratiu ha permès que les anàlisis quantitatives, qualitatives, i els discursos ètico-socials es complementin.
7. L'ús de models predictius permet fer polítiques proactives i no-reactives, permetent anticipar els problemes molt abans de que apareguin.

El projecte Pigall ha assolit de manera satisfactòria els objectius plantejats inicialment, consolidant-se com una eina estratègica per a la gestió dels Serveis Socials de l'Ajuntament de Girona. En l'àmbit de l'anàlisi, ha permès millorar significativament el coneixement de la realitat social del municipi, facilitant la identificació de problemàtiques a través de dades integrades i accessibles en temps real. Pel que fa a la planificació, l'eina ha contribuït a optimitzar la presa de decisions i l'assignació de recursos, reduint notablement el temps necessari per al disseny d'actuacions.

El fet que l'eina s'hagi construït a partir del diàleg constant entre tots els agents implicats, ha permès construir un instrument amb més funcionalitats que les inicialment previstes.

En relació amb la predicció, els models desenvolupats han demostrat un alt nivell de fiabilitat, amb una precisió del 94% en els models supervisats i bons resultats en els no supervisats. Això ha permès anticipar escenaris de risc de manera efectiva, amb deteccions recurrents al llarg de l'any i diferents horitzons temporals. Pel que fa a l'automatització, s'ha aconseguit una reducció molt significativa del temps dedicat a tasques mecàniques, passant de dies o setmanes a minuts, així com la implementació de múltiples processos automatitzats que milloren l'eficiència operativa.

Entre les principals fites assolides, destaca la creació d'un sistema integral que combina visualització de dades, generació d'aviso, models predictius i eines de suport a la gestió, tot desenvolupat amb recursos propis i amb una forta coordinació entre perfils tècnics, socials i

acadèmics. Aquest enfocament col·laboratiu ha estat clau per garantir l'adequació de l'eina a les necessitats reals i la seva adopció per part dels professionals, reflectida en un ús diari i un alt grau de satisfacció global.

No obstant això, el projecte també presenta alguns desafiaments. Entre aquests, destaca la necessitat de continuar ampliant i millorant la qualitat de les dades disponibles, així com garantir la constant actualització dels models predictius per mantenir-ne la precisió davant canvis socials i econòmics. També es planteja el repte de continuar formant els usuaris i fomentar l'ús avançat de l'eina, així com assegurar la seva escalabilitat i transferibilitat a altres àmbits municipals.

Finalment, com a àrees de millora, es proposa ampliar el nombre d'indicadors i fonts de dades, incorporar nous models predictius orientats a altres problemàtiques socials, millorar la personalització dels quadres de comandament i reforçar els mecanismes d'avaluació contínua de l'impacte de l'eina. Així mateix, es considera rellevant continuar aprofundint en els aspectes ètics i de protecció de dades, per garantir un ús responsable i transparent de la intel·ligència artificial en l'àmbit públic.

7. IMPACTES

7.1 IMPACTES INTERNS

1) Millora en la planificació i gestió dels serveis socials

A nivell d'impactes, el Projecte Pigall ha suposat una transformació significativa en la manera com els serveis socials municipals planifiquen, gestionen i anticipen les necessitats de la ciutadania. La incorporació d'eines d'anàlisi de dades i intel·ligència artificial ha permès disposar d'informació actualitzada en temps real sobre la realitat social del municipi, fet que facilita una presa de decisions més rigorosa i basada en evidències. Al mateix temps, l'automatització de tasques administratives i de generació d'informes ha reduït càrregues de treball internes i ha millorat l'eficiència dels equips professionals.

Un altre impacte rellevant és la capacitat predictiva del sistema, que permet detectar situacions de risc social —com possibles pèrdues d'habitatge o vulnerabilitat en persones grans— abans que esdevinguin problemàtiques greus, afavorint així polítiques públiques més preventives i proactives. Finalment, el projecte també ha reforçat la col·laboració entre perfils socials i tecnològics dins l'administració, consolidant un model innovador, replicable i alineat amb criteris d'ètica pública i ús responsable de la tecnologia.

2) Interès d'altres àrees de l'Ajuntament en les potencialitats de l'eina Pigall

La metodologia i les eines desenvolupades són aplicables a altres àrees de gestió municipal. En aquest sentit, diverses àrees de l'Ajuntament han mostrat el seu interès en el Projecte Pigall i han identificat necessitats a les que l'eina podria donar resposta.

3) Presa de consciència de la necessitat de millorar el registre per part dels professionals:

La creació de l'eina Pigall ha ajudat a visualitzar les potencialitats que té un registre acurat i complert de les dades de la intervenció social, la qual cosa té un impacte positiu en la presa de consciència dels professionals dels serveis sobre la necessitat de millorar el registre de les dades.

Per altra banda, el tractament de les dades realitzats per dissenyar el Pigall, ha permès identificar millores en el sistema de registre de la intervenció social. En aquest sentit, s'evidencia la necessitat de disposar d'un nou sistema de registre de la intervenció, buscant l'adaptabilitat al W-Social que impulsa el Departament de Drets Socials i Inclusió.

4) Elaboració de l'Anàlisi d'Impacte de Protecció de Dades i coordinació amb la Delegada de Protecció de dades per a la seva realització.

S'ha dut a terme un treball de col·laboració i definició de mesures de protecció de dades per tal que l'eina Pigall respongui als requeriments i normativa vigent, i alhora doni resposta a les necessitats analítiques i diagnòstiques que es proposa.

Aquestes experiències de treball col·laboratiu intern a l'Ajuntament generen nous reptes relacionats precisament amb la innovació que suposen, i que porten a la recerca de solucions creatives i adequades.

En aquest procés, també s'ha modificat els documents interns d'alta de serveis socials i els consentiments informats per adaptar-los al compliment de la protecció de dades en els usos de l'eina pigall.

5) L'avaluació ètica i legal del Projecte Pigall a partir del Model PIO.

L'Observatori d'Ètica en Intel·ligència Artificial de Catalunya ha realitzat l'avaluació ètica i legal del Projecte Pigall a partir del Model PIO. A l'informe de l'avaluació es conclou que el Projecte Pigall no només compleix amb els estàndards exigits pel Model PIO (que recull estàndards internacionals i requisits legals), sinó que en molts aspectes també els supera i, per tant, demostra que és possible innovar en l'àmbit social amb aplicacions d'IA que han estat dissenyades de forma ètica, legal i centrada en les persones.

Com a punts forts del projecte destaca:

- Governança participativa: La col·laboració social-tecnològica-ètica és el seu actiu principal.
- Transparència i explicabilitat: Documentació exhaustiva i comunicació honesta de limitacions.
- Autonomia humana: Dissenyat com a eina de suport, no de substitució, potenciant la feina social.
- Privacitat i protecció de dades: Tractament rigorós de les dades personals i utilització de salvaguardes, incloent-hi la realització d'una AIPD.
- Sostenibilitat institucional: Desenvolupament amb recursos propis, sense dependències privades.

I com a àrees de millora i reptes identifica:

- Accés a dades externes: La dependència d'APIs d'altres administracions limita l'ampliació de l'eina.
- Col·laboració amb la representació laboral: Incloure activament els sindicats en el procés per gestionar l'impacte en les condicions de treball.
- Formació continuada: La formació inicial és un bon punt de partida, però cal un pla de formació contínua per a tots els usuaris potencials per assegurar una adopció òptima.
- Monitorització post-desplegament: Cal planificar com es farà el seguiment de l'ús, la precisió dels models en el temps (per evitar el que es coneix com a drift) i la satisfacció dels usuaris després de la implementació completa.
- Avaluació d'impacte ambiental: Incorporar l'anàlisi de la petjada de carboni del projecte quan estigui en ple funcionament.

7.2 IMPACTES EXTERNS I DIVULGACIÓ

Jornades “Experiències d’aplicació de la intel·ligència artificial en la millora dels serveis socials”

El 27 de novembre de 2025, l’Ajuntament de Girona va organitzar una jornada per compartir experiències d’implementació de la intel·ligència artificial i les noves tecnologies als Serveis Socials. L’acte va consistir en dues ponències i una taula d’experiències per compartir experiències de projectes desenvolupats per administracions locals i un col·loqui final.

Les ponències van ser:

- La implementació de la IA en el sector social de Catalunya, a càrrec de Jordi Amat Batalla, director de l’àrea de Tecnologies i les Comunicacions del Departament de Drets Socials i Inclusió de la Generalitat.
- Serveis Socials i IA des d’una aproximació sociotècnica i ètica, a càrrec l’Albert Sabater, professor Serra Húnter i Director de la Càtedra OEIAC de la Universitat de Girona.



A la taula d’experiències van participar:

- L’Institut Municipal de Serveis Socials de l’Ajuntament de Barcelona, que van presentar la iniciativa “Usos d’IA en Serveis Socials d’atenció al sensellarisme de Barcelona”.
- El Consorci d’Acció Social de la Garrotxa, amb la presentació de “L’eina predictiva de situacions de vulnerabilitat a la Garrotxa”
- Osona, Acció Social, que van presentar “Portal Social, ecosistema de transformació digital a Osona Acció Social”
- Equip Pigall de l’Ajuntament de Girona, que van presentar “projecte Pigall, eina de planificació i suport a la presa de decisions”.

La jornada va finalitzar amb un col·loqui amb els participants que havien presentar un projecte al voltant dels reptes i oportunitats que ofereix la IA i les noves tecnologies pels serveis socials.



Xerrada per donar a conèixer l'eina Pigall a l'alumnat del Màster en Ciència de Dades de la Universitat de Girona.

L'Equip Pigall va ser convidat per la Universitat de Girona per presentar el Projecte Pigall a l'alumnat del màster en Ciència de Dades de la UdG. El propòsit d'aquesta activitat va ser oferir als futurs professionals de la ciència i l'enginyeria de dades una visió completa del cicle de vida d'un projecte de *Business Intelligence* desenvolupat dins de l'administració pública: des de la seva concepció a la seva implementació i documentació posterior. També es van explicar les maneres que aquesta eina d'intel·ligència artificial es fa servir per treballar en benefici de la nostra ciutat.

Participació en el Mobile World Congress Barcelona 2026

La Secretaria de Polítiques Digitals de la Generalitat de Catalunya va convidar l'Ajuntament de Girona a presentar el projecte Pigall en l'edició del 2026 del Mobile World Congress. Durant aquesta participació es va poder presentar l'eina a un representant de la Secretaria i intercanviar experiències amb altres participants del congrés.

https://web.girona.cat/relacionsexternes/noticies/-/asset_publisher/eU0FC7fioS14/content/l-ajuntament-de-girona-presenta-l-eina-pigall-al-mobile-world-congress



Publicacions

S'ha col·laborat en la redacció de l'article "Modelling Together for AI-Based Social Services", que s'ha publicat com a capítol del llibre "Participatory Modelling an Simulation to Improve AI-based Public Social Services", editat per Springer Nature. En aquest article s'explora la integració de tecnologies algorísmiques a la provisió de serveis socials a tot el conjunt de l'estat Espanyol.

ISBN: 978-3-032-15282-4

ANNEXOS

ANNEX 1

Glossari de termes

En aquest annex es recull un petit glossari definint alguns dels termes relacionats amb aquest projecte més utilitzats en els àmbits social i tecnològic.

Àmbit social

Aquests termes pertanyen al vocabulari habitualment utilitzat pels tècnics socials i persones que es dediquen a l'atenció.

- **SIAGI Servei Socials:** l'acrònim SIAGI significa Serveis Informàtics Ajuntament de Girona. SIAGI té diferents aplicatius per cada àrea de gestió de l'Ajuntament. El SIAGI de Serveis Socials és l'eina de registre de les intervencions que realitzen els serveis d'atenció social de l'àrea de Servei a les Persones (Serveis Socials).
- **Usuari/a:** persona atesa per Serveis Socials.
- **Nucli familiar:** persones amb relació de parentesc que conviu en un mateix domicili.
- **Tipus de família:** es distingeix entre:
 - **Família nuclear:** unitat de convivència formada per progenitors i descendents
 - **Família extensa:** unitat de convivència formada per familiars de diferents graus de consanguinitat.
 - **Persona sola**
 - **Parella**
 - **Altres:** qualsevol nucli familiar amb una configuració diferent a les anteriors.
- **Nucli de convivència:** persones amb relació de parentesc o no que conviu en un mateix domicili.
- **Família monoparental:** família formada per un sol progenitor que conviu amb un o més fills i n'assumeix en solitari la cura i la responsabilitat
- **Àmbits:** cadascun dels espais de la vida d'una persona que influeixen en el seu desenvolupament. Es distingeixen diferents àmbits bastats en el model d'inclusió:
 - Àmbit econòmic
 - Àmbit d'habitatge
 - Àmbit escolar/formatiu
 - Àmbit de salut física i mental
 - Àmbit relacional
- **Valoració (de l'àmbit):** es distingeixen 3 nivells de valoració de cada àmbit:
 - **Exclusió:** situació en què la persona queda fora de la participació plena en la societat per dificultats greus en aquest àmbit.
 - **Vulnerabilitat:** situació de risc en caure en exclusió social per les dificultats que pateix en aquest àmbit
 - **Inclusió:** autonomia en aquest àmbit
- **EBAS:** Equip Bàsic d'Atenció Social. L'Ajuntament de Girona, té 7 equips territorials: Barri Vell, Santa Eugènia, Sant Narcís, Palau, Taialà, Pont Major i Girona Est. Cada un

d'aquests equips està format, com a mínim, per les figures professionals de director/a, treballador/a social, educador/a social, integrador/a social i administratiu/a.

- **SAGGID:** Servei d'Atenció a la Gent Gran i Dependències: són el servei bàsic que dissenya els PIA (Pla Individual d'Atenció) i atén les persones amb dependència i persones amb discapacitats. Gestiona els serveis de Teleassistència, Serveis d'Atenció a Domicili, el Banc de Productes de Suport i els menjadors Socials i càtering
- **Porta d'entrada:** Servei d'atenció telefònica que realitza la primera acollida a les persones ateses als SBAS. Centralitza i vehicula les demandes d'atenció a SBAS.
- **Risc de pèrdua d'habitatge:** risc de ser desnonat o d'haver de deixar l'habitatge per un procés de desnonament, finalització de contracte de lloguer o per altres causes.
- **Emergència habitacional:** situació que es dona quan un nucli familiar ja ha perdut l'habitat i no disposa d'una solució habitacional estable.
- **Desnonament:** procés judicial que finalitza amb la pèrdua d'habitatge
- **Vulnerabilitat de gent gran:** situació de pèrdua d'autonomia per edat o estat de salut que afecta a una persona gran que no disposa de xarxa de suport.

Àmbit tecnològic

Aquests termes són habitualment utilitzats per aquells professionals que es dediquen al desenvolupament dels aspectes informàtics i matemàtics del Pigall.

- **Variable:** contenidor que representa una característica o dada d'un objecte. Per exemple, una persona pot tenir la variable "Pes" que, a la vegada, pot adoptar un valor numèric.
- **Variable categòrica:** variable que representa una classe o categoria, per exemple "Gènere" (pot ser Home/Dona/NB), o "Barri".
- **Variable numèrica:** variable que es pot representar amb un nombre. Per exemple, "Edat", "Pes" o "Ingressos".
- **Variable binària:** variable que només pot prendre dos valors: 0 / 1, Sí / No, Cert / Fals, Jove / Adult, etc.
- **Variable booleana:** variable que només pot adoptar els valors 0 / 1. A la pràctica, però, moltes vegades s'utilitza el terme "variable booleana" i "binària" indistintament, donat que qualsevol variable binària es pot expressar com a booleana.
- **Intel·ligència Artificial:** conjunt de mètodes, disciplines i camps que s'ocupen de construir màquines i algorismes autònoms, capaços de resoldre problemes complexos per sí mateixes.
- **Machine Learning / Aprenentatge Automàtic:** subcamp de la intel·ligència artificial dedicat a crear algorismes capaços d'aprendre a resoldre problemes de manera autònoma, a partir de dades o a partir de l'experiència.
- **Xarxa neuronal:** objecte matemàtic, inspirat en el funcionament del cervell dels animals, que es fa servir habitualment per resoldre problemes d'aprenentatge automàtic.
- **Aprenentatge supervisat:** tipus de problema d'aprenentatge automàtic on la màquina intenta aprendre sobre un conjunt de dades etiquetades. Per exemple, suposem que es vol entrenar una xarxa neuronal aprendre a identificar imatges de gossos i gats; per fer-ho, se li proporciona un conjunt de fotografies etiquetades d'aquests dos animals sobre els quals pugui aprendre.

- **Aprenentatge no-supervisat:** tipus de problema d'aprenentatge automàtic on la màquina intenta aprendre sobre un conjunt de dades no-etiquetades, sobre les quals ha de trobar-hi patrons per sí mateixa.
- **Entrenament:** procés pel qual un programa d'aprenentatge automàtic aprèn a resoldre un problema en concret.
- **Problema de clústering:** problema d'aprenentatge no-supervisat consistent en agrupar un conjunt de dades en diferents perfils segons les seves característiques estadísticament més significatives respecte el total del conjunt. El procés d'agrupar les dades en perfils s'anomena *clusteritzar*.
- **Índex (clústering):** hi ha moltes maneres de fer clústering, amb resultats molt diferents. A l'hora de fer un clústering d'unes dades, s'intentarà que els clústers resultants compleixin algun tipus de criteri. Normalment, aquests criteris consisteixen en calcular algun tipus d'índex a partir dels elements de cada clúster. Dos dels índex més habituals són l'índex de Davis-Bouldin i el de la Silueta, que intuïtivament es pot considerar que medeixen si els clústers estan prou diferenciats, són prou explicatius del conjunt de dades i no són redundants.
- **Dataset:** conjunt de dades utilitzades per entrenar un programa de Machine Learning.
- **Base de dades:** col·lecció ordenada de dades emmagatzemades dins d'algun tipus d'infraestructura especialitzada.
- **PostgreSQL:** un dels molts tipus de bases de dades que existeixen.
- **Vector:** informalment parlant, és objecte matemàtic que pot servir per representar direccions i magnituds, coordenades, informació, etc. En el nostre context, cada dada individual és representada com a un vector (podem pensar un data set com un conjunt de vectors).
- **Dimensió:** components d'un vector. Un punt en un pla es pot representar com a un vector de dues dimensions (x,y), un punt a l'espai com un vector de tres dimensions (x,y,z), etc. En el cas de les dades, si les representem com a vectors, aquests tindran tantes dimensions com camps hi hagi a les dades.
- **API:** interfície de connexió entre ordinadors, programes o bases de dades. S'utilitza habitualment per poder extreure informació de bases de dades de forma senzilla.
- **Model:** representació física o abstracta d'un objecte, persona, sistema o procés. Per exemple, un model pot representar un edifici, un procés químic entre dos reactius, etc.
 - **Model matemàtic/estadístic:** model que consisteix en una representació matemàtica abstracta, normalment donada per un conjunt d'equacions. Per exemple, les equacions que descriuen l'evolució d'una població.
 - **Model (aprenentatge automàtic):** a diferència dels models matemàtics, en el context de l'aprenentatge automàtic s'acostuma a utilitzar la paraula "model" per referir-se a un algorisme d'aprenentatge ja entrenat per resoldre un problema concret.
- **Debugger:** procés de solucionar errors (habitualment anomenats "bugs") en el codi d'un programa.
- **Algorisme:** instruccions "pas a pas" per resoldre un tipus de problema o realitzar una tasca.
 - **Algorisme (matemàtic):** algorisme escrit en llenguatge matemàtic.
 - **Algorisme (informàtic):** algorisme escrit en un llenguatge interpretable per un ordinador que serveix per realitzar un cert còmput o transformar unes dades.

Moltes vegades són traduccions a llenguatge màquina d'algorismes matemàtics.

- **HDBSCAN (*Hierarchical Density-Based Spatial Clustering of Applications with Noise*):** un dels molts algorismes de clustering que existeixen. Té una sèrie de propietats matemàtiques molt atractives que el fan idoni per treballar amb les dades que s'han utilitzat pel Pigall.
- **Projecció / reducció de dimensionalitat:** molts algorismes que s'utilitzen en ciència de dades i aprenentatge automàtic tenen problemes per treballar amb dades que tinguin moltes dimensions (i.e. molts camps). Per això, de vegades es redueix la dimensionalitat de les dades, agafant els vectors que les representen i projectan-los sobre un espai de menys dimensions. Per exemple, si les dades es poden representar com a vectors de tres components, es busca trobar alguna manera d'ajuntar informació per poder representar-los com a vectors d'una o dues components.
- **UMAP (*Uniform Manifold Approximation and Projection*):** és un algorisme de projecció. Matemàticament, té la propietat de poder reduir considerablement la dimensionalitat d'un conjunt de dades i, a la vegada, respectar les estructures i relacions subjacents entre les dades originals.
- **Classificador:** model d'aprenentatge supervisat que, donades unes dades, prediu una categoria. Un exemple de classificador pot ser una xarxa neuronal que, donada la imatge d'un animal, digui de quin animal es tracta.
- **Random forest:** algorisme d'aprenentatge supervisat. Té l'avantatge de poder oferir prediccions precises aprenent de poques dades.
- **Frontend:** part d'un programa informàtic "que es veu", és a dir, amb la que interactua directament l'usuari.
- **Backend:** part d'un programa informàtic que queda "per darrere" i que permet que aquest funcioni, ja sigui fent còmput, llegint/escrivint dades, etc.
- **Llenguatge de programació:** conjunt d'instruccions, semblants a un llenguatge humà, que poden ser enteses per un ordinador i permeten expressar algorismes informàtics.
- **Codi:** "text" escrit en un llenguatge de programació.
- **Python:** un dels molts llenguatges de programació que existeixen.
- **Llibreria:** conjunt de codis "prefabricats" que poden ser utilitzats pels programadors per fer una tasca específica, estalviant haver de tornar a escriure programes per realitzar certes tasques. Per exemple, existeixen llibreries que donen eines als programadors per realitzar certes operacions matemàtiques.
- **Apache airflow:** conjunt d'eines i llibreries que permeten analitzar i processar dades en temps real.
- **Scripts ETL:** conjunt de programes que s'ocupen de recollir i processar dades, així com realitzar càlculs computacionalment complicats.
- **Drift (dades temporals):** quan es parla de dades temporals, ens referim a drift com al canvi de tendència al llarg del temps.
- **Drift (models de predicció):** suposem que tenim un model d'aprenentatge automàtic que hem entrenat sobre un sub-conjunt de dades agafat d'una població, i que estem fent servir per fer prediccions sobre aquesta població. Amb el temps, a mesura que les característiques de la població canviïn, es pot donar la situació de que les prediccions del model empitjorin. Aquest fenomen es coneix com a drift, i la manera més simple de solucionar-lo consisteix en tornar a entrenar el model cada cert temps.

- **One-hot encoding:** tècnica de processament de dades que consisteix en convertir una dada categòrica en binària, creant un nou “camp” per cada categoria que pugui adoptar la dada. Per exemple, si volem fer un *one-hot encoding* d’una dada “Gènere” que pugui adoptar les categories “Home”, “Dona” o “No-binari”, es crearien tres columnes binàries (que només poden adoptar els valors Sí/No) amb els noms “És Home?”, “És Dona?”, “És No-Binari?”.
- **Distribució de probabilitat:** objecte matemàtic que permet descriure els resultats d’un experiment aleatori (per exemple, una tirada de dau, obtenir una mà concreta en una partida de pòquer, la probabilitat de que demà faci sol i no núvol, etc).
- **Mètrica de Hellinger:** funció matemàtica que permet dir com de semblants o diferents són dues distribucions de probabilitat.
- **Replicabilitat:** capacitat d’un algorisme, programa o mètode de poder ser copiat per altres persones, obtenint resultats similars. En l’àmbit de l’aprenentatge automàtic, és important que els algorismes siguin replicables, tant per poder entendre els resultats que donen, com per poder adaptar-los a altres usos.
- **Sèrie temporal:** seqüència de dades recollides en un interval de temps. Per exemple, una sèrie temporal podria ser la temperatura recollida diàriament a un punt d’una ciutat.
- **Tendència:** donada una sèrie temporal, la tendència és la direcció general que aquesta segueix (a l’alça, a la baixa, etc).
- **Estacionalitat:** donada una sèrie temporal, l’estacionalitat és la component cíclica d’aquesta que es repeteix cada cert període de temps. Un exemple podria ser observar el volum de turistes que venen a les costes catalanes. La tendència global pot ser a l’alça o a la baixa, però hi ha una component estacional que es repeteix anualment: el volum sempre augmenta a l’estiu i disminueix a l’hivern, creant així un cicle anual.
- **ADWIN:** algorisme capaç de detectar automàticament canvis abruptes en la tendència d’una sèrie

temporal. Aquests canvis acostumen a prendre la forma de pics de creixement/decreixement, o estabilitzacions.

- **SARIMAX:** un dels molts algorismes que existeixen per predir com evolucionarà una sèrie temporal, tenint en compte la seva estacionalitat i tendència.

ANNEX 2

Informe tècnic de les dades de Serveis Socials

ÍNDEX

1. INTRODUCCIÓ.....	63
2. OBJECTIUS	63
3. METODOLOGIA.....	63
4. ARQUITECTURA DE L INFORMACIÓ DE SERVEIS SOCIALS	68
5. ANÀLISIS PRELIMINAR DE LES DADES.....	70
5.1 Taules de Serveis Socials (anàlisi GENERAL).....	70
5.2 Camps recollits a SIAGI	70
5.2.1 Exactitud.....	72
5.2.2 Integritat.....	72
5.2.3 Actualització	74
5.2.4 Confiabilitat	76
6. CONCLUSIONS	77

1. INTRODUCCIÓ

En aquest informe es presenta l'anàlisi preliminar del conjunt de dades recollides a SIAGI, el programari desenvolupat per l'equip de Serveis Informàtics de l'Ajuntament, que permet registrar i gestionar la informació diària de Serveis Socials. Aquest programari emmagatzema totes les dades en una base de dades anomenada SIAGI, que conté la major part de la informació generada per l'Ajuntament.

Cal destacar que la majoria de les taules referenciades en aquest informe comencen amb les lletres "SS", fent referència a les taules específiques de Serveis Socials.

Així doncs, el que es pretén amb aquesta anàlisi preliminar és proporcionar una contextualització analítica de la situació actual de les dades generades per Serveis Socials. Per fer-ho, l'anàlisi s'ha dividit en dues parts:

- **Anàlisi sistemàtica de les taules:** Es fa una anàlisi a gran escala per oferir una visió macro de les dades.
- **Anàlisi detallada dels valors recollits a SIAGI per grups:** En aquesta fase s'avalua la qualitat de les dades tenint en compte aspectes com l'exactitud, la integritat, l'actualització i la confiabilitat, amb l'objectiu d'identificar possibles mancances i millores, o descartar camps que no aportin valor.

A continuació, s'exposen els objectius específics que es volen assolir amb aquest informe.

2. OBJECTIUS

Els objectius d'aquest informe són els següents:

- Comprendre l'arquitectura de les dades de Serveis Socials.
- Obtenir una visió general de la distribució dels registres en cadascuna de les taules de dades.
- Avaluar l'exactitud, la integritat, l'actualització i la confiabilitat de les dades recollides a SIAGI.

3. METODOLOGIA

La metodologia que s'ha seguit per l'anàlisi de les taules s'ha fet amb dos enfocaments diferenciats: general i detall

Quan s'ha utilitzat l'enfoc general, s'ha avaluat:

- Si la taula correspon a una SERIE_TEMPORAL o a una CATEGORIA
- En les sèries temporals s'ha calculat:
 - Data inici i final
 - Completesa de les dades. Per a cada atribut
 - Número de registres
 - Número de nulls
 - % de nulls respecte el total
 - Mitjana de valors inserits utilitzant tota la sèrie
 - Anàlisi de valors atípics (outliers) utilitzant 3 mètodes de diagnosi:

- Z method
- IQR method
- RS method

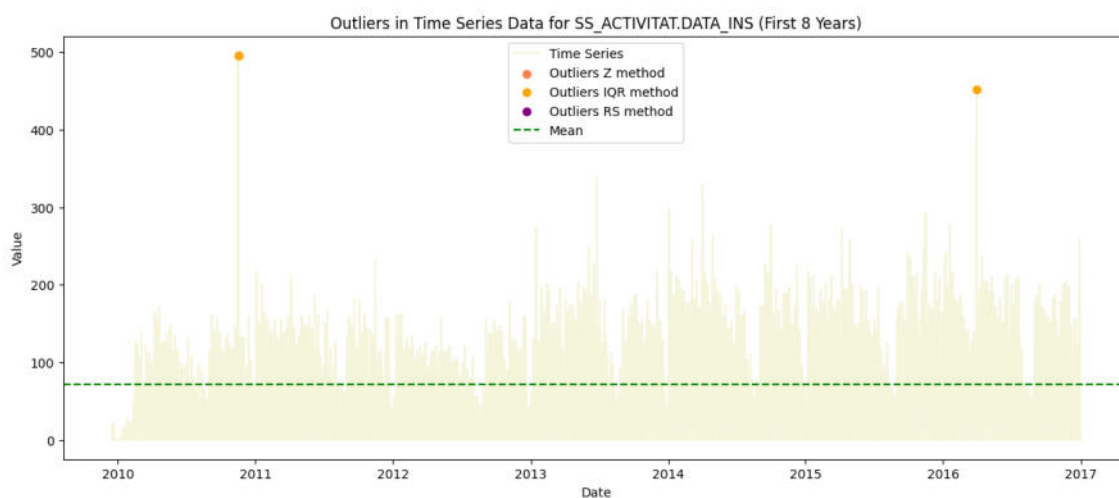


Figura 3.1: Valors atípics de la taula SS_ACTIVITAT (primers 8 anys).

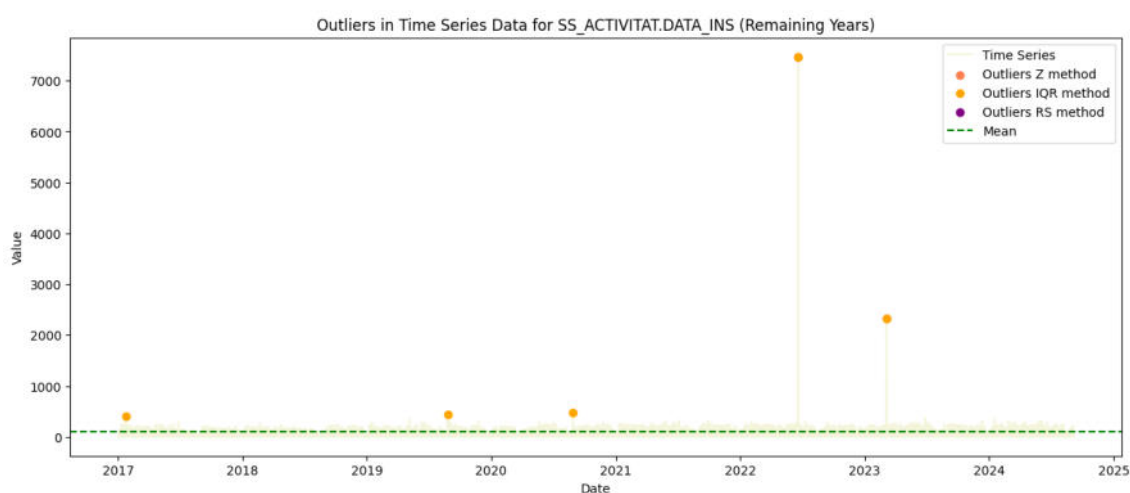


Figura 3.2: Valors atípics de la taula SS_ACTIVITAT (altres anys).

SERIES TEMPORALS		CATEGORIES	
TOTAL: 64 series temporals (amb possibilitat de reduir)		TOTAL:98 taules categoriques (inclou taules que no tractem)	
NOM_TAULA	NUM_ROWS	NOM_TAULA	NUM_ROWS
Veure fitxer: Classificació_taulas_udg.xlsx			

Figura 3.3: Distribució de les taules.

EXEMPLE D'ANÀLISI

TAULA: SS_MOTIU_TANCAMENT

SSMTA_CODI DESCRIPCIO

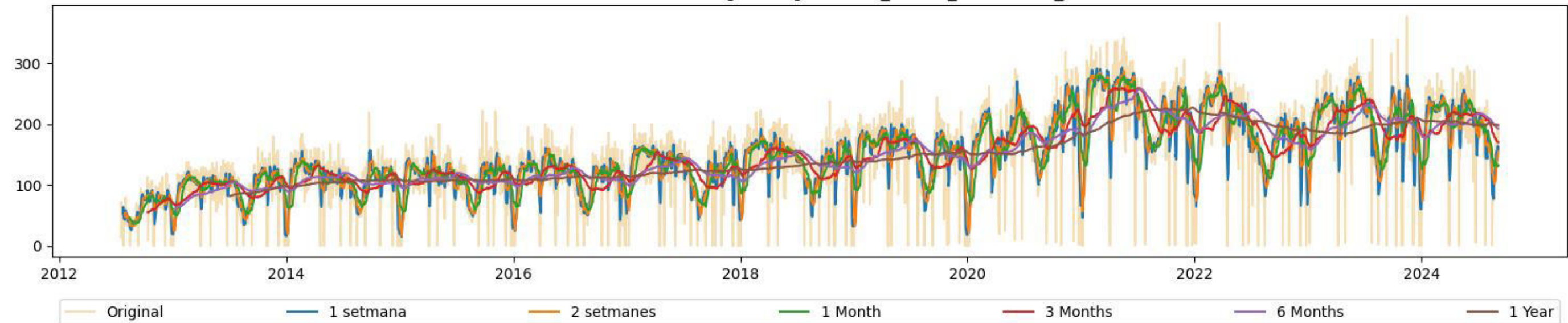
FIN Finalització intervenció
CAD Caducitat de la intervenció
CS Canvi de situació
ASSOB Assoliment objectiu

SSMTA_CODI DESCRIPCIO

MCOL Manca col·laboració interessat/da
MREC Manca de recursos
CEST Canvi d'estratègia o modificació objectiu
ALTR altres motius

TAULA: SS_ACCES_SIAGI

Time Series and Moving Averages for SS_ACCES_SIAGI.DATA_INS



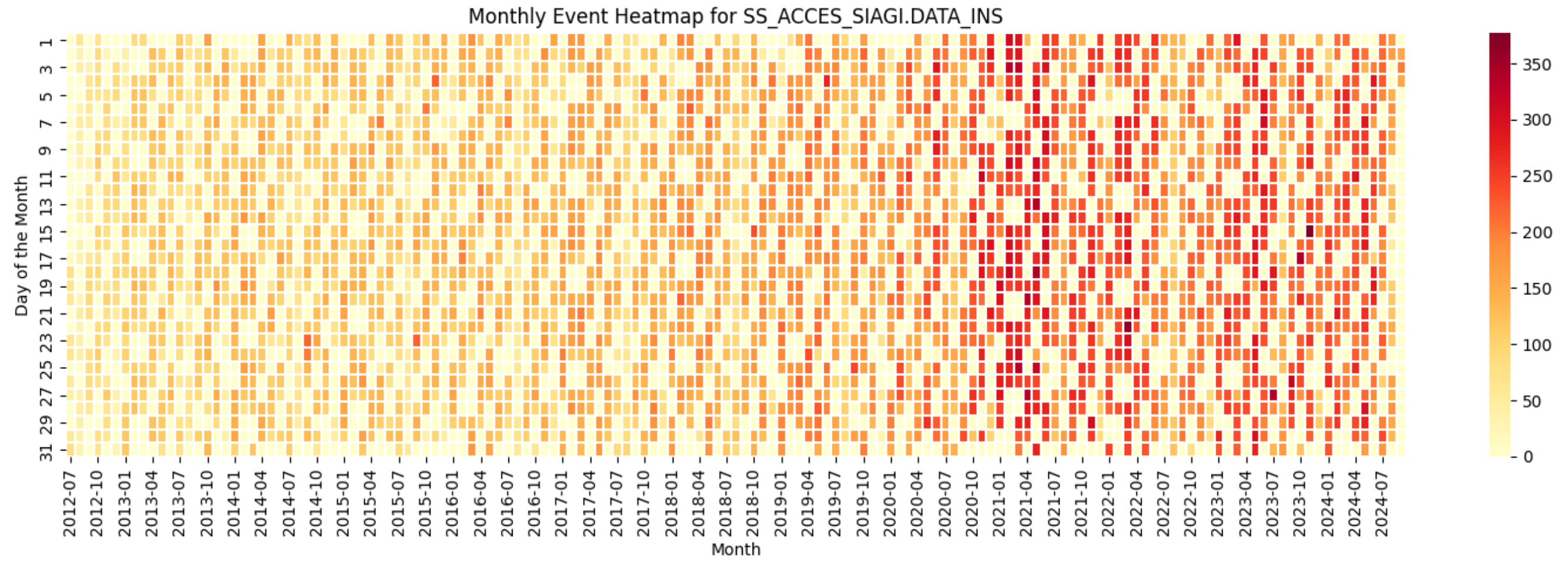


Figura 3.4: Exemple d'anàlisi.

A continuació, detallarem la metodologia utilitzada per analitzar cada camp que es recull a SIAGI:

Per tal d'analitzar tota la informació que es recull a SIAGI s'ha recollit en l'Excel adjuntat "informe_tecnic_dades_ss".

L'estructura que s'ha seguit és la següent:

- **Gran grup de dades:** Engloba un grup de dades.
- **Grup de dades:** Pot englobar un subgrup de dades.
- **Subgrup de dades:** Pot englobar un conjunt de camps
- **Nom del camp:** Nom del camp que es recull.
- **Tipus de dada:** Tipus de format de la dada.
- **Assignació:** Fa referència a com s'assigna el valor.
- **Explicació dada/ per què serveix:** Explicació bàsica sobre el camp.
- **Data inici:** Data inici de quan es comença a registrar cada valor.
- **Data actual:** Correspon a la data més actual en fer l'anàlisi.
- **Mitjana per any:** Mitjana de registres per any.
- **Mitjana per mes:** Mitjana de registres per mes.
- **Mitjana per dia:** Mitjana de registres per dia.
- **Tendència:** Tendència del conjunt de registres al llarg de la sèrie.
- **Total atípics:** Càlcul de valors atípics en el cas dels camps numèrics.
- **Ajustament:** Ajustament que es realitza en el cas dels camps numèrics amb valors nuls.
- **Min:** Valor mínim d'un camp numèric.
- **Max:** Valor màxim d'un camp numèric.
- **Mitjana:** Mitjana d'un camp numèric.
- **Mediana:** Mediana d'un camp numèric.
- **Desviació estàndard:** sd d'un camp numèric.
- **Nº categories:** Nombre de categories d'un camp categòric.
- **Distribució de categories:** Distribució percentual de les categories d'un camp categòric.
- **Valors nuls:** Total de nuls en valors absoluts.
- **% Valors nuls:** Percentatge de valors nuls respecte el total de registres.
- **Total registres:** Total de registres d'un camp.
- **Taula:** Nom de la taula on s'ha extret la informació del camp.
- **Integritat:** Si hi ha més del 5% de valors nuls en un camp, es determina que no es compleix la integritat.
- **Observacions:** Altres observacions a destacar sobre el camp analitzat.

Pel que fa als càlculs que s'han realitzat, s'ha fet ús de les funcions incorporades a Oracle, amb SQL.

L'únic càlcul que s'ha realitzat de manera manual ha sigut la tendència de la següent manera:

$$y = \frac{y_2 - y_1}{t_2 - t_1} \cdot (t - t_1) + y_1$$

On:

- **y**: El valor estimat de la variable dependent (o el valor de la tendència) en un temps *t*.
- **y1**: El valor inicial de la variable dependent en el temps inicial *t1*.
- **y2**: El valor final de la variable dependent en el temps final *t2*.
- **t1**: El temps inicial.
- **t2**: El temps final.
- **t**: El temps pel qual es vol calcular el valor de la tendència.

Si ho passem a codi SQL:

```
WITH data AS (
  SELECT
    ROW_NUMBER() OVER (ORDER BY <data>) AS x,
    count(*) AS y
  FROM <nom taula>
  WHERE <categoria> is not null
  group by <data>
)
SELECT
  (COUNT(x) * SUM(x * y) - SUM(x) * SUM(y)) /
  (COUNT(x) * SUM(x * x) - SUM(x) * SUM(x)) AS tend
FROM data
```

Per últim, per tal d'avaluar la qualitat de les dades s'ha analitzat l'exactitud, la integritat, l'actualització i la confiabilitat. Tanmateix, si es vol afinar l'anàlisi sobre algun camp en particular us podeu adreçar a l'Excel "informe_tecnic_dades_ss" per tal de poder veure un resum estadístic i de la descripció més detallada del camp.

4. ARQUITECTURA DE LA INFORMACIÓ DE SERVEIS SOCIALS

Les dades de SIAGI es poden estructurar en tres categories:

- Dades generals
 - Dades individuals:
 - dades identificatives
 - dades sensibles (visibles només pels tècnics de Serveis Socials)
 - Dades de nucli
 - dades identificatives
 - dades sensibles (visibles només pels tècnics de Serveis Socials)
- Dades d'intervenció: registre d'activitats
- Dades de gestió: Informes i expedients.

Aquestes dades s'organitzen en quatre grans grups:

- Serveis d'atenció social
- Serveis d'atenció GG i DP
- Centres Oberts
- Dades d'Atenció

El Bloc de dades **Serveis d'Atenció Social** és el més extens, ja que conté dades generals i identificatives, dades de la persona atesa i dades del nucli familiar.

Les dades identificatives són les dades habituals i concretes que es recullen des de l'Administració per identificar i poder-se comunicar amb la ciutadania.

Les dades generals permeten a les tècniques de serveis socials situar la persona atesa en un context vital i conèixer el seu entorn.

La informació complementària es divideix en 5 àmbits o dimensions de la vida de la persona atesa, que aporten informació sobre la necessitat o problemàtica del cas concret:

- Habitatge
- Situació econòmica
- Laboral o escolar
- Salut
- Relacional

Pel que les dades registrades d'aquestes dimensions serà més completa en tant que tingui relació directa amb la necessitat o problemàtica del cas concret, ja que si es tracta d'un tema estrictament econòmic, la tècnica registrarà les dades d'aquest àmbit però no registrarà dades de l'àmbit de salut o escolars dels fills, pel que de tots els casos registrats trobarem complerta la informació general en canvi la complementària, només estarà complerta en aquells àmbits que interessin per a la posterior intervenció, pel que poden contenir taules o camps buits.

D'aquestes dimensions de la vida de la persona, les que estiguin omplertes ens aporten informació rellevant per conèixer la situació del conjunt de les persones ateses i els seus nuclis. Ens permeten configurar una fotografia de quines són les problemàtiques econòmiques de les ciutadanes de Girona, de la vulnerabilitat residencial, de la infància en risc detectada a la ciutat i de problemàtiques associades.

Finalment, el pla de treball ens dona dades de quines són les accions que s'acorden amb la persona atesa per millorar la seva situació i l'Ajut ens en dona sobre la quantitat d'ajuts atorgats i del tipus, així com quants diners s'hi destinen.

5. ANÀLISIS PRELIMINAR DE LES DADES

5.1 Taules de Serveis Socials (anàlisi GENERAL)

L'anàlisi general de totes les taules ens ha permès:

1. Poder classificar les taules que tenen característiques de sèrie temporal vs. categòriques
2. En les sèries temporals s'han identificat outliers d'addició (no s'ha estudiat encara les eliminacions o modificacions dels registres).
3. A falta d'una anàlisi encara més profund del conjunt de dades, sospitem que aquests outliers poden estar causats principalment per:
 - a. Transferències massives d'un sistema d'informació a un altre (especialment en la franja d'anys 2009-2013)
 - b. Decisions preses a nivell de Serveis Socials, com per exemple: "A partir de XX/YY/YYYY cal registrar tots els casos de forma P com Q"
 - c. Accions automatitzades (triggers de la BD) generen molt registres. Per exemple:
 - i. Es creen +1000 registres per a tota la ciutat en una data concreta, corresponent a la factura de serveis del servei de TeleAssistència, considerant que cada registre (línia) és una hora a facturar del servei
 - d. Registres que tenen una temporalitat estacional, com poden ser els ajuts de les beques menjador (setembre, inici del curs escolar) o d'ajudes al lloguer que es donen en diferents moments concrets de l'any.
4. Fins ara, i sense aplicar tècniques avançades, només amb l'observació de les dades, s'han estimat els següents límits
 - a. Entre 40-60 registres diaris, poden ser considerats normals i són possibles de realitzar per un treballador en una jornada laboral
 - b. Entre 60-500-1000 registres poden ser fruit d'una acció massiva controlada (trigger)
 - c. +1000 registres poden ser fruit d'una acció massiva o transferència de dades entre sistemes
5. La correcció de les dades, en base a la seva mitjana d'ús, pot estar clarament influenciada per com s'utilitza la taula al SIAGI
 - a. Hi ha taules amb mitjanes d'ús i registres més altes (SS_REGISTRE_INF), ja que és un camp "curt i fàcil d'omplir" que s'utilitza com un sistema de missatgeria interna entre departaments.
 - b. Hi ha altres taules (SS_USUARIS_NUCLIS) que només s'afegirà informació quan es procedeixi a fer una entrevista de la persona usuària i el seu nucli de convivència.

5.2 Camps recollits a SIAGI

En aquest apartat analitzarem les dades que es recullen a SIAGI a partir dels diferents grups de dades i també valorarem l'exactitud, la integritat, l'actualització i la confiabilitat.

Primer de tot, observarem la distribució dels diferents tipus de dades dels camps recollits:

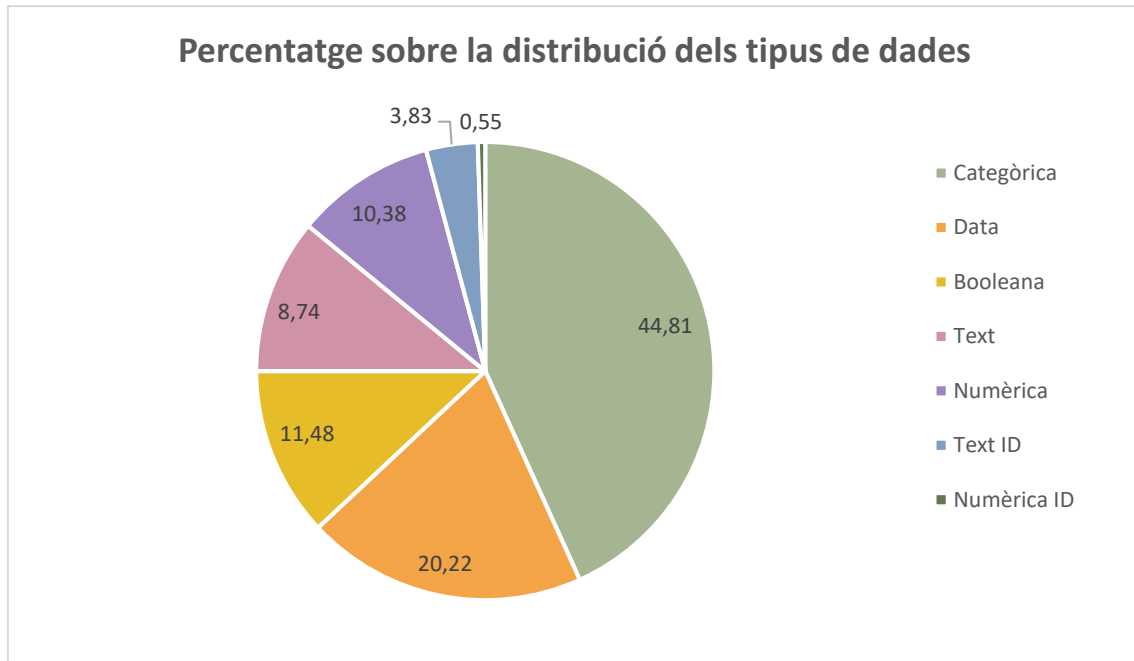


Figura 5.1: Distribució de les dades segons el seu tipus.

Tal com podem veure a la gràfica, més del 80% de les dades que es recullen són de tipus categòrica, data, booleana i camps ID. Aquest fet ajuda a minimitzar els errors de registre a SIAGI. D'altra banda, tenim camps de tipus text que es registra de manera manual i de tipus numèrica que es pot recollir de manera manual o és un càlcul automàtic.

Pel que fa a l'actualització de les categories, SIAGI té taules "tipus" on es guarden totes les categories úniques que s'utilitzen en les demès taules. Això permet actualitzar la categoria corresponent sense afectar la resta d'informació. No obstant, s'ha de tenir en compte que es pot donar el cas que en una taula hi hagi categories antigues que ja no s'utilitzin, i aquest fet pot afectar la nostra anàlisi. Per tant, s'hauria d'identificar correctament aquelles categories que no tenen una continuïtat constant al llarg de la sèrie de registres de les diferents taules.

Pel que fa a la informació, tal com hem comentat a l'apartat 4. Arquitectura de la informació de Serveis Socials, aquesta es divideix en quatre grups:

- Serveis d'atenció social
- Serveis d'atenció GG i DEP
- Centres oberts
- Dades d'atenció

Aquests grups els podeu veure distribuïts per les diferents fulles de l'Excel adjuntat (informe_tecnic_dades_ss).

5.2.1 Exactitud

Entenem per exactitud allò que mesura si les dades reflecteixen amb precisió la realitat o el fenomen que es vol mesurar. Les dades exactes són correctes i no contenen errors.

En el nostre cas, podem observar que la majoria de camps són categòrics, booleans i de tipus ID. Aquest fet implica que aproximadament el 80% de les dades són exactes, ja que aquests tipus de camps tenen poc marge d'error en el moment de registrar-les. No obstant això, s'ha de remarcar que hi ha usuaris duplicats i alguns de tipus test, el que afecta l'exactitud global de les dades.

5.2.2 Integritat

Avalua si les dades són completes i no hi falten valors importants. La integritat assegura que tota la informació necessària està present i és coherent.

A continuació, mostrarem la distribució de valors nuls per cada un dels grans conjunts de dades.

Serveis d'atenció social

Una persona atesa per Serveis d'atenció social pot tenir la informació registrada en aquests grans grups:

- General
- Informació complementària
- Registre d'informació
- Informació d'ajuts

La distribució de valors nuls seria la següent:

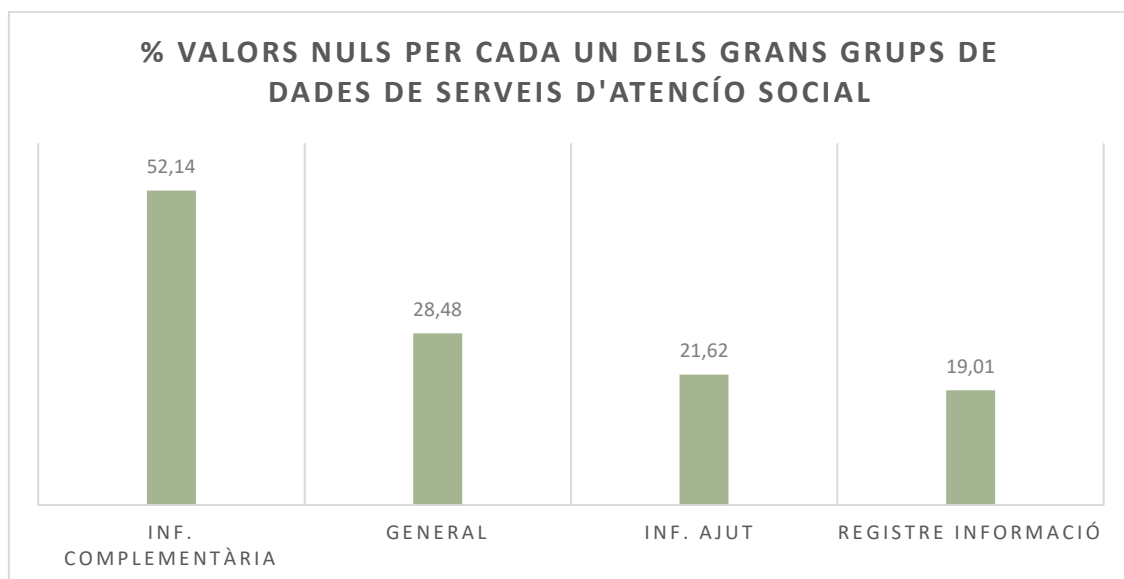


Figura 5.2: Distribució dels valors nuls a Serveis d'atenció social.

S'observa, doncs, com la informació complementària és on es troba una gran quantitat de registres sense dades. Això pot ser degut, ja que es pot donar el cas on una persona atesa que és registrada pot tenir un problema específic, com per exemple un ajut puntual. En aquest cas, no es registra tota la informació sobre aquesta persona a informació complementària i s'omplen els camps de la informació més rellevant. Tot i que si és un ajut econòmic sí que s'han d'omplir les dades sobre el nucli i sobre l'àmbit econòmic.

Serveis d'atenció a gent gran i de dependència

Dins d'aquests serveis es registra tota la informació corresponent a gent gran o persones amb dependència. És important entendre que tota persona que tingui informació tant en el SAGID com a l'EAD estaran registrades com a usuàries i tindran informació a Serveis d'atenció social.

Així doncs, la informació recollida es distribueix en aquests dos grans grups:

- Serveis d'Atenció a la Gent gran i de Dependència (SAGID)
- Equip d'Atenció a la Dependència (EAD)

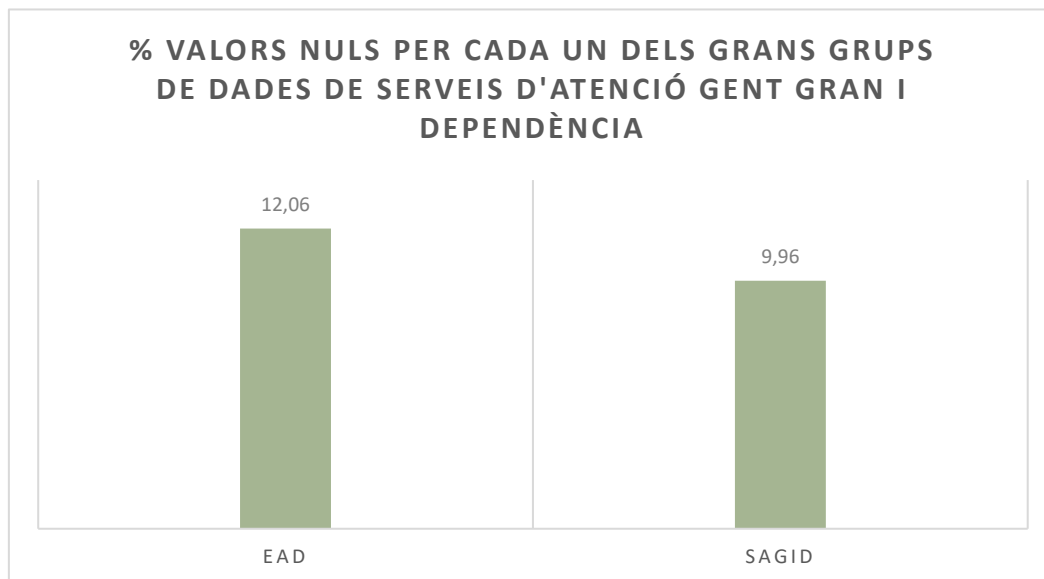


Figura 5.3: Distribució dels valors nuls dels Serveis d'atenció a gent gran i de dependència.

Podem observar que a **EAD** hi ha més d'un 10% de registres buits. No obstant això, molts d'aquests camps buits fan referència a camps relacionats amb la baixa de l'usuari sobre un expedient de Programa Individual d'Atenció (PIA). Així, la integritat de les dades no es veu tan afectada perquè es tracta d'usuaris que encara no han causat baixa, i per tant, és normal que aquests camps estiguin buits en aquests casos.

Pel que fa a **SAGID**, ens trobem en una situació similar. Molts usuaris encara no han causat baixa, i els camps de tipus data que fan referència a la baixa de la persona atesa estan buits. A més, observem que el camp sobre l'import manual pràcticament no es registra. Tot i que aquest camp

no és rellevant per a la majoria d'usuaris, és important tenir en compte que la manca de registre no afecta significativament la integritat general de les dades.

Centres Oberts

Els centres oberts registren infants o adolescents en situació de risc que participen en activitats. El conjunt de registres d'aquest gran grup presenta un 47,45% de valors nuls. Aquest alt percentatge de valors nuls es deu principalment al fet que molts dels camps han estat afegits posteriorment. Com a resultat, la integritat d'aquestes dades és deficient. Si es vol aprofundir en l'anàlisi d'aquestes dades, seria necessari depurar els registres, eliminant o ajustant aquells que presenten dades incompletes per aconseguir una integritat de dades adequada.

Dades d'atenció

Dins de les dades d'atenció es registren, principalment, totes les activitats realitzades a Serveis Socials. El percentatge de registres buits en aquest conjunt de dades és del 19,85%. Alguns camps, com el lloc o el canal pel qual es realitza l'activitat, presenten més d'un 20% de valors nuls. És important saber que els valors nuls del camp Lloc tendeixen a referir-se a la categoria "serveis".

D'altra banda, hi ha camps amb més del 80% de valors nuls, com el motiu d'anul·lació. Aquesta situació pot ser deguda al fet que moltes activitats no han estat anul·lades o que s'han realitzat, i això provoca que hi hagi un gran nombre de registres amb aquest camp buit.

A més a més, es nota que el camp "Tema" de l'activitat no es registra adequadament. Hi pot haver una tendència a escriure manualment el motiu de l'activitat al camp de text, en lloc de registrar-lo de manera estructurada, el que pot afectar la consistència i l'anàlisi posterior de les dades.

Finalment, s'ha observat que els serveis associats a cada professional s'han d'assignar manualment quan es fa un registre d'activitat. Això provoca que l'exactitud de les dades es vegi perjudicada, ja que pot haver-hi un marge d'error elevat. Una possible solució que es preveu a curt termini seria automatitzar la selecció del servei, una vegada s'ha afegit el professional corresponent.

5.2.3 Actualització

L'actualització indica si les dades són recents i es mantenen al dia. En aquest cas, el conjunt global dels camps és operatiu fins a la data actual. No obstant això, alguns camps sembla que s'han deixat d'utilitzar, amb els últims registres en les següents dates:

- **Seguiment pediàtric (Serveis d'Atenció Social):** Últim registre el 18/10/2023.
- **Import manual (Serveis d'Atenció GG i DEP):** Últim registre el 28/08/2023.

Pel que fa a la resta de camps, tots estan actualitzats amb registres fins a l'any 2024, el que indica que estan al dia.

A més, hem mesurat la tendència en el registre d'informació de cada camp on hi ha menys del 5% de valors nuls. Mesurar la tendència en el registre d'informació de cada camp és útil per identificar quina informació es registra amb més freqüència al llarg del temps. Aquesta anàlisi pot revelar patrons en la recollida de dades i ajudar a comprendre millor les variacions en les activitats i necessitats dels usuaris.

No obstant això, és important considerar que les tendències en els registres poden estar influenciades per diversos factors socials i contextos específics del territori. Això significa que els canvis en les dades poden reflectir variacions en les condicions socials, canvis en les polítiques o en les prioritats dels serveis socials, més que només variacions aleatòries.

Així doncs, podem veure la distribució de la tendència de la següent manera:

Serveis d'atenció social

	Tendència negativa	Tendència positiva	> 5% de valors nuls
INF. COMPLEMENTÀRIA	14	5	51
GENERAL	12	0	17
INF. AJUT	9	11	17
REGISTRE INFORMACIÓ	5	0	3

Figura 5.4: Distribució de la tendència en els diferents camps per Serveis d'atenció social.

Serveis d'atenció a gent gran i de dependència

	Tendència negativa	Tendència positiva	> 5% de valors nuls
SAGID	5	1	5
EAD	9	0	3

Figura 5.5: Distribució de la tendència en els diferents camps per Serveis d'atenció a gent gran i de dependència.

Centres Oberts

	Tendència negativa	Tendència positiva	> 5% de valors nuls
CENTRES OBERTS	3	0	7

Figura 5.6: Distribució de la tendència en els diferents camps per Centres Oberts.

Dades d'atenció

	Tendència negativa	Tendència positiva	> 5% de valors nuls
REGISTRE ACTIVITATS	1	7	5

Figura 5.7: Distribució de la tendència en els diferents camps per Dades d'atenció.

5.2.4 Confiabilitat

La confiabilitat de les dades avalua la precisió, la consistència i l'estabilitat de les dades al llarg del temps.

Precisió i Consistència: En aquest cas, hem detectat la presència de registres duplicats, el que pot afectar la precisió i la consistència de les dades. Els usuaris duplicats poden conduir a una informació inconsistent o errònia, ja que els registres duplicats poden contenir dades contradictòries o incompletes. A més, com que només es registra la informació rellevant per cada cas i no tota la informació sobre cada usuari nou, això pot afectar la completesa i, en conseqüència, la precisió de les dades, en el cas que es vulgui observar el cas des d'una visió global.

Mesures Correctives: És crucial processar les dades per eliminar els duplicats i establir mesures per prevenir la creació de registres duplicats en el futur. Això pot incloure la implementació de procediments de verificació de dades i la millora de les pràctiques de registre per assegurar que cada usuari sigui únic i que les seves dades es registrin de manera precisa i consistent.

Estructuració de l'Anàlisi: Per a les anàlisis posteriors, s'ha de tenir en compte que no tots els usuaris tenen dades completament registrades en tots els camps analitzats. Per tant, és important estructurar les investigacions tenint en compte les limitacions de les dades i treballar amb els diferents grans conjunts de dades i els diversos àmbits de Serveis Socials. Això permetrà una anàlisi més acurada i una interpretació correcta de la informació disponible.

6. CONCLUSIONS

En aquest anàlisi preliminar de les dades s'ha estructurat l'avaluació en dues parts.

En la primera part s'ha analitzat de manera agnòstica totes les taules de SIAGI. Aquesta anàlisi de les taules ha permès classificar-les en sèrie temporal i categòriques. En les sèries temporals s'han identificat outliers, potencialment causats per transferències massives de dades, decisions institucionals, accions automatitzades, o temporalitat estacionària. Els registres normals varien entre 40-60 diaris, mentre que entre 60 i 1000 podrien ser resultat d'accions massives controlades, i més de 1000 podrien indicar accions massives o transferències de dades. La correcció de dades pot estar influenciada per com s'utilitza cada taula, amb variacions en les mitjanes d'ús segons la seva funció i tipus d'informació registrada.

Pel que fa a la segona part. S'ha analitzat més detalladament tots els camps recollits a SIAGI, centrant-se en l'exactitud, la integritat, l'actualització i la confiabilitat de les dades, per així poder valorar la qualitat d'aquestes.

Distribució de dades: Aproximadament el 80% de les dades recollides són categòriques, booleans, ID i dates, cosa que redueix els errors. Els camps de text i numèrics poden presentar més variabilitat, amb registres manuals o automàtics.

Exactitud: La majoria de camps són exactes, però hi ha problemes amb usuaris duplicats i alguns camps de tipus test que poden afectar la precisió global.

Integritat: Alguns grans conjunts de dades, com els Serveis d'atenció social i els Centres Oberts, presenten valors nuls significatius. La informació complementària i els Centres Oberts mostren una alta proporció de dades mancants, cosa que indica problemes amb la integritat. S'ha de tenir en compte, però, que de tots els casos registrats trobarem complerta la informació general. No obstant, la complementària només estarà complerta en aquells àmbits que interessin per a la posterior intervenció, pel què poden contenir taules o camps buits.

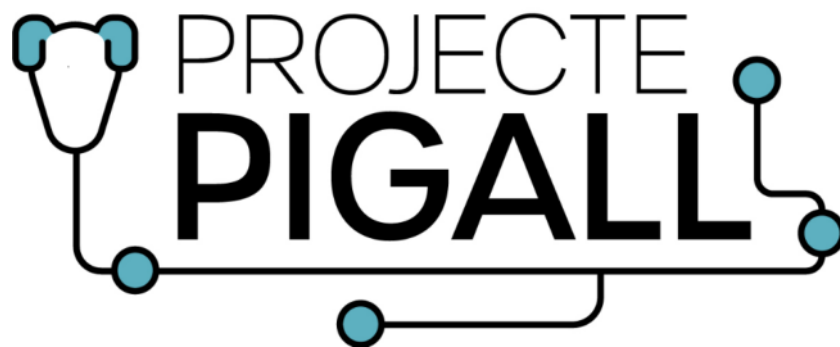
Actualització: La majoria dels camps estan actualitzats fins al 2024, però alguns camps específics no es registren des de dates recents. L'anàlisi de tendències mostra variacions en la freqüència de registre que poden reflectir canvis socials o en les polítiques dels serveis socials.

Confiabilitat: La presència de registres duplicats afecta la precisió i consistència. És essencial eliminar duplicats i establir procediments per evitar futurs errors. La informació ha de ser estructurada tenint en compte les limitacions de les dades per assegurar una anàlisi acurada.

En resum, el sistema SIAGI presenta un sistema de minimització d'errors de registre a partir de la captació d'informació amb valors preestablerts. Tot i això, es necessita implementar millores en la integritat de les dades i en la gestió dels registres duplicats per garantir la qualitat i precisió de la informació recollida. A més, en futures anàlisis serà important estructurar la diagnosi sobre els diferents àmbits de Serveis Socials per reduir els biaixos i obtenir resultats més precisos.

ANNEX 3

ANÀLISI DE LES DADES DELS ÀMBITS D'ACTUACIÓ DE SERVEIS SOCIALS: *PREPROCESSAMENT DE LES DADES*



ÍNDEX

1. ÀMBIT ECONÒMIC.....	80
2. ÀMBIT LABORAL	93
3. ÀMBIT RESIDENCIAL	95
4. ÀMBIT EDUCATIU	98
5. ÀMBIT RELACIONAL.....	101
6. ÀMBIT SALUT EMOCIONAL	106

1. ÀMBIT ECONÒMIC

L'àmbit econòmic contempla informació de cinc fonts diferents: Ajuts, Personal, Nucli, Valoració econòmica i Activitat econòmica.

Per tal de seleccionar les persones que han estat beneficiàries d'ajuts d'urgència municipal, s'han seleccionat aquelles que han demanat ajut i se'ls hi ha aprovat. Com que una persona pot ser beneficiària de més d'un ajut en tot el seu recorregut per Serveis Socials, s'ha seleccionat com a referència l'últim ajut atorgat.

OBJECTIU DEL DOCUMENT

- Entendre com s'ha construït la taula
- Obtenir una visió general de l'estat de les dades

FONTS

En aquest apartat explicarem com es registra cada font, els camps, i com es van unint a la font principal AJUTS.

AJUTS

EL SIAGI conté informació sobre els diferents ajuts que una persona ha rebut durant el llarg de la seva trajectòria de relació amb els Serveis Socials. La informació referent als ajuts es troba en gran part a la taula SS_PROPOSTA_AJUTS. És important entendre que fins al 2019, els ajuts aprovats es guardaven a partir del camp 'EXPEDIENT' emplenat. No obstant, a partir del 2020, es comença a utilitzar un altre camp 'NUMERO_INFORME' que s'ha de mapejar amb el 'NUM_EXP' de la taula INFORMES per saber si l'ajut s'ha aprovat finalment.

Així doncs, les dades que es recullen en aquest apartat es resumeixen en les següents:

- **DATA_AJUT:** Data de l'ajut
- **IMPORT_TOTAL:** Import de l'ajut
- **ALTRES_AJUTS:** Si la persona ja ha rebut ajuts en els últims 12 mesos
- **CONCEPTE:** Concepte de l'ajut
- **IMPORT_AJUTS_TOTALS:** Import total de tots els ajuts demanats per persona
- **AJUTS_TOTALS:** Nombre d'ajuts demanats per persona

PERSONAL

La informació personal de la persona també s'ha tingut en compte. A continuació, mostrem les dades referents a Personal:

- **ID_USUARI**
- **DATA_BAIXA_GENERAL**
- **MOTIU_BAIXA_GENERAL**
- **GENERE**
- **DATA_NAIXEMENT**
- **EDAT_AJUT**
- **SECTOR**
- **PAIS_NAIXEMENT**
- **NACIONALITAT**
- **SITUACIO_LEGAL**

NUCLI

La informació referent al nucli familiar també s'ha tingut en compte per la construcció d'aquesta taula. Les dades que es recullen a Nucli són les següents:

- **DATA_INS_NUCLI:** Data d'inscripció del nucli
- **SSNCO_ID:** ID del nucli
- **SSTIU:** Tipus d'unitat del nucli
- **DATA_INS_PERSONA:** Data d'inscripció de la persona al nucli
- **DATA_BAIXA_NUCLI**
- **MOTIU_BAIXA_NUCLI**
- **ES_TITULAR:** Mostra si una persona és la persona titular del nucli. És a dir, la persona atesa per Serveis Socials.
- **TOTAL_PERSONES_NUCLI**
- **TE_FAMILIA:** Mostra si la persona té família, ja sigui a la mateixa ciutat o fora.

Per tal d'afegir les dades de nucli, es filtren els registres per usuari i se selecciona un de coherent segons la relació entre la data d'inscripció al nucli familiar i la data de l'ajut.

En general, el que es fa és:

1. **Identificar usuaris únics:** Es treballa individualment amb cada usuari per analitzar els seus registres.
2. **Gestionar usuaris amb múltiples nuclis:** Es té en compte l'estat del nucli en el moment en que es percep l'ajut.

3. **Conservar registres únics sense conflictes:** Si un usuari només està associat a un únic nucli, es manté tota la informació sense fer cap canvi.

El resultat final és un nou conjunt de dades amb un únic registre per usuari, que prioritza la consistència entre la data d'inscripció al nucli i la data de l'ajut. A continuació, s'analitzen i es preparen les dades d'ajuts i nuclis familiars per obtenir una visió clara de quantes persones estaven actives en cada nucli familiar en el moment en què es va concedir un ajut. A grans trets, el que es vol aconseguir és:

1. **Identificar les dates rellevants d'ajuts:** Per cada nucli familiar, es detecten les dates en què es van concedir ajuts.
2. **Calcular el nombre de membres actius del nucli:** Es compten quantes persones estaven "actives" en el nucli en cada data d'ajut. Una persona es considera activa si:
 - No té una data de baixa general o aquesta és posterior a la data de l'ajut.
 - No té una data de baixa específica del nucli o aquesta és posterior a la data de l'ajut.
3. **Gestionar diferents situacions:**
 - Si només hi ha una data d'ajut per al nucli, s'utilitza aquesta data per comptar els membres actius.
 - Si hi ha més d'una data d'ajut per al mateix nucli, es realitza el càlcul per a cada data de forma independent.
4. **Consolidar la informació:** Finalment, es crea un conjunt de dades que inclou, per a cada ajut, el nombre total de persones actives al nucli en aquell moment.

Finalment, s'afegeix la informació respecte si la persona té família identificant la connexió més propera entre l'usuari i la seva família en el moment d'un ajut. Per cada ajut, es comprova si l'usuari estava associat a una família abans de demanar l'ajut o en un marge de 60 dies després de la data de l'ajut. Si hi ha diverses associacions, es selecciona la més recent.

VALORACIÓ ECONÒMICA

Quan una persona és atesa a Serveis Socials per tenir dificultats econòmiques, els tècnics fan una primera valoració. En aquest cas, es registra informació sobre el tipus de cobertura i situació econòmica, i si té deutes. Aquestes dades es registren en una mateixa taula (SS_NUCLI_DEUTE) i cada persona pot tenir el seu històric registrat. És important entendre que el criteri de valoració es farà tenint en compte el nucli familiar. Per tant, la informació serà la mateixa per a tots els membres del nucli familiar.

Així doncs, les dades que es recullen en aquesta valoració es resumeixen en les següents:

- **DATA_DEUTE:** Data de la valoració
- **SSTDE_CODI:** Tipus de deute
- **SSTSE_CODI:** Tipus de suport econòmic
- **SSTCE_CODI:** Tipus de cobertura econòmica
- **CODI_AMBIT_VAL:** El resultat final de la valoració

Per tal d'afegir les dades de valoració econòmica es té com a objectiu identificar la data de deute més propera a la data de l'ajut per a cada nucli familiar, i vincular aquesta informació amb altres detalls relacionats amb el deute.

En general, el que es fa és:

1. **Identificar la data de deute més propera a l'ajut:** Per a cada ajut, es busca la data de deute (DATA_DEUTE) més propera a la data de l'ajut (DATA_AJUT) que estigui creada abans de demanar l'ajut o en un marge de 60 dies després de la data de l'ajut. Si hi ha diverses dates de deute, es selecciona la més recent.

ACTIVITAT ECONÒMICA

A part de la informació que registra la tècnica en les valoracions, les dades sobre els ingressos i les despeses també es registren, però aquests no tenen perquè tenir la mateixa data de registre de la de valoració. De fet, aquestes dades es registren a una altra taula (SS_ACTIVITAT_ECONOMICA). Aquestes dades també es registren tenint en compte el nucli familiar i també tenen històric.

Així doncs, les dades que es recullen en aquesta valoració es resumeixen en les següents:

- **DATA_ACTIVITAT_ECO:** Data del registre de l'import
- **INGRESSOS**
- **DESPESES**
- **DATA_ACTIVITAT_ECO_ULTIMA:** Data més actual de l'últim registre
- **INGRESSOS_ULTIMS:** Ingressos més actuals
- **DESPESES_ULTIMS:** Despeses més actuals

L' objectiu és identificar la data de la valoració de l'activitat econòmica realitzada més propera a la data de l'ajut per a cada nucli familiar.

En general, el que es fa és:

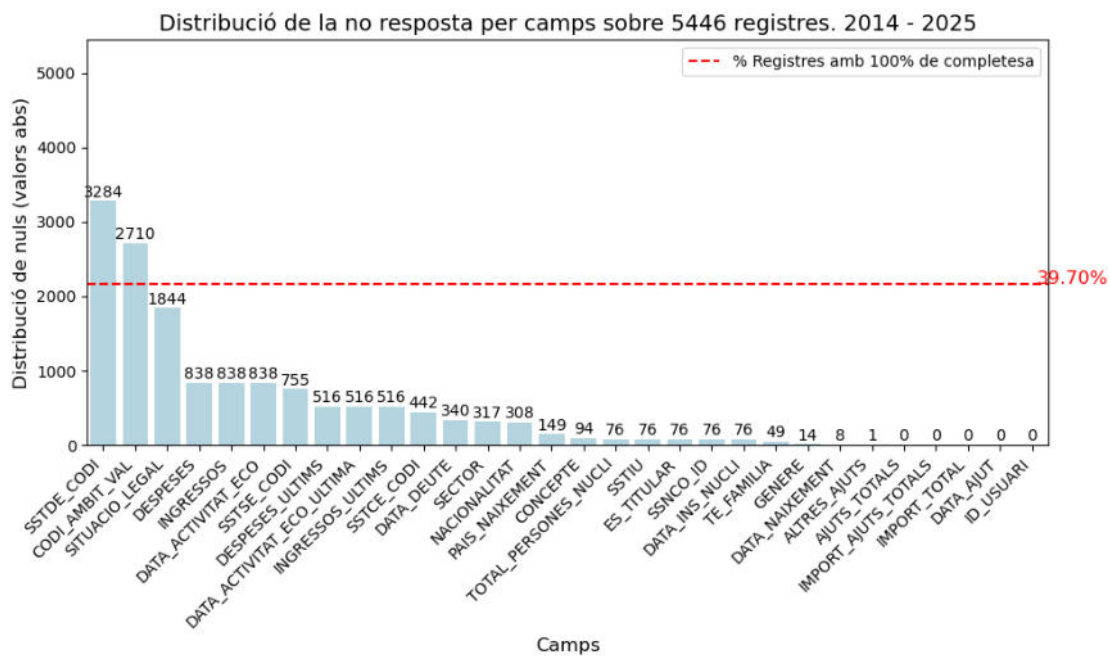
1. **Identificar la data de la valoració de l'activitat econòmica més propera a l'ajut:** Per a cada ajut, es busca la data de la valoració de l'activitat econòmica (DATA_ACTIVITAT_ECO) més propera a la data de l'ajut (DATA_AJUT) que estigui

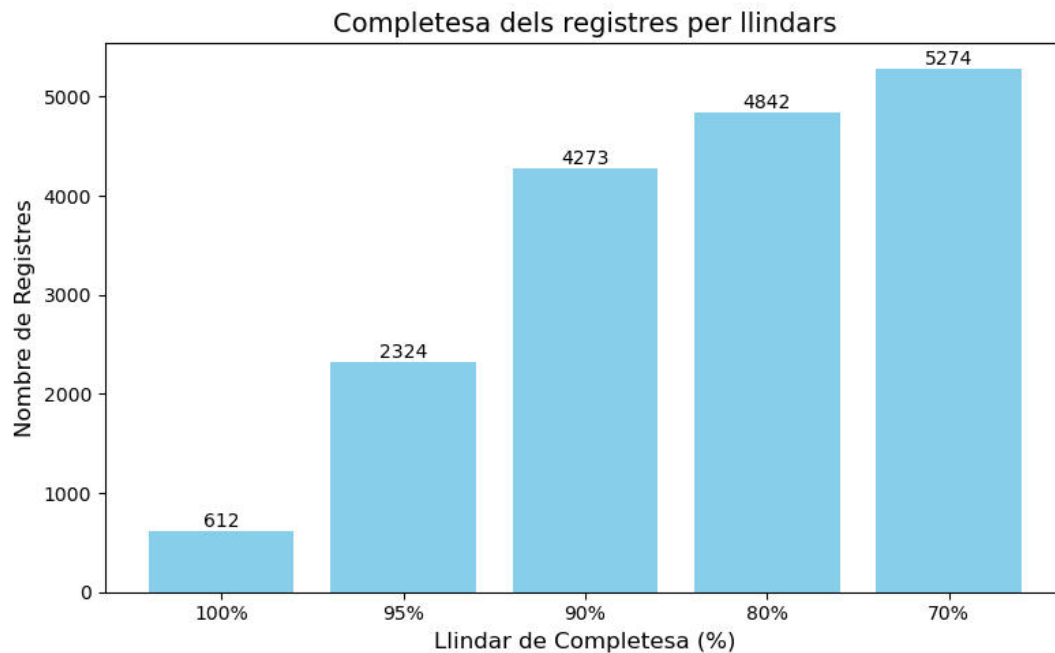
creada abans de demanar l'ajut o en un marge de 60 dies després de la data de l'ajut. Si hi ha diverses dates de valoracions de l'activitat, es selecciona la més recent.

Per aquestes dades també es té en compte l'última informació registrada.

A continuació, veurem l'estat actual de la no resposta.

ESTAT ACTUAL





INTEGRACIÓ DE FONTS

En aquest apartat explicarem com s'han completat els valors buits dels camps utilitzant altres fonts disponibles a l'Ajuntament.

SECTOR

Per tal d'integrar les dades del sector de cada persona s'ha utilitzat l'adreça del domicili el qual està registrada la persona actualment.

REGLES DE CONEIXEMENT

Les regles de coneixement són condicions utilitzades per completar els valors buits dels camps. Cada regla segueix una lògica definida prèviament, basada en criteris que es consideren vàlids i fiables.

CODI DE L'ÀMBIT DE VALORACIÓ

- SITE1: Exclusió econòmica
- SITE2: Vulnerabilitat o risc d'exclusió econòmica
- SITE3: Inclusió econòmica

Per omplir els valors buits de l'àmbit de valoració ens basem en una sèrie de condicions relatives als ingressos, despeses, nombre de persones del nucli familiar i l'IRSC

(Indicador Renda de Suficiència de Catalunya). Si el codi d'àmbit ja està definit, es manté sense modificar. Si no, s'aplica una lògica per determinar-lo.

Els barems de l'IRSC següents son ponderats amb el decret d'habitatge vigent segons la zona de confluència, és a dir, la ciutat o àrea de residència. En el cas de Girona, la zona de confluència és l'A.

- IRSC 1 membre = 2075.97€
- IRSC 2 membres = 2517.12€
- IRSC 3 membres = 2413.32€
- IRSC 4 membres = 2335.47€

A continuació, s'apliquen les regles de coneixement **per ordre de prioritat**:

1. **Si el codi d'àmbit ja està definit:**
 - Es retorna tal com està i no es recalcula.
2. **Casos específics on es retorna SITE1:**
 - Si el concepte de l'ajut és "Emergència habitacional"
 - Si els ingressos són **0**
 - Si els ingressos són menors que les despeses
3. **Assignació de SITE2:**
 - Si els ingressos són menors que l'IRSC per al nombre total de persones del nucli familiar.
 - Si el tipus de situació és que no té xarxa o que la xarxa ajuda parcialment. I el tipus de cobertura és que no té ingressos suficients.
4. **Cap condició compleix:**
 - Es retorna None.

TIPUS DE SUPORT ECONÒMIC

- NTX: No té xarxa
- TXI: Té xarxa i aquesta ajuda parcialment
- TXD: Té xarxa, però es desconeix si aquesta ajuda
- TX: Té xarxa i aquesta ajuda amb totes les despeses

Primer de tot s'han creat dues llistes: La primera llista correspon a aquelles persones que tenen xarxa i aquesta ofereix suport econòmic; la segona llista conté aquelles persones que tenen xarxa (amb influència positiva o neutre), independentment del suport que donen. Per totes dues llistes, les persones d'interès hauran de tenir una data d'inici (DATA_INICI) igual o anterior a la data de l'ajut (DATA_AJUT), i, en cas que hi hagi data de finalització (DATA_FI) ha de ser posterior a la data de l'ajut.

Una vegada s'han creat aquestes dues llistes es tindran en compte les següents regles per ordre de prioritats:

1. Si el camp SSTSE_CODI ja està definit, es retorna sense modificar
2. Si l'usuari no està a la xarxa (xarxa_list), s'assigna el codi 'NTX'
3. Si l'usuari té suport econòmic (suport_eco_list), s'assigna el codi 'TXI'
4. Si el nombre de persones del nucli familiar (TOTAL_PERSONES_NUCLI) és 1 o desconegut, s'assigna el codi 'NTX'.
5. En qualsevol altre cas, s'assigna el codi 'TXD'

TIPUS DE COBERTURA ECONÒMICA

- NI: Sense ingressos
- INNB: Ingressos insuficients per cobrir les necessitats bàsiques
- ICA: Ingressos insuficients per a un desenvolupament adequat
- IS: Ingressos suficients i adequats

Primer de tot, es creen dues llistes:

Els conceptes en vermell són aquells que es donen en un moment determinat en el temps i que, que una vegada finalitzats, ja no es continuen utilitzant.

- **necessitats_bas**: Conté una llista de conceptes d'ajut relacionats amb necessitats bàsiques: "Subsistència", "**Subsistència - Conflicte Ucraïna**", "Habitatge", "Consums llar", "Emergència habitacional", "Fiança", "**SUBSISTÈNCIA MOTIVADA PER COVID19**", "**AJUT TARGETA BON PREU**", "Llum", "Aigua", "Gas", "Llum i Gas combinada", "Allotjament temporal - Conflicte Ucraïna", "Regularització comptadors llum a Girona Est"
- **desenvolupament_adeq**: Conté una llista de conceptes d'ajut relacionats amb desenvolupament adequat: "Salut", "Transport", "Formació", "Programa de Vacances Municipal", "**Ajut esportiu i de lleure Pla Comunitari Construïm Ponts**", "Altres"

A continuació, s'apliquen les regles de coneixement **per ordre de prioritats**:

1. Si SSTCE_CODI ja està definit, es retorna el valor existent.
2. Si la situació econòmica és "TX", s'assigna el valor 'IS'.
3. Si els ingressos són 0, s'assigna el valor 'NI'
4. Si la situació econòmica no és "TX" i el concepte de l'ajut està dins de la llista necessitats_bas, s'assigna el valor 'INNB'
5. Si la situació és "TX" i el concepte de l'ajut està dins de la llista desenvolupament_adeq, s'assigna el valor 'ICA'

TIPUS DE DEUTE

- CP: Corrent de pagament
- ADH: Acumulació deutes habitatge, lloguer
- ADS: Acumulació deutes subministrament
- ADP: Acumulació de deutes de préstecs i targetes bancàries
- ALT: Altres

Per tal d'aplicar les regles de coneixement pel tipus de deute, s'ha creat primer una funció que mostri la tipologia d'ingressos i despeses d'un nucli familiar.

L'únic requeriment per aquesta funció és que la data d'inscripció del nucli ha de ser igual o anterior a la data de l'ajut + 60 dies.

A continuació, s'apliquen les regles de coneixement per ordre de prioritat seguint els següents criteris:

- 1. Si ja existeix un valor en SSTDE_CODI:**
 - Si el codi de deute (deute) no és None, es retorna directament.
- 2. Si els ingressos són menors o iguals a les despeses:**
 - Si l'activitat econòmica inclou el codi "HAB" o el concepte està relacionat amb **habitatge** s'assigna 'ADH'.
 - Si el concepte està relacionat amb **consums de la llar** s'assigna 'ADS'.
 - Si l'activitat econòmica inclou el codi "PRITA" (Préstecs i targetes de crèdit), s'assigna 'ADP'.
 - Si cap de les anteriors és certa, s'assigna 'ALT'.
- 3. Si els ingressos són majors que les despeses:**
 - Assigna 'CP'.
- 4. Si el tipus de cobertura és "IS" i els ingressos són majors que les despeses:**
 - Assigna també 'CP'.
- 5. Altres casos amb condicions addicionals:**
 - Si el codi "HAB" o el concepte de l'ajut és d'habitatge, s'assigna 'ADH'.
 - Si el concepte és de consums de la llar, assigna 'ADS'.
 - Si el codi "PRITA" està present, assigna 'ADP'.
- 6. Per defecte:**
 - Retorna None si no es compleix cap condició.

SITUACIÓ LEGAL

- REGULAR: Regular. Nova categoria afegida
- REG: Autorització de residència
- IRG: Irregular
- RSD: Amb autorització de treball i residència

- TRM: Tràmit
- DPI: Demandant de protecció internacional
- PID: Amb protecció internacional denegada
- DES: Desconegut

Per aplicar les regles de coneixement de situació legal s'ha creat primer una funció per veure si la persona té DNI o NIE. Els criteris d'aquesta funció són els següents:

1. Si comença amb les lletres X, Y o Z, es classifica com a NIE.
2. Si comença amb un dígit i té una longitud exacta de 9 caràcters, es classifica com a DNI.
3. Si no compleix aquestes condicions o no és una cadena, es retorna None.
4. Si no encaixa amb cap de les condicions anteriors, es classifica com a Altres.

Aquesta funció es fa servir per assignar el tipus de document a la columna DOCUMENT a partir de la columna NIF.

A continuació, s'apliquen les regles de coneixement:

Definim la llista de països comunitaris (paisos_comunitaris) que utilitzarem per aplicar les RC.

1. Si la situació legal ja està definida es manté el valor existent.
2. Si no està definida:
 - 'REGULAR':
 - Si la nacionalitat és qualsevol del llistat de paisos_comunitaris.
 - O si el tipus de document és "DNI".
 - 'REG':
 - Si la nacionalitat no es troba dins de paisos_comunitaris i el tipus de document és "NIE".
 - 'IRG':
 - Si la nacionalitat no es troba dins de paisos_comunitaris i el tipus de document és None, però la persona té un passaport
 - O si la nacionalitat no es troba dins de paisos_comunitaris i el tipus de document és "Altres".
 - Si cap de les condicions anteriors es compleix, s'assigna None.

CONCEPTE

Per tal d'omplir els valors buits del concepte de l'ajut s'apliquen les regles de coneixement següents:

1. **Si el concepte ja està definit** es retorna el valor existent sense modificar-lo.

2. Si el concepte no està definit

- Si el codi de deute és "ADH", s'assigna el concepte "Habitatge".
- Si el codi de deute és "ADS", s'assigna el concepte "Consums llar".
- Si el codi de deute és "ADP", s'assigna el concepte "Deutes préstecs".
- Si no es compleix cap de les condicions anteriors, s'assigna el concepte "Altres".

IMPORT DE L'AJUT

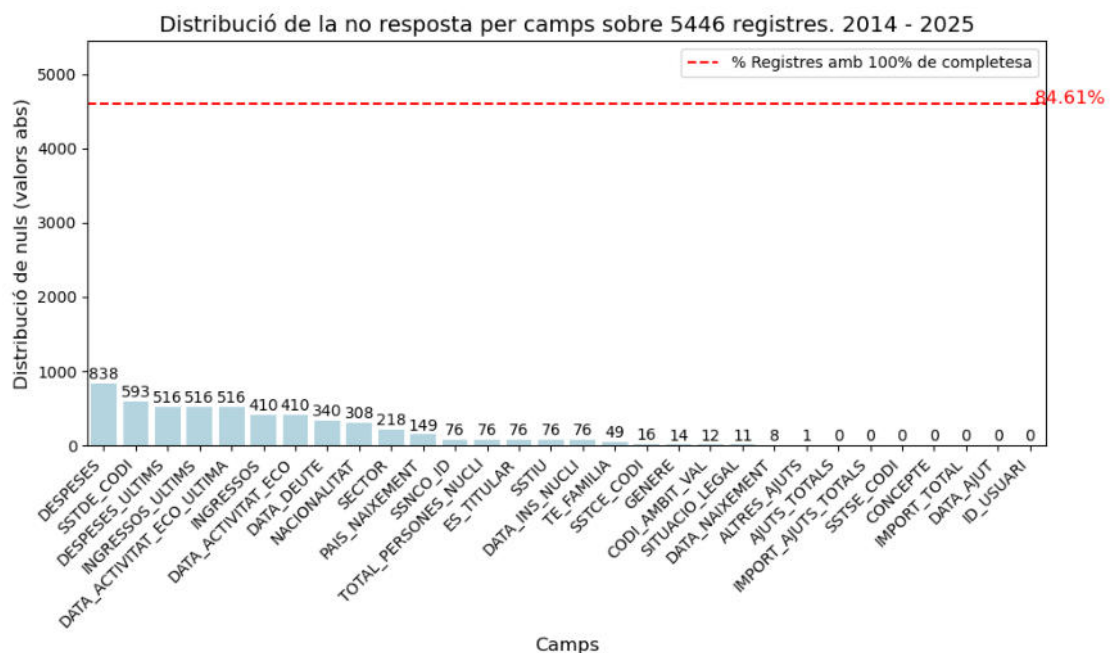
L'import de l'ajut que estigui buit s'afegirà l'import que hi ha a la factura, en cas que hi hagi un proveïdor.

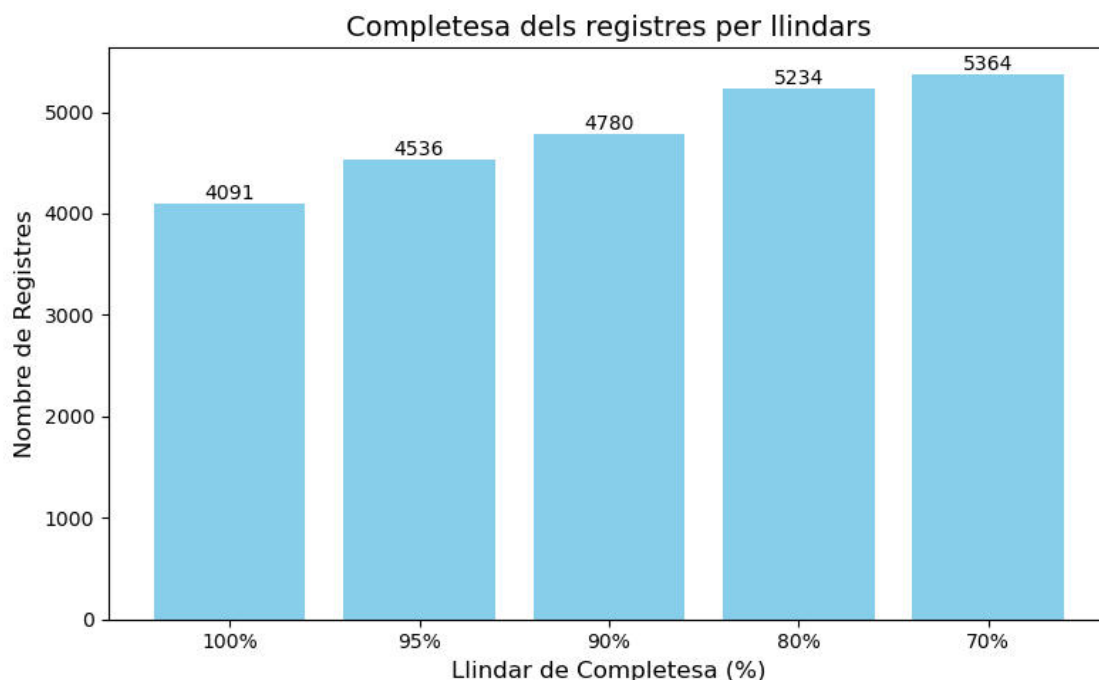
INGRESSOS

Si el tipus de cobertura és "NI" i no hi ha ingressos, s'omple DATA_ACTIVITAT_ECO amb la DATA_DEUTE i s'estableixen els ingressos com a 0.

Seguidament, veurem l'estat de la no resposta, una vegada aplicades les regles de coneixement.

ESTAT POST-REGLES DE CONEIXEMENT





LIMITACIONS EN LA COHERÈNCIA TEMPORAL DE LES DADES

Les diferents fonts que construeixen la taula d'Ajuts Econòmics Puntuals es recullen independentment de si es demana o no un ajut puntual. Aquesta característica genera una certa independència entre les diverses fonts, fet que implica que no tota la informació es mostri sincronitzada en el mateix període de temps en què es realitza la sol·licitud de l'ajut. Tal com s'ha observat, sempre es considera la informació compresa entre la data de l'ajut més 60 dies i la data anterior a l'ajut.

Aquesta metodologia pot donar lloc a situacions en què, per exemple, si un ajut puntual es va sol·licitar el 2018, la informació més propera i disponible de la font de nucli sigui del 2015, mentre que la valoració econòmica correspongui al 2017, etc.

Si ens fixem en les fonts de Valoració Econòmica i Activitat Econòmica, es pot observar que la diferència de dies entre el registre d'una font i l'altra varia. A continuació, es mostra una taula amb exemples:

Diferència de dies	Registres	%
0	2048	37.60%
>= -60 i <= 60	2416	44.36%
<= 60	4255	78.13%

S'observa que la majoria de registres (78.13%) tenen una activitat econòmica registrada en un període comprès entre la data de la valoració econòmica més 60 dies i la data anterior a aquesta valoració. Això indica que les valoracions realitzades han pogut tenir en compte l'activitat econòmica de la persona registrada prèviament o bé que aquesta informació ha estat actualitzada i registrada el mateix dia o dins dels 60 dies posteriors.

2. ÀMBIT LABORAL

L'àmbit laboral contempla informació de tres fonts: Valoració laboral, Nivell de formació i dades personals de les persones.

Per tal de seleccionar les persones ateses que han sigut registrades a l'àmbit laboral s'han tingut en compte la valoració i el nivell de formació.

OBJECTIU DEL DOCUMENT

- Entendre com s'ha construït la taula
- Obtenir una visió general de l'estat de les dades

FONTS

Valoració laboral

La tècnica valora si la persona usuària té la capacitat de treballar, si pot fer-ho ajustant alguns aspectes prèviament o bé, no té la disposició o capacitat per fer-ho.

Aquesta font té històric.

- **DATA_VALORACIO**
- **SSTSL_ID**

Nivell de formació

Es registra el nivell de formació per poder fer una orientació més adequada i saber en quin àmbit laboral es pot inserir.

S'ha agafat la formació més alta a partir de l'ordre de prioritat següent:

- 'DOCT' = 1 // Doctorat
- 'UNI' = 2 // Titulació universitària
- 'BATX' = 3 // Batxillerat
- 'FPMP' = 4 // Formació Professional
- 'GE' = 5 // Graduat Escolar
- 'CEP' = 6 // Certificat estudis primaris
- 'ALTRES' = 7
- 'AN' = 8 // Analfabetisme

Aquesta font té històric.

- **SSNFO_CODI**

CREACIÓ DE LA TAULA

Com que les tres fonts principals tenen històric, el que es fa és filtrar per les dades més actuals. En aquest cas, agafarem de referència les dades de Valoració laboral i després seleccionarem aquell nivell de formació més alt. Per últim, s'englobaran les dades personals de la persona: Gènere, Data de naixement, Sector SBAS, País de naixement, nacionalitat i Situació legal.

INTEGRACIÓ DE FONTS

S'espera que tinguem accés al Padró per poder emplenar els nuls de les dades identificatives.

REGLES DE CONEIXEMENT

SITUACIO_LEGAL

- Nacionalitat = Espanyola o amb DNI ☐ Regular
- Estrangers sense NIE i sí passaport ☐ Irregular
- Estrangers amb NIE ☐ Autorització de residència

3. ÀMBIT RESIDENCIAL

L'àmbit residencial contempla informació de dos fonts: Valoració residencial i dades personals dels nuclis.

Per tal de seleccionar les persones ateses que han sigut registrades a l'àmbit laboral s'han tingut en compte la valoració i el nivell de formació.

OBJECTIU DEL DOCUMENT

- Entendre com s'ha construït la taula
- Obtenir una visió general de l'estat de les dades

FONTS

Valoració residencial

La tècnica valora la situació de la persona en relació a la residència: la persona atesa disposa d'un espai digne per viure, es troba en situació de risc de pèrdua de l'habitatge o habita un espai que no compleix els requisits mínims per considerar-se d'habitatge i l'última situació és si es troba sense habitatge.

Aquesta font té històric.

- DATA_VALORACIO
- SSHAB_ID
- SSTTI_ID
- SSTHA_CODI
- FLG_DESNONAMENT
- FLG_AMUNTEGAMENT
- CODI_AMBIT_VAL
- ES_BAIXA

Dades de nucli

Fa referència a les persones que configuren el nucli familiar de la usuària. Per elaborar la taula s'han realitzat les següents accions: d'aquestes dades s'han agrupat les edats i nacionalitats. De la situació legal s'ha prioritzat la més favorable. I finalment del sector hem escollit el valor que no sigui nul. Aquesta font té històric.

- SSNCO_ID
- EDATS
- NACIONALITATS

- TOTAL_USUARIS
- TIPUS_NUCLI
- SITUACIO_LEGAL
- SECTOR

CREACIÓ DE LA TAULA

Com que les dues fonts principals tenen històric, el que es fa és filtrar per les dades més actuals. En aquest cas, agafarem de referència les dades de Valoració residencial i després la informació més actual del nucli.

Les dades passaran pels següents filtres:

- S'identifiquen aquelles persones que han causat baixa de SIAGI (general) o del domicili i es marquen com a "S" a ES_BAIXA. En cas que tots els membres d'un nucli hagin causat baixa, ES_BAIXA es mostrarà com a "S". Si es dona el cas que hi hagi persones que han causat baixa d'un nucli, només es tindrà en compte la resta de persones que segueixen en el nucli. Aquest últim criteri es tindrà en compte per treballar la resta de camps.
- S'eliminen els usuaris duplicats i tipus test
- S'eliminen els usuaris que compleixen els següents criteris:
 - La persona té una adreça diferent al titular. S'entén que aquesta persona no viu més al domicili.
 - En cas que no hi hagi titular es té com a referència l'adreça de la persona més gran del nucli. Les persones que tinguin una adreça diferent a aquesta persona s'eliminen.
 - En cas que hi hagi dues persones de la mateixa data de naixement que no hi hagi titular, i que tinguin adreça diferent, s'analitzaran un per un. Per aquest cas, no s'ha trobat cap cas que passi aquesta condició.
 - S'eliminen aquells nuclis on cap dels integrants és una persona atesa.

INTEGRACIÓ DE FONTS

S'espera que tinguem accés al Padró per poder emplenar els nuls de les dades identificatives.

REGLES DE CONEIXEMENT

SITUACIO_LEGAL

- Nacionalitat = Espanyola o amb DNI ☐ Regular
- Estrangers sense NIE i sí passaport ☐ Irregular
- Estrangers amb NIE ☐ Autorització de residència

S'ha prioritzat la situació legal més favorable tenint en compte que el número menor és el més favorable

Prioritat situació legal:

priority = {

- 'REGULAR': 1,
- 'REG': 2,
- 'RSD': 3,
- 'TRM': 4,
- 'DPI': 5,
- 'IRG': 6,
- 'PID': 7

4. ÀMBIT EDUCATIU

L'àmbit educatiu contempla informació de quatre fonts: Valoració educativa, Etapa educativa, Centres oberts i dades personals de les persones.

Per tal de seleccionar les persones ateses que han sigut registrades a l'àmbit educatiu s'han tingut en compte l'etapa i la valoració educativa.

OBJECTIU DEL DOCUMENT

- Entendre com s'ha construït la taula
- Obtenir una visió general de l'estat de les dades

FONTS

Etapa educativa

Aquesta font registra les dades sobre l'etapa educativa de l'infant i si s'està fent seguiment educatiu des de SBAS. Aquesta font té històric, atès que es poden registrar consecutivament les diferents etapes educatives de la persona.

- DATA_INICI_ESC
- SSEED_CODI
- FLG_SEGUIMENT

Valoració educativa

Una tècnica de Serveis Socials pot valorar el risc de fracàs escolar de la persona atesa i té en compte aspectes com l'assistència, el material, el rendiment i si existeixen dificultats d'aprenentatge. També es valora si l'acompanyament que s'està fent des de SBAS és o no suficient.

Aquesta font té històric.

- DATA_VALORACIO
- SSUES_ID
- ID_USUARI
- SSTME_ID
- SSTSI_ID
- SSTRE_ID
- FLG_NEE
- FLG_DIFICULTAT
- CODI_AMBIT_VAL

Centres oberts

Les tècniques d'SBAS poden valorar que un infant o adolescent ha d'assistir a Centre Obert per diferents motius, per reforçar l'adquisició d'hàbits, per necessitats educatives o perquè existeix una situació de vulnerabilitat o risc a la llar familiar.

- DATA_CENTRE_OBERT
- SSCEO_ID
- SSMAC_ID

Dades personals

- GERE
- DATA_NAIXEMENT
- SECTOR
- PAIS_NAIXEMENT
- NACIONALITAT
- SITUACIO_LEGAL

CREACIÓ DE LA TAULA

Com que les tres fonts principals tenen històric, el que es fa és filtrar per les dades més actuals. En aquest cas, agafarem de referència les dades d'Etapa educativa i després seleccionarem aquella informació de Valoració educativa i Centre Obert més propera a l'Etapa. Per últim, s'englobaran les dades personals de la persona.

INTEGRACIÓ DE FONTS

S'espera que tinguem accés al Padró per poder emplenar els nuls de les dades identificatives. La única dada que s'ha pogut agafar d'una altre font és el Sector.

El sector s'ha obtingut a partir de les operacions següents:

- S'ha utilitzat la funció NUCLI.OBTENIR_ID_DOMICILI() per així obtenir l'id_domicili de la persona.
- Seguidament s'ha mapejat aquest id amb la taula DOMICILIS per així agafar el CODI_CAR i NUM del domicili i així poder utilitzar la funció PKZONES.OBTENIR_ZONA_TXT() per obtenir el sector.

REGLES DE CONEIXEMENT

SSTRE_ID, SSTSI_ID i SSTME_ID

- Es pot determinar que la no resposta correspon a adequat

CODI_AMBIT_VAL

- Rendiment no adequat + absentisme + no material → fracàs escolar
- Si rendiment no adequat + assistència regular → vulnerabilitat
- Si rendiment adequat + assistència regular + Sí material → procés d'escolarització adequat

SITUACIO_LEGAL

- Menors de 18 anys, s'agafa la millor Situació legal dels pares. La prioritat de situació legal ha sigut aquesta:
 - Regular: 1
 - Autorització de treball: 2,
 - Autorització de residència: 3,
 - Tràmit: 4,
 - Demandant de protecció internacional: 5,
 - Irregular: 6,
 - Protecció Internacional Denegada: 7
- Nacionalitat = Espanyola o amb DNI → Regular
- Estrangers sense NIE i sí passaport → Irregular
- Estrangers amb NIE → Autorització de residència

SSMAC_ID

- Si el menor té FLG_NEE = "S" → Reforçar l'aprenentatge de dèficits detectats.

5. ÀMBIT RELACIONAL

L'àmbit relacional contempla informació de dos fonts: Valoració àmbit relacional i Persones Interès.

Per tal de seleccionar les persones ateses que han sigut registrades a l'àmbit Relacional s'han tingut en compte les dades identificatives de la persona i l'última valoració de l'àmbit relacional registrada.

OBJECTIU DEL DOCUMENT

- Entendre com s'ha construït la taula
- Obtenir una visió general de l'estat de les dades

FONTS

Valoració relacional

La tècnica valora la situació de la persona en relació al seu àmbit relacional: la persona usuària gaudeix d'una xarxa relacional consistent i que cobreix les seves necessitats socials, la seva xarxa és feble i es troba amb necessitats socials no cobertes i finalment no disposa de vincles sòlids pel què es troba en situació d'exclusió en l'àmbit social.

Aquesta font té històric.

- **FLG_FAMILIA**

** Es descarta la informació sobre la valoració final de l'àmbit per l'alt nombre de valors nuls.*

Persones interès

Fa referència a persones amb qui es relaciona la usuària. Es registra quin tipus de vincle i relació manté amb la persona d'interès, com també la data d'inici i finalització (en cas que la relació hagi finalitzat). Aquesta font té històric.

** Es descarta aquesta informació per l'alt nombre de valors nuls.*

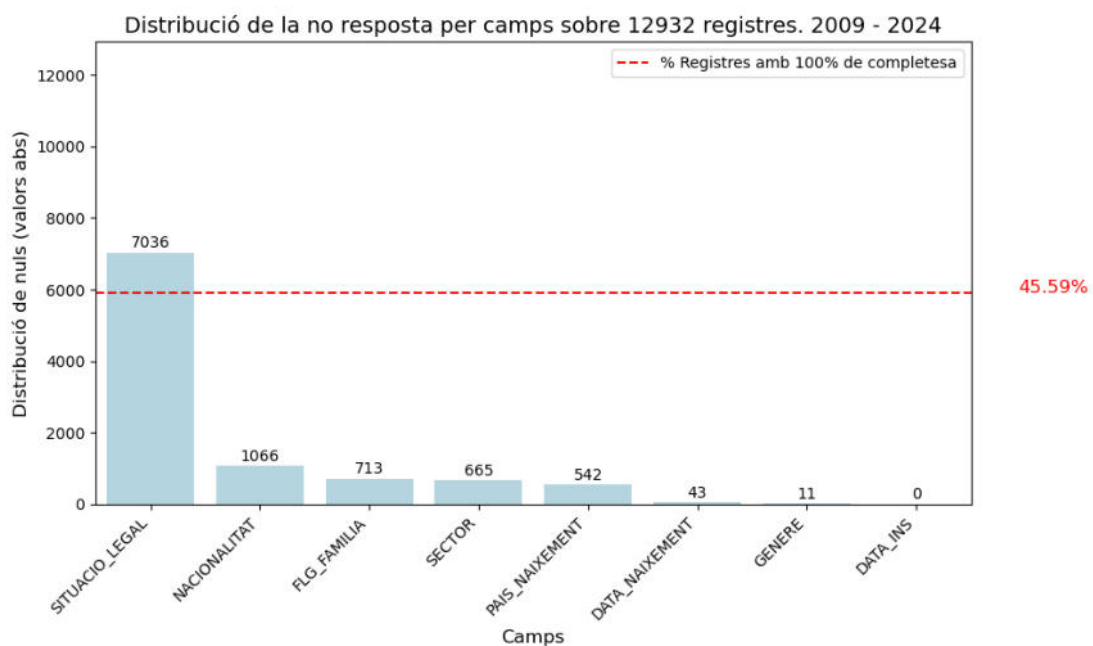
CREACIÓ DE LA TAULA

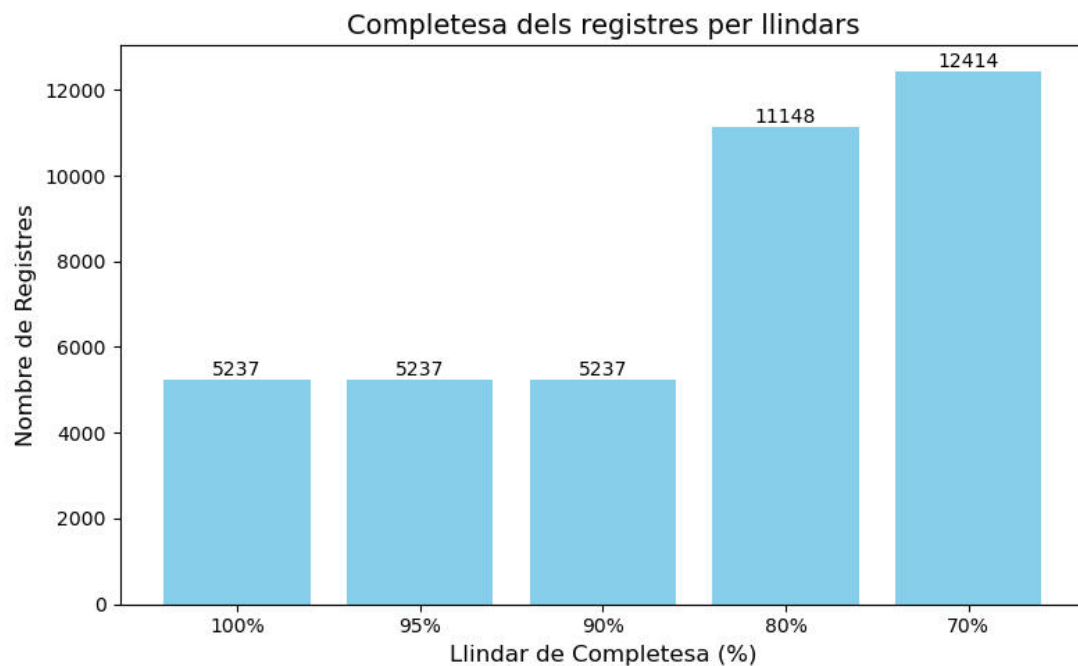
Com que les dues fonts principals tenen històric, el que es fa és filtrar per les dades més actuals. En aquest cas, agafarem de referència les dades de Valoració relacional i després la informació més actual del nucli.

Les dades passaran pels següents filtres:

- S'identifiquen aquelles persones que han causat baixa de SIAGI (general) o del domicili i es marquen com a "S" a ES_BAIXA. En cas que tots els membres d'un nucli hagin causat baixa, ES_BAIXA es mostrarà com a "S". Si es dona el cas que hi hagi persones que han causat baixa d'un nucli, només es tindrà en compte la resta de persones que segueixen en el nucli. Aquest últim criteri es tindrà en compte per treballar la resta de camps.
- S'eliminen els usuaris duplicats i tipus test
- S'eliminen els usuaris que compleixen els següents criteris:
 - La persona té una adreça diferent al titular. S'entén que aquesta persona no viu més al domicili.
 - En cas que no hi hagi titular es té com a referència l'adreça de la persona més gran del nucli. Les persones que tinguin una adreça diferent a aquesta persona s'eliminen.
 - En cas que hi hagi dues persones de la mateixa data de naixement que no hi hagi titular, i que tinguin adreça diferent, s'analitzaran un per un. Per aquest cas, no s'ha trobat cap cas que passi aquesta condició.
 - S'eliminen aquells nuclis on cap dels integrants és una persona atesa.

ESTAT ACTUAL





INTEGRACIÓ DE FONTS

SECTOR

Per tal d'integrar les dades del sector de cada persona s'ha utilitzat l'adreça del domicili el qual està registrada la persona actualment.

S'espera que tinguem accés al Padró per poder emplenar els nuls de les dades identificatives.

REGLES DE CONEIXEMENT

Les regles de coneixement són condicions utilitzades per completar els valors buits dels camps. Cada regla segueix una lògica definida prèviament, basada en criteris que es consideren vàlids i fiables.

FLG_FAMILIA

Aquest camp s'omple a partir de les dades que agafem de la taula SS_PERSONA_RELACIO i SS_TIPUS_VINCLE. En aquestes dues taules tenim les persones i el tipus de vincle de la persona atesa registrada. En cas que hi hagi un familiar, determinem que té família.

SITUACIÓ LEGAL

- REGULAR: Regular. Nova categoria afegida
- REG: Autorització de residència
- IRG: Irregular
- RSD: Amb autorització de treball i residència
- TRM: Tràmit
- DPI: Demandant de protecció internacional
- PID: Amb protecció internacional denegada
- DES: Desconegut

Per aplicar les regles de coneixement de situació legal s'ha creat primer una funció per veure si la persona té DNI o NIE. Els criteris d'aquesta funció són els següents:

5. Si comença amb les lletres X, Y o Z, es classifica com a NIE.
6. Si comença amb un dígit i té una longitud exacta de 9 caràcters, es classifica com a DNI.
7. Si no compleix aquestes condicions o no és una cadena, es retorna None.
8. Si no encaixa amb cap de les condicions anteriors, es classifica com a Altres.

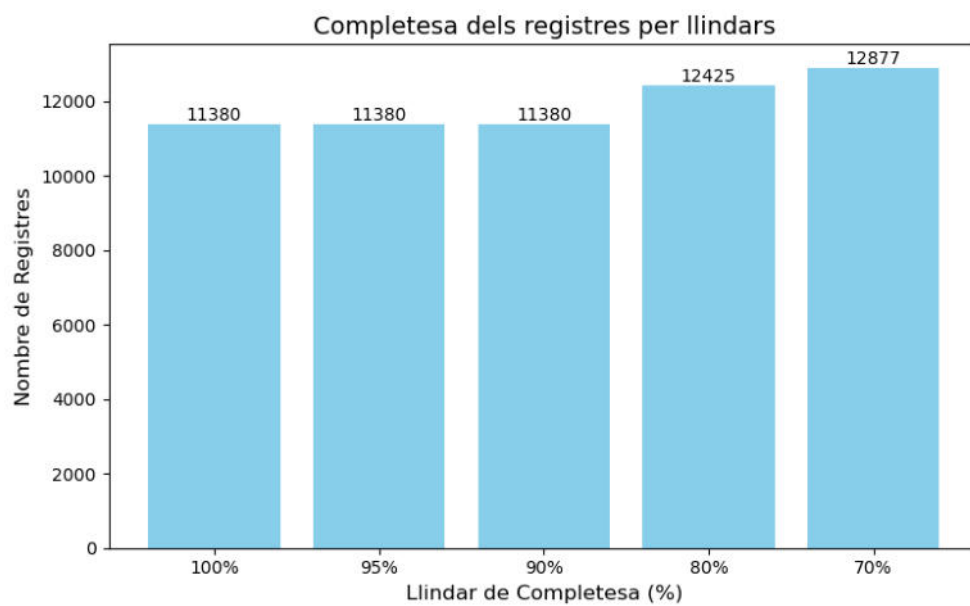
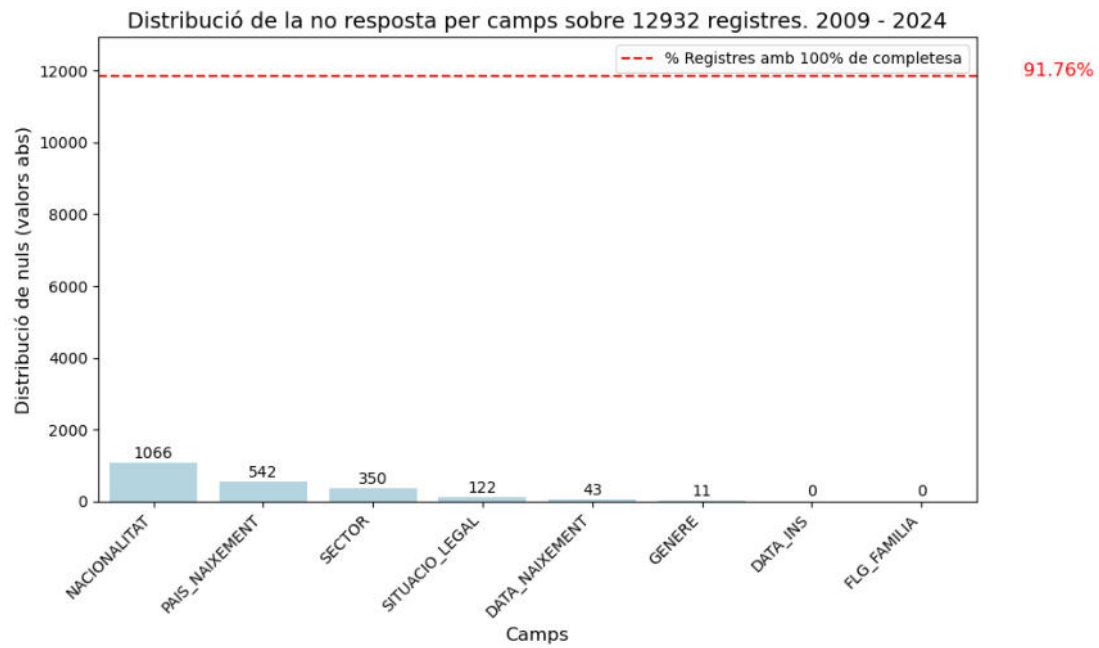
Aquesta funció es fa servir per assignar el tipus de document a la columna DOCUMENT a partir de la columna NIF.

A continuació, s'apliquen les regles de coneixement:

Definim la llista de països comunitaris (paisos_comunitaris) que utilitzarem per aplicar les RC.

3. Si la situació legal ja està definida es manté el valor existent.
4. Si no està definida:
 - **'REGULAR':**
 - Si la nacionalitat és qualsevol del llistat de paisos_comunitaris.
 - O si el tipus de document és "DNI".
 - **'REG':**
 - Si la nacionalitat no es troba dins de paisos_comunitaris i el tipus de document és "NIE".
 - **'IRG':**
 - Si la nacionalitat no es troba dins de paisos_comunitaris i el tipus de document és None, però la persona té un passaport
 - O si la nacionalitat no es troba dins de paisos_comunitaris i el tipus de document és "Altres".
 - Si cap de les condicions anteriors es compleix, s'assigna None.

ESTAT POST-REGLES DE CONEIXEMENT



6. ÀMBIT SALUT EMOCIONAL

L'àmbit Salut contempla informació de dues fonts: Valoració Salut i Dades Identificatives.

Per tal de seleccionar les persones ateses que han sigut registrades a l'àmbit Salut s'han tingut en compte Tipus de Salut de salut, disminució, dependència reconeguda.

OBJECTIU DEL DOCUMENT

- Entendre com s'ha construït la taula
- Obtenir una visió general de l'estat de les dades

FONTS

Valoració Salut

Aquesta font registra les dades sobre l'estat de salut emocional de la persona atesa. La tècnica registra una valoració sobre com afecta l'estat de salut emocional a la estabilitat en la vida de la persona.

Aquesta font té històric, atès que es poden registrar consecutivament els diferents estats o diagnòstics.

- DATA_INICI_ESC
- SSEED_CODI
- FLG_SEGUIMENT

Dades Identificatives

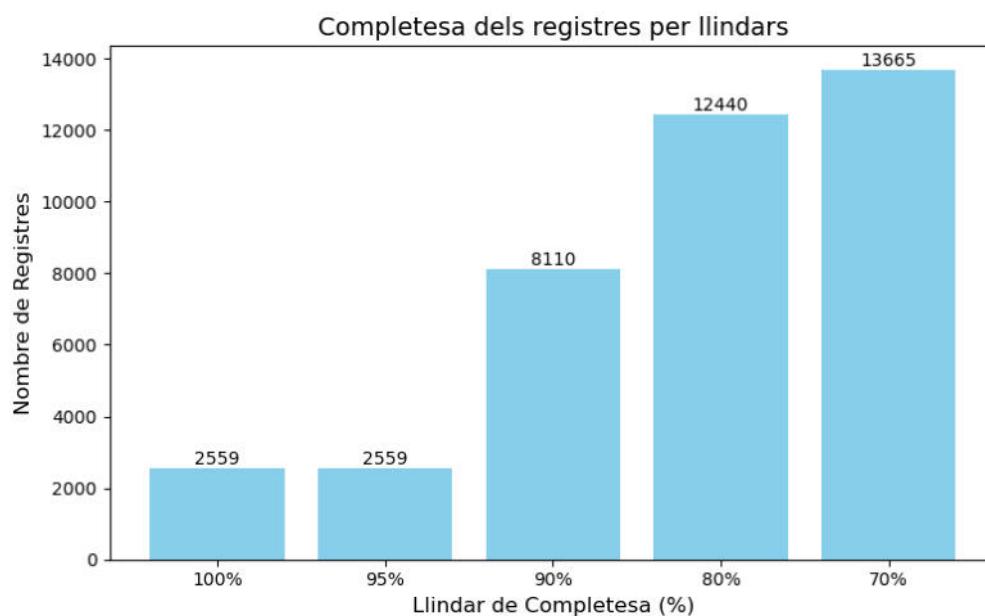
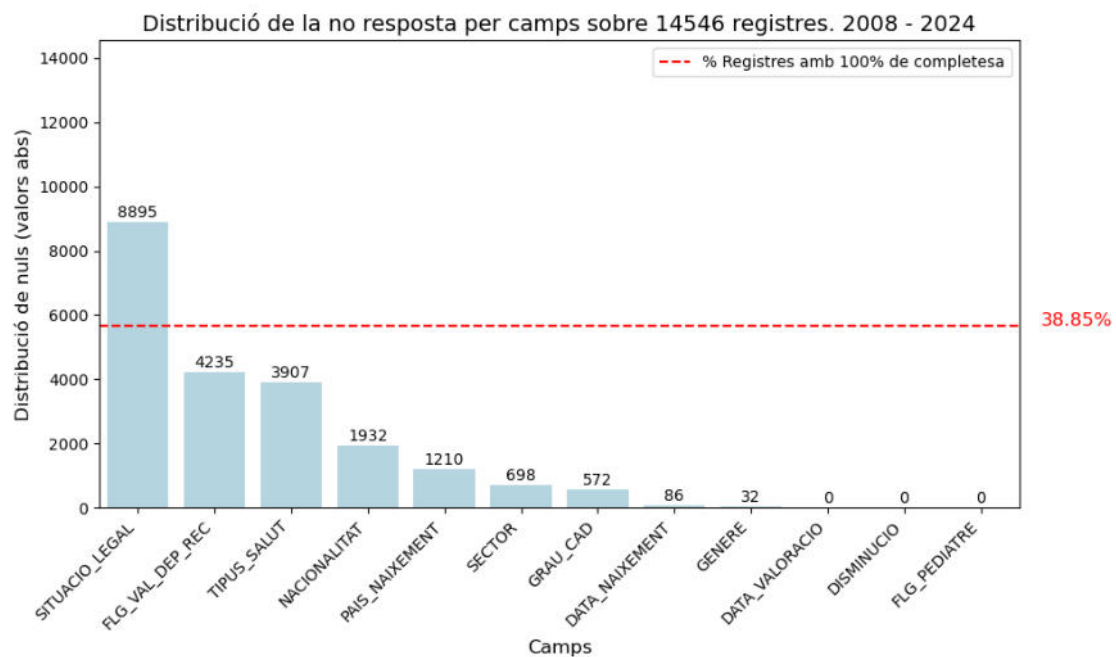
Aquesta font té històric.

- DATA_VALORACIO
- SSUES_ID
- ID_USUARI
- SSTME_ID
- SSTSI_ID
- SSTRE_ID
- FLG_NEE
- FLG_DIFICULTAT
- CODI_AMBIT_VAL

CREACIÓ DE LA TAULA

Com que les dues fonts principals tenen històric, el que es fa és filtrar per les dades més actuals. En aquest cas, agafarem de referència les dades d'Estat de Salut i després seleccionarem aquella informació de valoració de la data més recent. Per últim, s'englobaran les dades personals de la persona: Gènere, Data de naixement, Sector SBAS, País de naixement, nacionalitat i Situació legal.

ESTAT ACTUAL



INTEGRACIÓ DE FONTS

En aquest apartat explicarem com s'han completat els valors buits dels camps utilitzant altres fonts disponibles a l'Ajuntament.

SECTOR

Per tal d'integrar les dades del sector de cada persona s'ha utilitzat l'adreça del domicili el qual està registrada la persona actualment.

REGLES DE CONEIXEMENT

Les regles de coneixement són condicions utilitzades per completar els valors buits dels camps. Cada regla segueix una lògica definida prèviament, basada en criteris que es consideren vàlids i fiables.

S'espera que tinguem accés al Padró per poder emplenar els nuls de les dades identificatives.

TIPUS_SALUT

- Es pot determinar que la no resposta correspon ☐ adequat
- Estabilitat Constant+ segueix tractament ☐ No limitant
- Estabilitat Fluctuant + segueix tractament ☐ Limitant

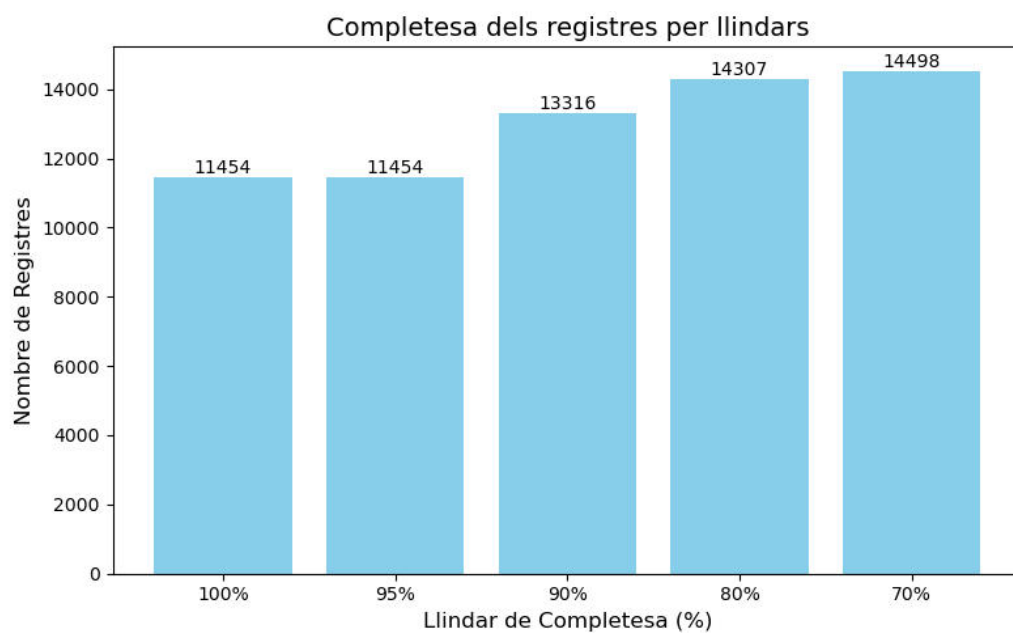
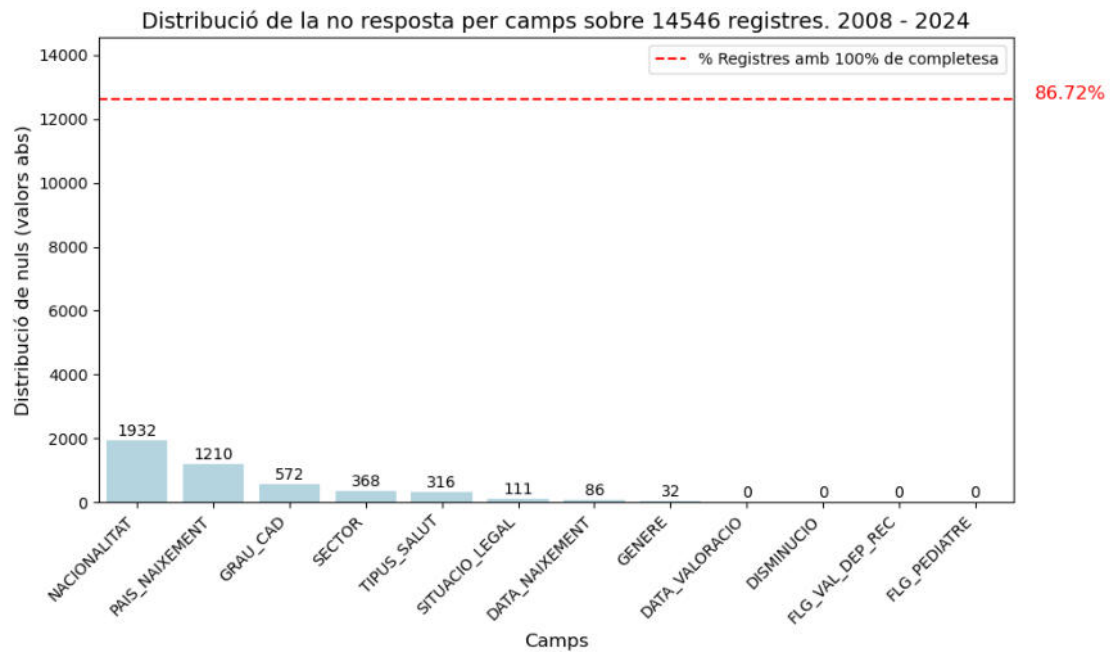
FLG_VAL_DEP_REC

- Si hi ha flag indica que Si hi ha una valoració de dependència
- No hi ha flag es pot inferir que no té dependència

DISMINUCIO

- Nul o no hi ha grau registrat no hi ha disminució

ESTAT POST-REGLES DE CONEIXEMENT



ANNEX 4

CLÚSTERING D'AJUDES ECONÒMIQUES

En aquest annex s'explica tècnicament com s'ha construït l'eina de clústering de perfils d'ajuts econòmics utilitzada per al Pigall.

1. Recollir i netejar dades

A la base de dades de SIAGI hi ha molta més informació de la que es necessita per fer aquest programa. El primer pas és, doncs, fer una selecció dels camps disponibles.

Després de moltes proves, s'ha seleccionat únicament la informació de l'àmbit econòmic i laboral, i s'han aplicat regles de coneixement per tal de completar els valors nuls.

També s'han separat les dades per anys, creant un data set diferent per les dades registrades a cada any des de 2015.

2. Normalització

Els algorismes de clústering, en la seva majoria, només poden treballar amb dades numèriques. Els datasets que s'ha generat són una barreja de dades numèriques (nivell de renda, despeses, etc) i categòriques (franja d'edat, gènere, situació legal, etc). Aquestes dades categòriques s'han de convertir a nombres reals, sense perdre informació rellevant. També cal normalitzar les dades numèriques perquè tinguin una escala similar.

a) Normalització de variables numèriques:

Les variables numèriques poden moure's a escales molt diferents, fet que pot afectar negativament a la precisió de qualsevol model d'aprenentatge automàtic que s'apliqui sobre aquestes dades. Per tant, s'aplica la següent estandardització a sobre de cada columna numèrica:

$$\hat{x} = \frac{x - \mu}{\sigma}$$

a on $\hat{x}$ representa el valor normalitzat d'una instància, x el valor original, μ és la mitjana aritmètica de tota la columna, i σ la desviació típica de la columna.

Amb això es transformen totes les columnes numèriques a una escala equiparable.

b) Normalització de variables categòriques

La normalització de columnes categòriques consisteix en diversos passos.

1. Aplicar *one-hot encoding* a cada columna.
2. Normalitzar cada columna, aplicant

$$\bar{x} = \frac{x}{\sqrt{s/n}}$$

on $\bar{x}$ és el valor normalitzat de cada instància, x és el valor sense normalitzar, s és la suma de tots els valors de la columna, i n és el total d'instàncies de la columna.

- Finalment, centrar les columnes aplicant

$$\hat{x} = \bar{x} - \mu$$

on $\hat{x}$ és el valor normalitzat de cada instància, x és el valor sense normalitzar, s és la suma de tots els valors de la columna, i μ és la mitjana de cada columna.

3. Reducció de dimensionalitat

S'aplica l'algorisme UMAP (Uniform Manifold Approximation and Projection) per tal de reduir-les a 5-10 dimensions (depenent del clústering que acabem fent; vegeu següent pas).

4. Clústering

Finalment, s'aplica l'algorisme HDBSCAN per clusteritzar les dades de cada any.

S'avalua la bondat dels clústers mitjançant dues mètriques:

- Índex de Daves-Bouldin
- Índex de la silueta

Es modifiquen els paràmetres tant de l'HDBSCAN com de l'UMAP per maximitzar el producte entre l'invers de l'índex de Daves Bouldin i l'índex de la silueta.

Per identificar quins clústers es repeteixen en dos anys consecutius, es tracta cada clúster com una funció de probabilitat (amb les probabilitats donades pels percentatges en cada una de les categories) i es computa com de similars són les funcions de probabilitat, usant la mètrica de Hellinger.

5. Explicabilitat

Al haver aplicat UMAP per fer la reducció de dimensionalitat s'ha perdut explicabilitat: els camps "projectats" són una barreja dels originals i ja no corresponen directament als camps originals.

Per millorar l'explicabilitat s'utilitzen dues metodologies complementàries:

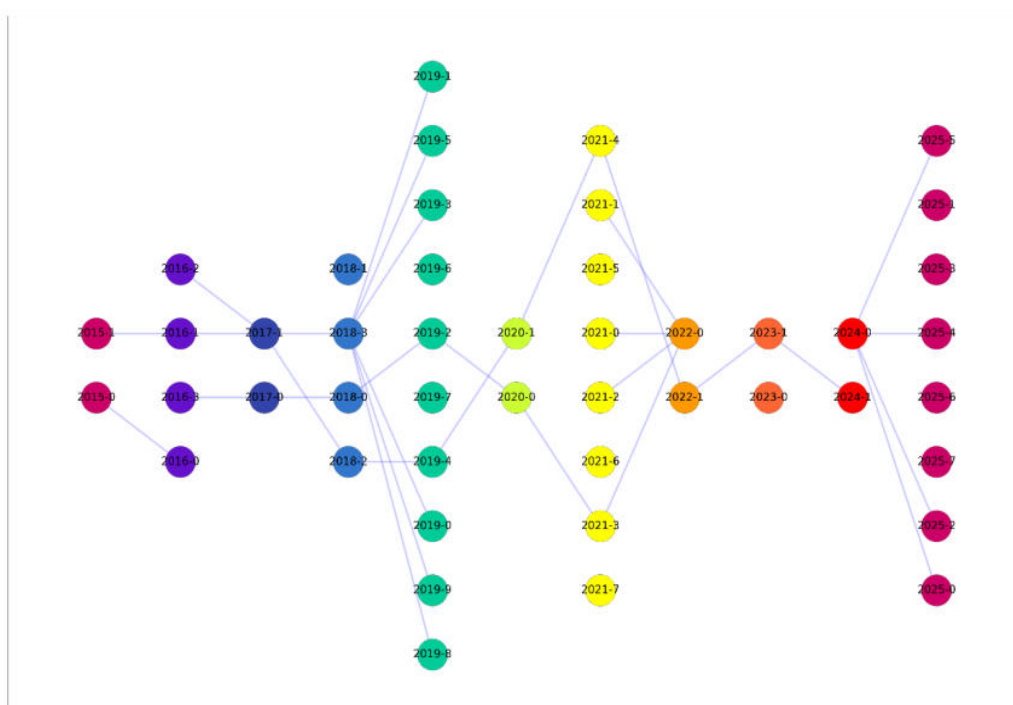
- Anàlisi manual. Es localitzen els elements de cada clúster en el data set original i inspeccionem quines característiques tenen.
- SHAP-values. S'utilitza una xarxa neuronal per explicar els grups.
 - S'entrena la xarxa per, donada una instància del data set normalitzat, predir en quin grup (clúster) es troba.
 - Una precisió alta significa que els clústers trobats són prou diferents i estadísticament significatius.
 - Agafant un exemple d'un clúster particular, pertorbant els inputs, i observant com varia la classificació, es poden veure quines variables són més significatives

a cada clúster. Aquest fet proporciona informació molt rellevant de les variables més significatives de cada clúster i, per tant, pot ajudar a explicar-lo.

Exemple

A continuació, s'exposa un exemple d'informe de clústering sobre ajuts econòmics. Es tracta de l'informe de clústering relatiu a les característiques de les persones que han sol·licitat ajuts econòmics l'any 2025.

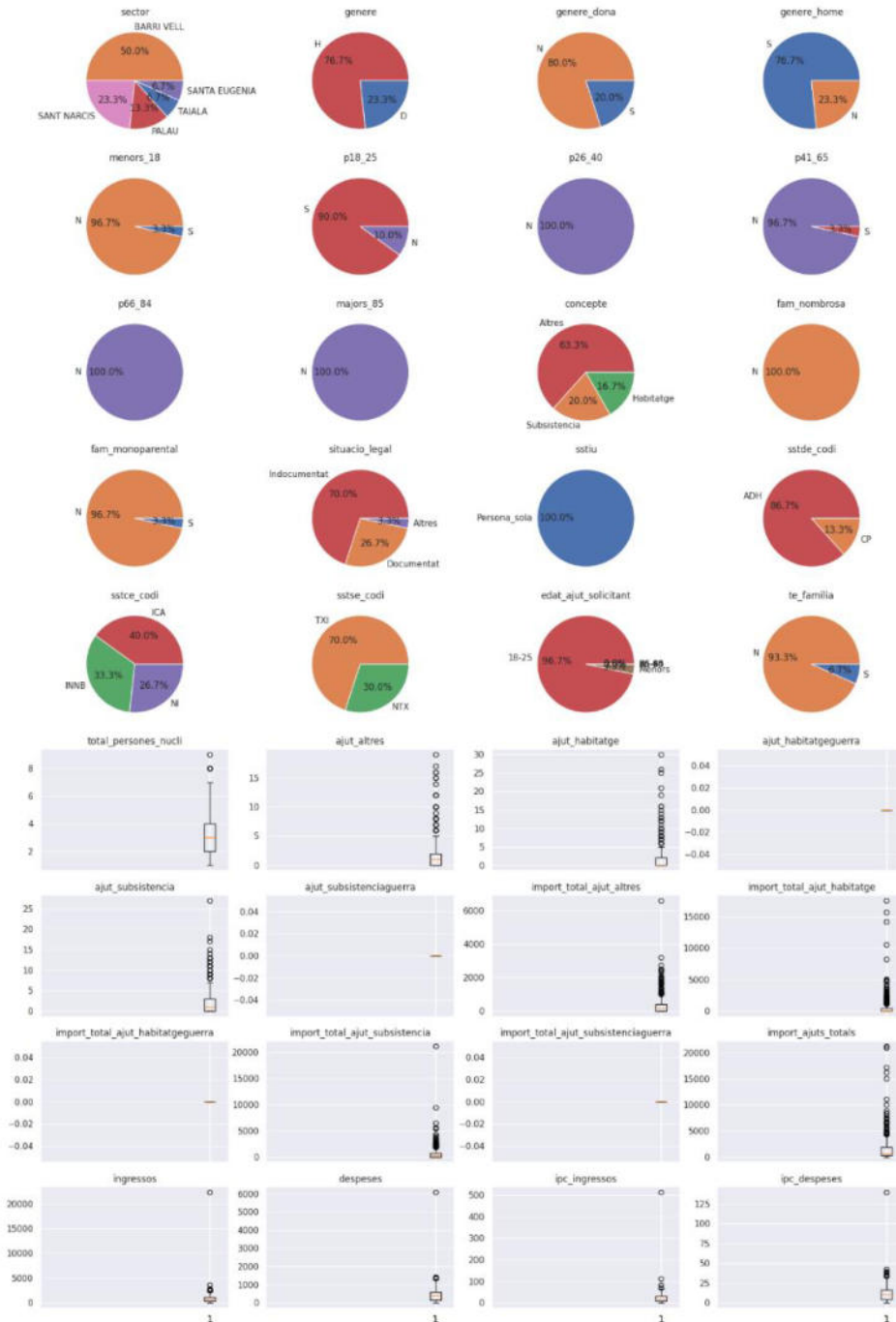
En primer lloc es mostra l'evolució dels grups en el transcurs dels anys:



Cada una de les columnes correspon a un any, essent la primera columna relativa a l'any 2014 i la darrera a l'any 2025. Si mirem en concret l'any 2025 comprovem que, si bé l'any 2024 les persones beneficiàries d'ajut s'agrupaven en dos grups principals que compartien característiques, l'any 2025 aquest dos grups s'han desglossat en 8. Les línies en diagonal identifiquen en quants grups s'ha disgregat cada un dels grups de l'any 2024.

Un cop identificat el comportament global dels grups respecte a l'any anterior, l'informe mostra les característiques de cada un dels grups l'any 2025. A tall d'exemple, es mostra l'informe relatiu al grup 1:

Any 2025 - Grup [1]



Veient els gràfics, es pot observar que en el Grup [1] de l'any 2025 es té que:

- La majoria són homes.
- La majoria de persones han demanat altres ajuts econòmics.
- El 25% o menys han demanat ajuts per subsistència.
- El 25% o menys han demanat ajuts per habitatge.
- El 25% o menys són famílies monoparentals.

Es pot comprovar que la característica principal que determina aquest grup és que són majoritàriament homes (76,7%), joves (90,0% tenen entre 18 i 25 anys) i que viuen sols (Persona sola 100%). De cada un dels grups generats se'n pot extreure una lectura similar si s'observen les seves gràfiques (per exemple, un altre grup podria ser el de les dones joves amb fills a càrrec, altre podria ser el grup de persones grans en situació de pobresa energètica, etc).

El clústering ens permet l'anàlisi de les característiques de les persones que sol·liciten un ajut econòmic d'urgència. Informalment, es pot considerar que cada clúster mostra una "fotografia" de quines són les casuístiques principals de cada any.

ANNEX 5

AVALUACIÓ ÈTICA I LEGAL DEL PROJECTE PIGALL A PARTIR DEL MODEL PIO

Observatori d'Ètica en Intel·ligència Artificial de Catalunya

Universitat de Girona

Introducció

El projecte Pigall consisteix en la creació d'una eina basada en el desenvolupament de tecnologies d'Intel·ligència Artificial (IA) i de sistemes de dades i indicadors que han de permetre disposar d'informació fiable i actualitzada per a la millora de la planificació de la intervenció social i la presa de decisions a l'àrea de Serveis Socials de l'Ajuntament de Girona. Això, i la capacitat dels equips tècnics, ha de suposar una millora la qualitat de l'atenció, amb criteris d'eficàcia, eficiència, justícia i inclusió posant la vida i la persona en el centre.

Per tant, el Projecte Pigall té el focus en l'ús responsable de la IA per als serveis socials en un context local que és el de la ciutat de Girona. L'ús va precedir d'un desenvolupament a càrrec de l'Ajuntament de Girona amb l'assessorament de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC) i el grup de recerca eXiT de la Universitat de Girona. El Projecte Pigall està inclòs en el Pla de recuperació, transformació i resiliència (PRTR), i està finançat per la Unió Europea – NextGenerationEU.

Context i resum de l'avaluació

Mitjançant l'explotació de dades pròpies (com les del sistema de registre SIAGI) i la combinació amb fonts externes, el Projecte Pigall ofereix funcionalitats de visualització (quadre de comandament), predicció de riscos (com ara la pèrdua d'habitatge o la vulnerabilitat en gent gran), generació d'avisos automàtics i automatització de tasques mecàniques, tot plegat dissenyat sota els principis d'IA responsable per actuar com a suport a la presa de decisions. Concretament, el projecte inclou:

- Integració de dades: Disposar d'un conjunt de dades a partir del qual es puguin desenvolupar tasques de suport a la presa de decisions.
- Anàlisi descriptiu: Conèixer la 'fotografia' de la provisió de serveis socials actual i la demanda.
- Monitorització: Definir identificadors claus per garantir una provisió de serveis satisfactòria (o actual) i detectar desviaments per corregir-los el més aviat possible.

- Predicció: anticipar escenaris de futur pel que fa a context sociodemogràfic i socioeconòmic i possibles situacions a afrontar des dels serveis socials. Conjunt d'eines que relacionant grans quantitats de dades i partint d'algoritmes ens mostren una predicció/pronòstic que ha de ser útil per prendre decisions i planificar polítiques i accions socials.
- Optimització: de processos i recursos.
- Creació d'un Simulador what-if (desenvolupament d'un simulador).

En aquest sentit, el Projecte Pigall representa un exercici exemplar en la implementació de la IA a l'administració pública i, particularment, en l'àmbit dels serveis socials. Des del seu inici, el Projecte Pigall ha estat concebut i desenvolupat amb un enfocament *ex ante* d'IA responsable, o sigui, integrant els principis ètics i legals com a eixos fonamentals en el disseny i consecució del mateix. La col·laboració entre professionals de l'àmbit dels serveis socials, de persones expertes en dades i intel·ligència artificial, així com de persones amb un coneixement en ètica aplicada, ha permès que els objectius i les eines concebudes i desplegades siguin no tan sols tècnicament sòlides, sinó socialment coherents per als serveis socials i, és clar, permeses des d'un punt de vista normatiu.

Aquest document detalla, a partir del Model PIO sobre Principis, Indicadors i Observables de l'Observatori d'Ètica en Intel·ligència Artificial de Catalunya (OEIAC), quin és el grau d'adequació i compliment des d'una perspectiva ètica i legal del Projecte Pigall. Per fer-ho, l'avaluació es basa en els 7 principis del mateix model i en els temes i subapartats que s'inclouen en el document "El Model PIO (Principis, Indicadors i Observables) sobre Obligacions i Drets: un enfocament d'autoavaluació per al compliment de la regulació en dades i sistemes d'intel·ligència artificial en el marc de la Unió Europea". Podeu consultar el model i el document a través de la pàgina web de l'OEIAC [<https://oeiac.cat/>].

A continuació, es presenta una síntesi de l'avaluació del Projecte Pigall per a cadascun dels set principis, estructurant les conclusions principals.

Principi 1: Transparència i explicabilitat

El Projecte Pigall demostra un compromís exemplar amb la transparència. La memòria del projecte documenta exhaustivament el procés de desenvolupament, l'arquitectura tecnològica, la procedència i qualitat de les dades, i les limitacions dels models predictius. Les funcionalitats de l'eina, els seus objectius i les decisions de disseny són públiques i reproduïbles, complint amb escriure els requisits dels articles 11, 13 i 53 del Reglament d'Intel·ligència Artificial (RIA).

Principi 2: Justícia i equitat

La justícia i equitat han estat incorporades *ex ante* en el disseny. La col·laboració multidisciplinària entre tècnics socials, experts en dades i en ètica ha estat clau per identificar i mitigar possibles biaixos cognitius, estadístics i de classificació. L'ús de regles de coneixement per tractar valors nuls i la validació contínua dels resultats per part de professionals socials asseguren que l'eina no reproduceix discriminacions estructurals. A més, el seu disseny com a

eina de suport a la decisió, i no de substitució, prevé el biaix d'automatització. Aquest element està contemplat en l'article 10 del Conveni marc del Consell d'Europa.

Principi 3: Seguretat i no maleficència

El principi de no maleficència és el pilar fonamental del projecte. La decisió explícita que l'eina no pugui prendre decisions de manera autònoma és la màxima garantia que no causarà dany directe. La supervisió per part de persones és constant i s'ha materialitzat en la validació de cada fase per part dels usuaris finals i els comitès d'experts. S'han avaluat els riscos potencials (per exemple, ús de dades no representatives) i s'han implementat mesures per mitigar-los, complint així amb els articles 9, 14 i 15 del RIA.

Principi 4: Responsabilitat i retiment de comptes

La responsabilitat recau clarament en l'Ajuntament de Girona com a entitat pública titular del projecte. La metodologia de treball, basada en la recerca-acció participativa i la col·laboració amb la UdG, estableix un marc de retiment de comptes robust davant la ciutadania. La documentació completa del procés i la realització d'una Anàlisi d'Impacte en la Protecció de Dades (AIPD) faciliten la demostració del compliment normatiu. L'eina està dissenyada perquè els seus usuaris tinguin un control total i puguin revertir qualsevol acció, en línia amb els articles 14 i 17 del RIA.

Principi 5: Privacitat

La privacitat és tractada amb el màxim rigor ja que, entre altres, s'ha realitzat una AIPD preceptiva i s'han aplicat mesures de protecció des del disseny, com ara l'agregació de dades a nivell d'"illa" per a les consultes del Padró o l'anonimització de les dades per a la seva descàrrega. L'ús d'infraestructures internes i segures (PostgreSQL, Oracle) garanteix la integritat i confidencialitat de les dades personals, complint estrictament amb el Reglament General de Protecció de Dades (RGPD) i els articles 10 i 27 del RIA.

Principi 6: Autonomia

La IA s'utilitza per alliberar els tècnics de tasques mecàniques i proporcionar-los informació agregada i prediccions, per tal d'augmentar la seva capacitat per prendre decisions informades i dedicar-se a les tasques com el treball comunitari i els vincles personals. L'eina, per tant, no substitueix ni limita l'autonomia professional, sinó que la potencia, tal com promou l'article 7 del Conveni marc del Consell d'Europa.

Principi 7: Sostenibilitat

El projecte aborda diferents aspectes. Primer, la sostenibilitat institucional, ja que l'eina es desenvolupada íntegrament amb recursos propis, sense dependències del sector privat, la qual cosa en garanteix la perdurabilitat i el manteniment.

Segon, la sostenibilitat social, ja que en millorar l'eficiència en la planificació i l'assignació de recursos, contribueix a una gestió més justa i responsable dels serveis socials. Tercer, cal subratllar que el Projecte Pigall no inclou la sostenibilitat ambiental (petjada de carboni dels

servidors, eficiència energètica dels models), i aquest és un aspecte a considerar en futures iteracions d'acord amb l'article 95 del RIA i les recomanacions de la UNESCO.

Avaluació detallada a través del Model PIO

Principi 1: Transparència i explicabilitat

L'equip del Projecte Pigall mostra un compromís molt elevat amb la transparència, entesa tant en la seva vessant tècnica (com funciona l'eina) com en la seva vessant organitzativa (qui l'utilitza i per què). La memòria reflecteix un esforç conscient per documentar i comunicar el procés de desenvolupament, les dades utilitzades i les limitacions de l'eina, alineant-se amb el RIA en matèria de transparència i per a sistemes d'alt risc (els quals, en l'àmbit social, ho són). Més concretament, tenint en compte les preguntes clau del Model PIO, podem assenyalar el següent:

-Anàlisi preliminar de l'ús del sistema:

1.1.1 (Documentació i reproductibilitat): Complert en alt grau. La memòria documenta de manera exhaustiva el procés de generació de resultats (models de clustering, Random Forest, detecció de canvis de tendència) i l'arquitectura tecnològica (APEX, PostgreSQL, Airflow). Aquesta documentació permet la reproductibilitat i el descobriment de problemes futurs.

1.1.2 (Precisió de l'ús): Complert. Es defineixen amb claredat les quatre finalitats principals de l'eina: Anàlisi, Planificació, Predicció i Automatització.

1.1.5 (Identitat del proveïdor): Complert. Es tracta d'un projecte de l'Ajuntament de Girona, amb la col·laboració explícita de la UdG (OEIAC i grup eXiT), que actuen com a assessors. La responsabilitat recau clarament en l'ens públic.

1.1.6 (Alfabetització i accessibilitat): Complert o en procés. La memòria esmenta la formació als usuaris i la creació d'un manual. La participació dels tècnics en el testeig és una forma de garantir que l'eina s'ajusta a les seves competències.

La creació d'eines de consulta per simplificar tasques també contribueix a la usabilitat.

Explicabilitat:

1.2.1 i 1.2.2 (Documentació pública del procés i informació tècnica): Complert en alt grau. La memòria és una peça central de documentació. Explica el perquè del projecte, les persones implicades (equips, àrees, UdG), els processos de recollida de dades i les decisions tècniques preses (ús de SIAGI, regles de coneixement, models).

1.2.3 (Instruccions clares i accessibles): Complert o en procés. Es menciona la creació d'un manual d'usuari i icones d'informació a la interfície, que són bones pràctiques. La formació directa a caps de secció i direccions reforça aquesta comprensió.

1.2.5 i 1.2.6 (Capacitat d'explicar els resultats i les limitacions): Complert en alt grau. La memòria és molt transparent sobre aquest punt. Explica els models utilitzats (Random Forest, HDBSCAN) i, molt important, admet explícitament les limitacions, com ara la manca de disponibilitat de

dades externes que va condicionar el producte final. Es reconeix que la precisió del model de risc d'habitatge és del 94%, establint un nivell d'expectativa realista.

Traçabilitat i representativitat:

1.3.1 i 1.3.2 (Representativitat i documentació de les dades): Complert amb anàlisi crítica. Es va realitzar un anàlisi profund de la qualitat de les dades de SIAGI (exactitud, integritat, etc.), identificant duplicats i valors nuls. Es van crear "regles de coneixement" per tractar aquests nuls, un esforç notable per millorar la representativitat. La decisió d'usar el nucli de convivència com a unitat d'anàlisi és un exemple de tractament intel·ligent de les dades. Es documenta la procedència (SIAGI, Padró, fonts externes).

1.3.4 (Validació de dades demogràfiques): Complert. Es va optar per utilitzar la "illa" com a màxim nivell de concreció per a les dades del Padró, una decisió de compromís entre granularitat i protecció de dades que demostra un bon ús del context.

Protocols de comunicació:

1.4.1 (Eines per a comentaris d'usuaris): Complert. El procés de recerca-acció participativa i les reunions mensuals amb l'equip assessor són mecanismes robustos per recollir comentaris. El testeig amb les directores d'SBAS i SAGGID ho consolida.

1.4.4 (Sistema de notificació d'incidents): Complert. La funcionalitat d'"Avisos" està dissenyada per notificar anomalies (canvis de tendència) i situacions de risc (pèrdua d'habitatge, vulnerabilitat en gent gran), actuant com a sistema d'alertes integrat.

Principi 2: Justícia i equitat

El projecte aborda la justícia i equitat de manera proactiva, no com un requisit posterior. La col·laboració amb l'OEIAC i el grup eXIT ha assegurat que aquests principis estiguessin presents des del disseny, buscant prevenir biaixos i fomentar la inclusió, tal com promou el RIA per a sistemes d'alt risc. Més concretament, tenint en compte les preguntes clau del Model PIO, podem assenyalar el següent:

Inclusió i diversitat:

2.1.1 (Mesures d'inclusió i equitat): Complert. L'eina es va dissenyar per millorar la planificació i la presa de decisions basada en evidències. En centrar-se en l'anàlisi de dades agregades i prediccions, busca una distribució més equitativa dels recursos. El fet de no prendre decisions individuals automàtiques és la mesura d'inclusió més important.

2.1.4, 2.1.5, 2.1.6 (Diversitat en l'equip de treball): Complert. La memòria destaca explícitament el "diàleg transversal entre professionals socials i tecnòlegs" com un dels punts forts. Aquest equip multidisciplinari (diferents disciplines, gènere, i àmbits de coneixement) és una garantia contra els punts cecs en el disseny.

2.1.9 (Mesures per a persones vulnerables): Complert. Les funcionalitats d'"Avisos" estan específicament orientades a detectar riscos per a col·lectius vulnerables com ara persones grans en situació de soledat o famílies en risc de pèrdua d'habitatge.

Identificació i mitigació de biaixos:

2.2.1, 2.2.2, 2.2.3 (Mitigació de biaixos estadístics, de classificació i cognitius): Complert en alt grau. El projecte demostra una consciència molt alta d'aquests biaixos. L'anàlisi de la qualitat de les dades de SIAGI i la creació de "regles de coneixement" per tractar valors nuls són mesures directes per mitigar biaixos estadístics.

La participació activa dels tècnics socials en la definició de les variables i en la validació dels resultats és la principal mesura per mitigar els biaixos cognitius i de classificació, assegurant que els models no aprenguin patrons esbiaixats pels registres històrics.

2.2.4 (Avaluació d'impacte de biaixos en col·lectius vulnerables): Complert. L'avaluació es va fer de manera continua, no com un punt aïllat. El fet que els models es testeessin i validessin per tècnics socials experts en els col·lectius afectats (gent gran, famílies amb risc d'habitatge) demostra una avaluació pràctica de l'impacte potencial dels biaixos.

2.2.9 (Advertència contra el biaix d'automatització): Complert. La memòria és clara en un punt crucial: l'eina és de suport a la presa de decisions, no substitutiva. L'enfocament metodològic i el disseny de l'eina, que requereix la interpretació humana dels avisos i la validació dels resultats, és la millor defensa contra aquest biaix.

Principi 3: Seguretat i no maleficència

El principi de no maleficència (no causar dany) està integrat en el nucli del projecte. La decisió de dissenyar l'eina com a suport a la presa de decisions en l'àmbit dels serveis socials, evitant qualsevol automatització que pugui perjudicar directament les persones, és la màxima expressió d'aquest principi que s'alinea amb el RIA. Més concretament, tenint en compte les preguntes clau del Model PIO, podem assenyalar el següent:

Funcionament del sistema:

3.1.2 i 3.1.3 (Informació sobre precisió i influència de l'usuari): Complert. Es comunica la precisió del model de risc d'habitatge (94%). La naturalesa de l'eina, que és consultiva i no autònoma, fa que la influència de l'usuari sigui total: l'eina no actua, només informa.

3.1.4 (Mecanismes de supervisió i control): Complert. La supervisió és el pilar del projecte. La validació constant per part dels tècnics socials i els assessors de la UdG (OEIAC) actua com un mecanisme de control intern i extern robust.

3.1.5 (Coneixement de comportaments prohibits): Complert. L'eina està dissenyada per no poder prendre decisions individuals. Està "en gàbia de Faraday" des d'aquest punt de vista, incomplint així el seu propi propòsit si intentés fer-ho.

Traçabilitat per a la Gestió del Risc:

3.2.1, 3.2.4, 3.2.5, 3.2.7 (Avaluació d'impactes i gestió de riscos): Complert en alt grau. El projecte ha realitzat una avaluació d'impacte sobre la salut, la seguretat i els drets de les persones, no només en la fase de disseny, sinó mitjançant la validació contínua amb professionals. Es van identificar riscos (com l'ús de dades no representatives) i es van prendre

mesures específiques per mitigar-los (regles de coneixement). L'avaluació de prediccions inexactes és inherent a la validació del model.

3.2.9 (Avaluació que les dades no es puguin utilitzar per discriminar): Complert. Aquesta és una de les principals preocupacions del projecte. L'equip ha treballat activament per evitar-ho mitjançant la governança de dades i la supervisió humana. L'ús de models interpretables i la transparència en els criteris són defenses addicionals.

Principi 4: Responsabilitat i retiment de comptes

El projecte exemplifica la responsabilitat pública. La titularitat de l'Ajuntament, la col·laboració acadèmica i la metodologia participativa estableixen un marc clar de retiment de comptes davant la ciutadania, complint així amb el RIA. Més concretament, tenint en compte les preguntes clau del Model PIO, podem assenyalar el següent:

Control:

4.1.1 i 4.1.5 (Comprensió del control i capacitació): Complert. La participació dels tècnics en el disseny i la formació planificada asseguren que les persones que supervisen l'eina la comprenen adequadament.

4.1.3 i 4.1.4 (Eines per intervenir i revertir decisions): Complert per disseny. L'eina no pren decisions, per tant, no hi ha cap decisió automàtica que revertir. Tota actuació final queda en mans dels tècnics. És una decisió de disseny que reforça el control humà.

Retiment de comptes:

4.2.1 (Determinació de la responsabilitat legal): Complert. La responsabilitat recau clarament en l'Ajuntament de Girona com a entitat pública que desenvolupa i utilitza l'eina per a la gestió dels serveis socials.

4.2.2 i 4.2.3 (Avaluació de la influència en la presa de decisions i revisió periòdica): Complert. La memòria és una avaluació de la influència que l'eina tindrà. L'objectiu explícit és "facilitar la presa de decisions basada en evidències". Les reunions periòdiques i els tallers amb tècnics són mecanismes de revisió contínua.

Figures de protecció:

4.3.1 (Demostració del compliment legal): Complert. La realització d'una AIPD (Anàlisi d'Impacte en la Protecció de Dades) i la documentació exhaustiva del projecte (aquesta memòria) són les principals eines per demostrar el compliment, especialment del GDPR.

4.3.3 (Formes de reparació o compensació): Complert o en procés. Tot i no estar explicitat, com que l'eina no pren decisions, la responsabilitat de les actuacions recau en l'administració, que ja disposa de mecanismes de reclamació i responsabilitat patrimonial.

Capacitats de la IA:

4.4.1 i 4.4.2 (No exagerar capacitats): Complert. La memòria és honesta sobre les limitacions i la precisió real. L'enfocament del projecte és pràctic i basat en la millora de la gestió, no en una visió tecnocràtica.

Principi 5: Privacitat

La privacitat és tractada amb el màxim rigor ja que un aspecte crític per a un projecte que gestiona dades de serveis socials. El compliment del GDPR i la realització d'una avaluació en aquesta matèria (AIPD) són evidències clares d'aquesta prioritat, alineant-se amb el RIA. Més concretament, tenint en compte les preguntes clau del Model PIO, podem assenyalar el següent:

Protecció de la privacitat:

5.1.1, 5.1.3, 5.1.9 (Mecanismes de protecció, millora contínua i preservació de la privadesa): Complert. L'AIPD realitzada és el mecanisme principal. La decisió de desagregar les dades del Padró fins a nivell d'"illa" i no més enllà és una aplicació pràctica del principi de minimització i de protecció des del disseny.

5.1.4 (Mecanisme per sol·licitar l'eliminació de dades): Complert per llei. Com a responsable del tractament, l'Ajuntament està obligat a respectar els drets ARSOL (Accés, Rectificació, Supressió, Oposició i Limitació) dels ciutadans segons el RGPD. L'eina no altera aquesta obligació legal.

5.1.10 (Ús de tècniques d'anonimat i xifratge): Complert. La memòria esmenta explícitament que les dades del Padró són "anonimitzades" per a les funcionalitats de descàrrega. L'ús de bases de dades segures (PostgreSQL, Oracle) i la capa d'API impliquen la implementació de mesures de seguretat estàndard.

Governança de dades:

5.2.1 (Registre d'accés a dades): Implícitament complert. L'ús de bases de dades institucionals (PostgreSQL, Oracle) i de l'APEX per a la interfície suggereix l'existència de registres d'accés i logs, pràctica estàndard en sistemes d'aquest tipus.

5.2.4, 5.2.5 i 5.2.8 (Documentació, adequació de dades i limitació de reutilització): Complert en alt grau. El procés de documentació és robust. L'anàlisi de la representativitat de les dades respecte al context (barris de Girona) és una part central de l'avaluació de la qualitat. Les "regles de coneixement" ajuden a garantir l'adequació i a tractar dades sensibles.

Principi 6: Autonomia

Aquest principi és un dels pilars fonamentals del projecte. L'objectiu declarat és alliberar al personal de les tasques mecàniques perquè puguin dedicar-se a treballar els vincles, i un ús de la IA que augmenta l'autonomia i el valor humà de la feina social, en línia amb el RIA. Més concretament, tenint en compte les preguntes clau del Model PIO, podem assenyalar el següent: Impacte en les capacitats de les persones usuàries:

6.1.1 i 6.1.2 (No disminuir opcions ni restringir llibertats): Complert. L'eina està dissenyada per augmentar el coneixement, no per limitar opcions. La informació agregada que proporciona permet als tècnics entendre millor el context, ampliant el seu camp d'acció i possibilitats de planificació.

6.1.3 (Protocol per ajudar a prendre millors decisions): Complert. Aquest és el seu objectiu principal. Proporcionar informació i prediccions basades en dades històriques i actualitzades permet prendre decisions més informades.

6.1.4 (Consciència que es tracta d'una IA): Complert. La formació i el manual d'usuari asseguren que els tècnics són plenament conscients que estan utilitzant una eina d'IA com a suport.

6.1.5 (Capacitació i informació per a usuaris finals): Complert o en procés. La formació està planificada per als tècnics (usuaris finals de l'eina). L'accessibilitat de la interfície es va testejar amb professionals.

Percepció del sistema:

6.2.1, 6.2.2 i 6.2.4 (No sobreestimar capacitats i evitar opacitat): Complert en alt grau. La memòria és transparent sobre les capacitats reals. El disseny de l'eina com a suport, no com a oracle, evita la sobreestimació. L'ús de models estadístics i d'aprenentatge automàtic, juntament amb la validació humana, ajuda a evitar que les operacions siguin opaques.

Principi 7: Sostenibilitat

El projecte aborda la sostenibilitat principalment des de dues perspectives: la sostenibilitat institucional (desenvolupat amb recursos propis, sense dependències privades) i la sostenibilitat social (millora de la planificació per a una gestió més eficient i justa dels recursos). L'impacte ambiental, tot i ser rellevant, no es detalla a la memòria. Més concretament, tenint en compte les preguntes clau del Model PIO, podem assenyalar el següent:

Desenvolupament sostenible:

7.1.1 (Impacte en drets laborals i qualitat del treball): Complert. És un objectiu central: “alliberar al personal de Serveis Socials de les tasques mecàniques”. L'automatització de tasques administratives allibera temps per a tasques de més valor, millorant la qualitat del treball.

7.1.3 (Accés social i bretxa digital): Complert o en procés. L'eina, en millorar la planificació dels recursos, pot contribuir a reduir desigualtats en l'accés als serveis. No obstant això, la memòria no explora com l'eina aborda la bretxa digital entre els mateixos tècnics o els ciutadans atesos.

7.1.4 (Participació amb parts interessades): Complert en alt grau. Aquest és el model de treball del projecte (recerca-acció participativa). La implicació de tècnics, direccions, acadèmia i altres àrees municipals és el seu punt més fort.

7.1.7 (Col·laboració amb sindicats): No esmentat. La memòria no detalla si hi ha hagut una col·laboració activa amb la representació laboral del personal de serveis socials, un aspecte clau per a una transició justa.

Impacte mediambiental:

7.2.1 a 7.2.9 (Enregistrament i avaluació de l'impacte ambiental): No avaluat. La memòria no conté informació sobre el consum energètic dels servidors, la petjada de carboni associada als models d'aprenentatge automàtic o l'ús de fonts d'energia renovables. Aquest és un punt de millora important, tot i que és comprensible que en un projecte pilot social, aquest aspecte no fos prioritari. També cal destacar que els sistemes d'IA que s'utilitzen no són grans models de llenguatge (LLM en anglès) sinó aplicacions d'IA que són molt més eficients en comparació des d'un punt de vista energètic.

Conclusions

A continuació us oferim unes breus conclusions de l'avaluació del Projecte Pigall a través de l'aplicació del Model PIO de l'OEIAC. Aquestes es centren en un seguit de punts forts destacables, així com en les àrees de millora i reptes a tenir en compte. En aquest punt, cal subratllar que el Projecte Pigall no només compleix amb els estàndards exigits pel Model PIO (que recull estàndards internacionals i requisits legals), sinó que en molts aspectes també els supera i, per tant, demostra que és possible innovar en l'àmbit social amb aplicacions d'IA que han estat dissenyades de forma ètica, legal i centrada en les persones.

Punts forts destacables:

- Governança participativa: La col·laboració social-tecnològica-ètica és el seu actiu principal.
- Transparència i explicabilitat: Documentació exhaustiva i comunicació honesta de limitacions.
- Autonomia humana: Dissenyat com a eina de suport, no de substitució, potenciant la feina social.
- Privacitat i protecció de dades: Tractament rigorós de les dades personals i utilització de salvaguardes, incloent-hi la realització d'una AIPD.
- Sostenibilitat institucional: Desenvolupament amb recursos propis, sense dependències privades.

Àrees de millora i reptes:

- Accés a dades externes: La dependència d'APIs d'altres administracions limita l'ampliació de l'eina.
- Col·laboració amb la representació laboral: Incloure activament els sindicats en el procés per gestionar l'impacte en les condicions de treball.

- Formació continuada: La formació inicial és un bon punt de partida, però cal un pla de formació contínua per a tots els usuaris potencials per assegurar una adopció òptima.
- Monitorització post-desplegament: Cal planificar com es farà el seguiment de l'ús, la precisió dels models en el temps (per evitar el que es coneix com a *drift*) i la satisfacció dels usuaris després de la implementació completa.
- Avaluació d'impacte ambiental: Incorporar l'anàlisi de la petjada de carboni del projecte quan estigui en ple funcionament.