

Emancipació juvenil a Espanya: una anàlisi dels determinants en la transició residencial.

Nom de l'estudiant: Jordi Pérez i Esclarin

Nom de la tutora: Catalina Llana Hesse

9 de juny del 2024

MEMÒRIA DEL TREBALL DE FINAL DE GRAU

Curs: 5è.

Estudis: Doble Grau en Administració d'Empreses i Gestió de la Innovació / Turisme i Gestió de l'Oci.

Resum

El present estudi aborda l'emancipació juvenil a Espanya en el període postpandèmic, que es troba en un moment crític. A través de l'aplicació metodològica *Random Forest*, s'aprofundeix en el perfil sociodemogràfic dels joves espanyols emancipats, centrant-se en les característiques que permeten la seva transició residencial. Els resultats mostren que la situació laboral, l'entorn familiar (l'estructura de la llar), així com les condicions residencials i econòmiques, desenvolupen un paper fonamental en l'abandonament de la llar dels progenitors. En contraposició, es constata que elements com la situació educacional o característiques idiosincràtiques, exceptuant l'edat de l'individu, no exerceixen un efecte significatiu en el procés d'emancipació.

Resumen

El presente estudio aborda la emancipación juvenil en España en el periodo pospandémico, que se encuentra en un momento crítico. A través de la aplicación metodológica *Random Forest*, se profundiza en el perfil sociodemográfico de los jóvenes españoles emancipados, centrándose en las características que permiten su transición residencial. Los resultados muestran que la situación laboral, el entorno familiar (la estructura del hogar), así como las condiciones residenciales y económicas, desempeñan un papel fundamental en el abandono del hogar de los progenitores. En contraposición, se constata que elementos como la situación educacional o características idiosincráticas, exceptuando la edad del individuo, no ejercen un efecto significativo en el proceso de emancipación.

Abstract

The present study addresses youth emancipation in Spain in the post-pandemic period, which is at a critical juncture. Through the methodological application of *Random Forest*, it delves into the sociodemographic profile of emancipated spanish youth, focusing on the characteristics that facilitate their residential transition. The results demonstrate that employment status, family environment (household structure), as well as residential and economic conditions, play a fundamental role in leaving the parental home. In contrast, it is observed that factors such as educational status or idiosyncratic characteristics, excluding age, do not have a significant effect on the emancipation process.

Paraules clau: joves, emancipació, determinants, Espanya, habitatge.

Taula de continguts

1. Delimitació i rellevància de la investigació	4
1.1. Justificació científica	4
1.2. Motivació personal per a la investigació i importància de la recerca en el grau	6
2. Antecedents	7
2.1. Revisió dels determinants de la transició residencial	7
2.2. Contextualització de l'emancipació juvenil a Espanya: l'ovella negra d'Europa.	10
2.3. Conclusions del marc teòric	14
3. Objectius i hipòtesis	15
3.1. Objectius	15
3.2. Hipòtesis	16
4. Metodologia emprada	16
4.1. Recopilació i processament de les dades	16
4.2. Confecció del model predictiu	19
5. Resultats obtinguts	22
5.1. Model predictiu: aplicació del Random Forest.	22
5.2. Model predictiu: mètode SHAP.	27
6. Conclusions de la recerca	35
7. Referències	38
8. Apèndixs	46
8.1. Apèndix I	46
8.2. Apèndix II	48
8.3. Apèndix III	60
8.4. Apèndix IV	61
8.5. Apèndix V	70

1. Delimitació i rellevància de la investigació

En aquest capítol, s'exposa una panoràmica inicial del tema, així com les motivacions personals i la importància de l'estudi per l'àmbit acadèmic.

1.1. Justificació científica

No és casualitat que un nombre considerable d'investigadors hagi centrat la seva atenció en l'emancipació juvenil a Espanya (Merino & Garcia, 2006; Maeso & Méndez, 2008; Aparicio & Oppedisano, 2015; Colom & Molés, 2016; Moreno, 2018; Donat & Martín, 2020; Soler et al., 2023). Un interès que s'ha vist reforçat per les creixents dificultats que els joves espanyols afronten en la transició a l'adulthood, resultant, entre altres qüestions, en l'alteració del calendari emancipatori (Torres, 2023).

En aquest sentit, diverses són les causes, específiques en el context espanyol, que han contribuït a la situació. Destaca la creixent precarització del mercat laboral (Úbeda et al., 2020) i els múltiples obstacles per accedir-hi (Julià et al., 2012), que ha afectat particularment els joves i la seva llibertat econòmica. La limitada oferta d'habitatge en el territori tampoc ajuda, una condició que s'agreuja en zones d'alta concentració poblacional, on els preus són gairebé inaccessibles per a la majoria, però en especial per als joves a causa de la manca de solvència (Echaves, 2017). Altres acadèmics assenyalen la ineficàcia – o, més aviat, l'absència – de polítiques que facilitin l'emancipació d'aquest grup (Soler et al., 2014). A propòsit d'això, cal considerar el període d'inestabilitat que ha definit el segle XXI, inicialment per la recessió de 2007 (Carballo-Cruz, 2011), i posteriorment per la pandèmia de la Covid-19 (Royo, 2020; González-Sanguino, 2020), les quals han exacerbat la tendència d'emancipació tardana en les noves generacions.

No obstant, atès que Espanya s'emmarca dins del model mediterrani d'emancipació, algunes veus atribueixen la demora en la transició residencial als factors culturals. És el cas de Comas (2015), que introdueix el concepte de la "síndrome de la dependència familiar" per descriure els patrons dels joves autòctons en l'abandonament de la llar familiar. L'autor exposa que, aquest estat de dependència, és producte d'un acord implícit que està profundament arrelat en la societat espanyola, del qual deriven actituds i pràctiques que reforcen la permanència dels joves en el si familiar. I aquest fet, de forma inexorable, ha

acabat prolongant el moment en què els joves assolien la plena maduresa i s'emancipen. Com tot relat, compta amb els seus detractors (Gentile, 2015; Noriega, 2023) com defensors (Requena; 2006a, Gil Calvo, 2014; Alberdi, 2014), tot i que el debat se centra més en la proporció en què els aspectes culturals, respecte al total d'elements involucrats, han induït a aquest estat d'estancament.

Com afirma Moreno (2012, p. 24), "el retard en deixar la llar dels pares en els països del sud d'Europa i específicament a Espanya...[és] el resultat d'una combinació de factors econòmics, institucionals i culturals". De manera anàloga, investigadors com Julià et al. (2012) qualifiquen aquest fenomen com a estructural, amb la conseqüència directa de què les parelles posposen la decisió d'engendrar fills.

Més enllà de les repercussions que això implica, resulta innegable que l'emancipació dels joves espanyols es troba en una etapa crítica (de Miguel, 2024), quelcom alarmant si es contrasta a escala continental. Al cap i a la fi, la mitjana d'edat d'emancipació a Espanya va ser de 30,03 anys en el 2022, màxims històrics d'ençà que hi ha registres, mentre que l'índex europeu corresponent a aquest mateix període se situava entorn els 26,4 anys (Eurostat, 2023). En contraposició, els països nòrdics s'han posicionat a la capdavantera dels índexs, amb una mitjana d'emancipació als vint-i-un anys (ibid). Aquesta evidència, que no és més que la punta de l'iceberg, adquireix un nou matís quan es considera l'evolució percentual de l'indicador. Designant el 2012 com l'any base o de referència, s'extreu un augment del 0,689% interanual en l'àmbit espanyol i, en canvi, una variació del només -0,0367% en el panorama europeu, deixant patent la projecció futura (ibid).

El retard en l'emancipació dels joves espanyols no és fruit d'un factor aïllat, sinó més aviat d'una complexa xarxa de circumstàncies. Esdevé evident que l'autonomia d'un individu variarà en funció el seu perfil sociodemogràfic (Donat & Martín, 2020), el moment en el temps, la regió i la comunitat autònoma (CJE, 2023), els factors estructurals (Molgat, 2007), així com per característiques intrínseques de cada persona (Spatscheck, 2007; Yates & Grey, 2012), entre altres motius. Certament, la combinació d'aquests elements contribueix a una realitat complexa i multifactorial, la qual es troba en un punt d'inflexió. Només cal fixar-se en les taxes d'emancipació de l'estat espanyol, que segons les darreres dades eren només del 16,3% (CJE, 2023), per comprendre el repte que planteja el calendari emancipatori per al país.

1.2. Motivació personal per a la investigació i importància de la recerca en el grau

En relació amb l'equip investigador, aquest està integrat per un únic membre, Jordi Pérez i Esclarin, qui en l'actualitat es troba cursant un doble grau en Administració d'Empreses i Gestió de la Innovació / Turisme i Gestió de l'oci. La transversalitat d'ambdós vessants ha dotat l'estudiant d'una perspectiva àmplia i multidisciplinària, assimilant no només els conceptes requerits per l'anàlisi estadística, sinó també atorgant-li enteniment dels moviments econòmics i de mercat que, tan directa com indirectament, pauten l'emancipació juvenil. Més concretament, la formació en Fonaments d'Estadística i Inferència Estadística, així com en Anàlisi de l'Entorn Macroeconòmic i Investigació de Mercats, ha conferit a l'investigador les competències necessàries per dur a terme el present estudi.

De tota manera, aquest plantejament conceptual ressalta la rellevància de la investigació en el grau, puix que, a més de ser un tema d'actualitat, pot aportar significat a l'àmbit acadèmic i obrir noves línies de recerca. A tal efecte, es realitzarà un diagnòstic dels joves emancipats, en procés d'emancipació o, per defecte, en edat per emancipar-se. És a dir, aplicant el mètode científic, es pretén extreure les suficients evidències per assolir una major comprensió de la casuística, discernint tant el perfil dels que s'emancipen com les respectives barreres que sorgeixen en el procés.

La voluntat d'indagar en els determinants de l'emancipació juvenil a Espanya, essent un tema prou específic, neix de la inquietud de l'estudiant arran l'aparent incompatibilitat entre l'habitatge i aquest segment poblacional. Per això, procurant una concepció holística, es pretén estimar el grau d'impacte d'aquells aspectes sociodemogràfics associats a l'emancipació residencial espanyola, amb l'objecte d'establir un punt de partida per al desenvolupament d'estratègies que abordin aquesta qüestió.

2. Antecedents

Segons Soler et al. (2023, p. 1), basant-se en les investigacions de Jones & Wallace (1992):

“La transició de la joventut a l'edat adulta és un procés social marcat per tres dimensions interrelacionades: la dimensió familiar, entesa com la formació d'una família pròpia; la dimensió econòmica i laboral, que es produeix al final de l'etapa educativa i, per tant, quan és possible l'accés al mercat laboral; i la dimensió residencial, és a dir, abandonar el domicili parental per desenvolupar una vida independent en un habitatge diferent”.

Partint d'aquesta idea, en aquest apartat es revisen les evidències científiques respecte a l'emancipació juvenil, en concret els factors que determinen la transició residencial dels joves. Consecutivament, s'aprofundeix en el context de l'emancipació juvenil a Espanya amb una perspectiva europea.

2.1. Revisió dels determinants de la transició residencial

En relació amb les característiques sociodemogràfiques que poden influir el calendari emancipatori, tal com exposen Billari & Liebroer (2007), l'edat emergeix com un element quelcom rellevant. Tot i les variacions regionals en l'edat mitjana d'emancipació, aquesta es concentra en uns anys específics, amb edats d'emancipació més primerenques al nord d'Europa que al sud i a l'est.

Stone et al. (2011) assegura que hi ha una relació entre la nacionalitat i l'emancipació entre els joves adults. Bàsicament, aquells amb nacionalitat estrangera tenen tendència a emancipar-se en un moment més tardà de les seves vides en comparació amb els joves autòctons. Això es deu a diversos factors, inclosos els desafiaments associats a l'adaptació a un nou entorn i les diferències en les oportunitats laborals i educatives.

Altres acadèmics, com Molgat (2007), Colom & Molés (2016) i Moreno (2018), assenyalen que el gènere exerceix un rol important en l'abandonament del nucli familiar. Els tres estudis arriben a conclusions similars, revelant diferències considerables en el comportament d'homes i dones. Per exemple, entre els homes, la situació laboral i la independència econòmica són factors que influeixen significativament en la seva emancipació. En canvi, per a les dones, aquests mateixos factors no juguen un paper tan destacat.

Miret (2006) sosté que el moment en què s'ingressa en el mercat laboral afecta de manera diferencial entre els homes i les dones; mentre que per aquests primers iniciar aviat la seva carrera professional els impulsa a deixar abans el domicili patern, per a les dones fa l'efecte invers.

A part, tant Moreno (2018) com Molgat (2007) apunten al fet que un dels principals determinants en l'emancipació de les dones és la convivència amb parella, independentment si és en cohabitació o matrimoni. En la mateixa línia, South & Lei (2015) exposen que tenir una relació de parella estable encoratja l'emancipació, de la mateixa manera que formar una nova família també n'augmenta la probabilitat.

Elder & Shanahan (2007) o Moreno (2012) mantenen que les dones, però, tenen una probabilitat més alta de deixar el domicili parental que els homes. Per regla general, el sexe femení és més procliu a formar una nova família, ergo, a abandonar la casa dels progenitors més d'hora (ibid).

Pot assumir-se que un dels elements més explicatius de la transició residencial és el nivell d'estudis assolit. No obstant, la literatura existent presenta resultats contradictoris. Per exemple, Miret (2006) i Moreno (2018) varen determinar que aquesta variable, almenys en el context espanyol, no s'associava amb una emancipació anticipada o diferida en el temps. És a dir, no tenia cap efecte. En contraposició, Colom & Molés (2016) varen constatar que, en els anys 1990, 2000 i 2010 a Espanya, allargar l'etapa de formació, o bé assolir un major nivell d'estudis, postergava l'abandonament del domicili familiar. Podria argumentar-se que la continuació dels estudis suposa una demora en l'ingrés al món laboral, moment en el qual, en principi, l'individu pot assolir la independència econòmica i la capacitat d'automantenir-se.

Altres *papers* també corroboren aquesta observació, indicant que allargar l'etapa estudiantil repercuteix negativament en l'emancipació dels joves (Stone et al., 2011; Donat & Martín, 2020), i fins i tot en la paternitat, posposant-la en el temps (Kneale & Joshi, 2008).

Ara bé, Stone et al. (2011) varen descobrir algunes peculiaritats, aplicables al Regne Unit, amb relació al nivell d'estudis assolit i la situació laboral. En aquest sentit, els joves amb estudis superiors s'independitzaven en major proporció a un habitatge compartit, mentre que els desocupats i econòmicament inactius tenien més probabilitats de viure sols. La

co-residència amb els pares, a més, era comuna tant entre dones amb estudis universitaris com sense. Tot i que als 30 anys era més habitual entre aquells amb estudis mínims, desocupats i econòmicament inactius, sense importar el gènere. Per contra, al voltant de la trentena, l'educació superior fomentava més l'emancipació que a principis i mitjans de la vintena d'edat. Per tot això, es podria afirmar que el nivell d'estudis i la situació laboral afecten de manera diferent a les persones en l'etapa dels 20 als 30 anys.

L'estat ocupacional (Bellart & Oller, 2005; Stone et al., 2011; Moreno, 2012; Colom & Molés, 2016; Donat & Martín, 2020), així com l'estabilitat de la feina per se (Becker et al., 2010), és un dels predictors del calendari emancipatori, puix que estar actiu professionalment incrementa les oportunitats d'independència, mentre que estar desocupat o tenir unes males condicions laborals, fa l'efecte contrari. Hi ha excepcions, però; altres estudis, com el de Cordón (1997), demostren que al sud-oest d'Europa, al voltant dels anys noranta, estar ocupat laboralment no constituïa una raó de pes per a l'emancipació.

Relatiu a l'efecte de la renda, Ghidoni (2002) planteja que el no disposar d'una font d'ingressos estable al llarg del temps influeix de manera adversa en les possibilitats dels joves per emancipar-se. Tot i això, com indica Moreno (2018), dependre econòmicament dels pares mostra una correlació negativa.

Manacorda & Moretti (2006) i Ghidoni (2002) coincideixen que, a mesura que augmenta l'estrat d'ingressos dels progenitors, les possibilitats que els fills surtin de la llar familiar disminueixen. Requena (2006b) i Stone et al. (2011) demostren que la transició a la vida adulta està estretament lligada a la classe social que pertany l'individu. Contràriament, estudis com els de Kuhar & Reiter (2014) suggereixen que la categoria social, la situació econòmica i la formació dels pares no sempre són factors explicatius de l'emancipació dels seus descendents. Per tant, la correlació entre la posició socioeconòmica de la família respecte al calendari emancipatori es posa, si més no, en qüestió.

Mentrestant, els resultats obtinguts de l'estudi de Blaauboer & Mulder (2010) permeten afirmar que un entorn acollidor en el domicili d'origen inhibeix la determinació dels joves de canviar de pis. En canvi, créixer en habitatges monoparentals comporta una sortida més prematura del nucli familiar.

D'acord amb Spatscheck (2007) i Yates & Grey (2012), les característiques personals, psicològiques i ambientals poden condicionar la transició a la vida adulta. Considerant la seva naturalesa qualitativa, no és d'estranyar que en la literatura no hi hagi prou resultats empírics que permetin determinar-ne el grau d'impacte. Tal com es desprèn de l'estudi conduït per Rosas et al. (2006), la quantificació i estandardització d'aquesta mena de variables és un procés intricat, fet que podria dificultar l'anàlisi de la seva contribució relativa en l'emancipació.

Els factors estructurals, com ara el mercat laboral i el mercat de l'habitatge, també configuren el trajecte de vida dels joves. Per citar un cas, la disponibilitat d'oportunitats laborals i d'habitatge a un preu assequible modelen la capacitat dels joves adults per adquirir autonomia (Molgat, 2007).

Igualment, Aparicio & Oppedisano (2015) posen de manifest que la presència de subsidis en el lloguer està correlacionada positivament amb l'emancipació juvenil. Els científics conclouen que, en la maquetació d'aquestes polítiques, és bàsic considerar tant els salaris dels joves com els preus dels habitatges de la regió, puix que, si no es fa, les mesures podrien no ser del tot efectives, especialment per als joves que es troben en el llindar de pobresa.

En darrer lloc, les dinàmiques econòmiques també són identificades com a factors decisius en l'emancipació juvenil, ja que els períodes de recessió o crisi deriven en una disminució dels índexs d'emancipació (Stone et al., 2011; Ballesteros et al., 2012). Així doncs, unes condicions contextuais adverses minven les opcions d'emancipar-se.

2.2. Contextualització de l'emancipació juvenil a Espanya: l'ovella negra d'Europa.

El procés d'emancipació a Europa s'ha sotmès a un endarreriment progressiu en les últimes dècades, convertint-se en un fenomen més complex, que alhora requereix més temps per completar-se un cop s'inicia (Billari & Liefbroer, 2010). És indubtable, a més, que els models del món laboral i familiar s'han desvinculat dels rols i les expectatives tradicionals de la societat, alterant el que abans seguia un curs seqüencial de les diferents etapes de la vida (Beck et al., 1994).

Com s'ha constatat en l'apartat 1.1., a Espanya aquesta tendència s'ha intensificat en relació amb els estats veïns, convertint-se en un dels països europeus amb una de les mitjanes d'edat d'emancipació més elevades.

Podria ésser objecte de consideració que certs canvis culturals, tecnològics i econòmics que han acompanyat l'era de la globalització (Colom & Molés, 2016), juntament amb altres singularitats del context espanyol, han catalitzat aquesta situació. Amb tot, la presa de decisions dels joves espanyols vers l'emancipació s'ha vist compromesa.

Per contextualitzar, l'entorn laboral a Espanya, amb una forta inclinació al sector serveis, es defineix per la presència generalitzada de contractes temporals, retribucions baixes i limitades perspectives de desenvolupament professional (CJE, 2023). Una situació que es fa especialment prevalent entre els joves, amb la comuna situació que es troben sobrequalificats per als llocs de treball que ocupen (ibid). A part, per aquest segment, l'accés a una feina ja és, per si mateix, un desafiament. Aprofundint en les dades, els índexs d'atur entre els joves va ser aproximadament del 28% l'octubre de 2023, gairebé duplicant la mitjana continental, que era del 14,4% en el mateix període (OECD, 2024).

Certament, aquesta conjuntura no fa sinó accentuar la disparitat entre els salaris dels joves espanyols i el cost de vida, una divisió que es fa present en moltes àrees, però sobretot en el lloguer o l'entrada d'un pis. En aquest sentit, un jove hauria de destinar el 93,9% del seu salari per poder llogar un habitatge de manera individual, mentre que el cost de l'entrada per adquirir un pis equivaldria a 4,5 anys del seu salari íntegre (CJE, 2023). Com a resultat, els joves espanyols es troben amb escasses perspectives per aspirar a la independència residencial.

De fet, a Espanya, els joves poden ser considerats com un grup vulnerable, ja que es troben en major risc de pobresa que altres cohorts d'edat. Una conjuntura que revela una bretxa generacional, en la que les condicions econòmiques i les oportunitats dels joves són significativament menys favorables que generacions anteriors (Cantó, 2019). Com a dimensió addicional, també s'hi suma l'efecte de la inflació. Malgrat que s'ha moderat des del 2022, quan va experimentar un augment a causa de l'escalada dels conflictes geopolítics a Europa, en última instància hi ha hagut una pèrdua de poder adquisitiu de la població (CJE, 2023).

El mercat immobiliari espanyol és, sens dubte, un dels altres causants del retard en el calendari emancipatori. Echaves (2017) exposa que, a diferència d'alguns països del centre i nord d'Europa, on els desajustos entre l'oferta i la demanda es regulen mitjançant la intervenció pública, a les societats del sud, com Espanya, aquesta funció recau en les forces del mercat, amb un rol més secundari de l'Estat. Un *modus operandi* que ha acabat provocant que l'habitatge espanyol ja no sigui un dret universal, sinó més aviat un bé sotmès a especulació. Inevitablement, això ha propiciat l'auge de grans corporacions i fons d'inversió que controlen una important quantitat d'habitatges en determinades zones urbanes, encarint-les i gentrificant-les en el procés. De nou, els joves són dels que més es veuen afectats per aquest sistema, ja que representen una demanda poc solvent i que, per extensió, no poden fer front a les regles de joc del mercat.

Alguns investigadors, com ara Delgado (2018), especulen que la solució a aquesta problemàtica podria trobar-se en l'Espanya rural, col·loquialment anomenada com 'l'Espanya Buida', on la disponibilitat d'habitatges més assequibles podria proporcionar als joves l'oportunitat d'establir-se. En l'actualitat, no obstant, no hi ha prou incentius, com la manca d'infraestructures i llocs de treball, per atraure'ls a viure i establir-s'hi.

En còmput global, la resposta dels ens públics davant del fenomen tampoc ha sigut la més encertada. És veritat que hi ha hagut intents com la Renda Bàsica d'Emancipació (d'ara endavant RBE) de 2007, la qual tenia com a finalitat impulsar l'accés a l'habitatge als joves (BOE, 2007). Malauradament, com argumenta Gentile (2015), la lentitud burocràtica en la concessió de les ajudes, l'augment dels lloguers per part dels arrendadors i les retallades dutes a terme per reduir el deute públic arran el xoc del 2007, incloent-hi la de la RBE, van tenir una conseqüència fatal en la seva aplicació. A raó de l'autor, la introducció d'aquesta política, que va ser una temptativa de l'estat d'imitar estratègies efectives d'altres països de la Unió Europea, no va aconseguir els resultats esperats, ja que el nombre de joves adults emancipats va continuar disminuint gradualment.

No va ser fins a l'any 2022 que es va adoptar la que seria la successora de la RBE, coneguda com el Bo de Lloguer Jove, que també té per objectiu secundar l'emancipació de les persones entre 18 i 35 anys amb pocs recursos econòmics (BOE, 2022). Encara és aviat per conèixer els seus efectes, a més que no s'ha realitzat cap estudi exhaustiu, de moment, per analitzar-ne l'impacte. Així i tot, a causa de la insuficiència en la dotació pressupostària, i

que les ajudes no són ajustades o adaptades a les diferents àrees geogràfiques, Sala i Roca (2022, p. 24) declara que:

“...ni ens trobem davant d'una mesura estructural de la política d'habitatge,... ni som davant d'un mecanisme adaptat a les realitats socials i territorials,... ni som davant d'una eina que pugui afavorir de manera justa, transformadora i eficaç les necessitats d'emancipació de la població jove”.

Al mateix temps, convé considerar els aspectes culturals que poden contribuir a la dependència dels joves espanyols. Segons Esping-Andersen (1999), en els països del sud d'Europa, com Espanya, hi predomina un model “familista” o de dependència familiar, ja que les famílies es responsabilitzen del suport econòmic i social dels seus fills durant períodes més extensos, endarrerint així l'entrada a la vida adulta. Aquesta mateixa idea és defensada per Gil Calvo (2002), qui manté que la forta cohesió de les famílies espanyoles és un dels principals causants de la poca autonomia dels seus joves.

Com a contraargument, Moreno (2012) defensa que la prolongació de l'etapa juvenil, tan característica a Espanya, és més atribuïble als factors estructurals que a la mateixa cultura. Fonamentalment, la inestabilitat del mercat laboral, les múltiples barreres per accedir a un habitatge i unes polítiques insuficients convergeixen en el retard dels joves a establir-se en un habitatge propi. És a dir, segons l'autora, els joves no opten tant per romandre a la llar familiar com es veuen obligats a fer-ho a causa de les nul·les alternatives que se'ls presenten.

En aquest context, és evident que la societat ha experimentat un canvi de paradigma, donant lloc a nous patrons de comportament, molts dels quals com a resposta a les adversitats que han sorgit en la transició a l'adulthood (Billari & Liefbroer, 2010). Per exemple, a Espanya els joves opten per romandre més temps a casa dels pares per poder d'acumular el màxim capital possible (Moreno, 2012). De la mateixa manera, segons els resultats obtinguts per Bernardi (2007), els joves espanyols no s'independentzen fins a un cop que han aconseguit, com a mínim, la mateixa posició benestant que els seus progenitors. Al marge d'això, també es pot atribuir una prolongació del període d'escolarització en les noves generacions, que ha anat en augment des dels inicis del segle (INE, s.d.b).

Per últim, cal destacar que, en l'àmbit europeu, també han aparegut noves tendències de convivència, com seria la independència individual, la cohabitació amb altres individus de diferent nucli familiar, o bé la convivència amb parella sense estar casats (Toulemon, 1997; Levin, 2004; Duncan & Philips, 2008). A contrapunt, els matrimonis espanyols han sofert un declivi, amb una disminució del seu nombre amb el pas del temps, i, com a resultat, s'ha donat una contracció en la formació de noves famílies i dels índexs de natalitat (Tobío, 2001).

2.3. Conclusions del marc teòric

En definitiva, hi ha múltiples factors, tant a nivell macro com a nivell micro, que intervenen en la decisió o capacitat d'emancipar-se. Dins d'aquests, diverses investigacions apunten a una àmplia gamma de variables, com són l'edat, la nacionalitat, la cultura, el gènere, la presència d'una parella estable, el nivell d'estudis assolit, l'estat civil, tenir fills i l'estat ocupacional. Però també altres com la classe social i la formació dels progenitors, la renda, l'estructura de la llar, el mercat de l'habitatge, les polítiques públiques i el context socioeconòmic que afecten el calendari emancipatori.

Com s'ha constatat en la revisió de la literatura, dins de la Unió Europea, Espanya és un cas paradigmàtic quant a l'emancipació juvenil, trobant-se a la cua continental en els índexs d'emancipació. A més, malgrat estudis a escala estatal com els de Moreno (2018) o Colom & Molés (2016), per citar-ne alguns, es fa patent la poca literatura en què s'aprofundeix en els condicionants de l'emancipació residencial, sobretot en el moment present en el temps. La falta d'investigació del fenomen no només restringeix la seva comprensió global, el qual es troba en constant evolució, sinó que també obstaculitza l'elaboració de polítiques que abordin les necessitats reals d'aquest col·lectiu.

Per consegüent, esdevé necessari prosseguir l'estudi d'aquesta àrea, posant el focus en els determinants de la transició residencial. Com a qüestió de fet, la perseverança en aquesta labor és fonamental per poder identificar les característiques sociodemogràfiques tant dels joves que estan independitzats com dels que no. Només així es podrà focalitzar, i ulteriorment esmenar, l'*statu quo* de l'emancipació juvenil a Espanya.

3. Objectius i hipòtesis

Conceptualitzada la casuística i basant-se en les diferents línies obertes de la literatura existent, aquest estudi té com a objectiu explorar, en la nova realitat postpandèmica, la capacitat d'emancipació residencial dels joves espanyols, aprofundint en el perfil sociodemogràfic dels que estan emancipats. D'aquesta manera, la pregunta de recerca esdevé la següent:

En el context postpandèmic espanyol, quins són els factors que influeixen en major mesura el fet que un jove o una jove s'emancipi residencialment?

3.1. Objectius

En consonància amb la pregunta de recerca, l'objecte de la investigació és el desenvolupament d'un Treball de Final de Grau (TFG) científic en el qual s'identifiquin els determinants de la transició residencial dels joves espanyols d'entre 18 a 29 anys. En altres paraules, es pretén quantificar el grau d'impacte d'aquelles característiques sociodemogràfiques que estan associades a l'emancipació. Per això, els objectius de l'estudi són els següents:

- **Objectiu 1:** Identificar els determinants sociodemogràfics en la transició residencial dels joves en el context postpandèmic espanyol.
- **Objectiu 2:** Determinar el grau de correlació de cadascuna de variables, en especial èmfasi en la condició laboral, el gènere, el nivell d'estudis i la situació residencial, amb el procés d'emancipació dels joves espanyols d'entre 18 a 29 anys.

S'ha d'aclarir que, tot i haver seleccionat l'època postpandèmica com a període d'anàlisi, l'estudi no pretén centrar-se explícitament en els efectes de la pandèmia en si mateixa. D'altra banda, per a la maquetació del segon objectiu, s'han considerat aquelles variables que han rebut més atenció per part dels investigadors i que estan més relacionades amb el calendari emancipatori o que, per defecte, poden ser rellevants pel context espanyol.

3.2. Hipòtesis

Tot seguit, es presenten les hipòtesis de la recerca, les quals estan intrínsecament vinculades als objectius establerts i seran subjectes a contrast o refutació a partir dels resultats obtinguts.

H1. Estar actiu laboralment afavoreix la probabilitat d'emancipació.

H2. Haver assolit estudis superiors disminueix la probabilitat d'emancipació.

H3. La convivència amb altres persones incrementa la probabilitat d'emancipació.

H4. El fet de ser dona augmenta la probabilitat d'emancipació.

4. Metodologia emprada

4.1. Recopilació i processament de les dades

Atès a les limitacions de l'estudi¹, s'ha optat per recórrer a una base de dades d'una font secundària. Malauradament, a Espanya, les enquestes que permeten prosseguir amb el treball de camp són escasses. Per començar, és essencial que les enquestes procedeixin d'una font oficial per garantir la fiabilitat de les dades. A part, és igualment important que aquestes incorporin una àmplia gamma de característiques sociodemogràfiques dels joves, i que al mateix temps siguin prou desagregades per ser compatibles amb els objectius definits.

No obstant això, hi ha una enquesta que compleix amb tots aquests paràmetres: l'Enquesta de Característiques Essencials de la Població i Habitatges (d'ara endavant ECEPOV). Realitzada per l'Institut Nacional d'Estadística entre el 2021 i el 2022, va recopilar informació sobre diversos habitatges, les persones que hi residien, així com les característiques dels edificis en què es trobaven aquests habitatges (INE, 2023).

¹ Les restriccions d'aquest estudi deriven de la manca de recursos i temps per a la realització d'una enquesta de font primària.

L'enquesta en si va cobrir tot el territori nacional i constava de més de cent preguntes que incloïen un ampli espectre de temàtiques, com ara la situació laboral, el nivell d'ingressos i el d'estudis, la nacionalitat, el gènere, així com aspectes relacionats amb la salut i la situació residencial, entre altres magnituds. Es va aplicar un mostreig aleatori estratificat per als municipis amb més de 50.000 habitants o capitals de província, i un mostreig bietàpic estratificat per a la resta de municipis. A través de qüestionaris electrònics i en paper, es van recollir dades de 172.444 habitatges, els quals albergaven un total de 424.493 persones.

Quant a la presentació de les dades, aquestes estan disposades en tres bancs; en un s'emmagatzema la informació relativa a cada habitatge, en un altre les respostes obtingudes de tots dels residents de cada llar i, en paral·lel, un tercer en el qual es troben respostes que concerneixen els residents de 16 o més anys d'aquests habitatges. De tota manera, els tres bancs de dades poden ser unificats en un de sol, utilitzant la variable *identificadora* com a nexa entre els diferents arxius.

Tot i això, per aconseguir una base de dades útil per a la investigació, és a dir, que contingui informació rellevant sobre el fenomen estudiat, és necessari realitzar uns passos previs. Mitjançant el programa Stata, s'ha filtrat per les següents variables: *edat*, per la franja de valors d'entre 18 i 29 anys, i per *motiu de canvi*, que és la raó de l'últim canvi de residència (vegeu apèndix I). En aquest cas, s'assigna la categoria de valor 4, que és la d'*emancipació*, però també la del valor 0, que són els joves que encara residien en la llar familiar al moment de ser enquestats.

D'aquesta manera, s'obté una mostra de 1.248 individus de 18 a 29 anys que estan emancipats, i 20.070 de la mateixa franja d'edat que no ho estan, que són els que actuaran com el grup de control. De mode que seran un total de 21.318 subjectes que conformaran la BBDD de la investigació.

Això sí, prèviament a la unificació dels tres bancs de dades esmentats, s'ha realitzat un cribratge de les variables que, segons la literatura existent, han demostrat certa capacitat predictiva en relació amb l'emancipació, o que, per defecte, podrien estar estretament relacionades amb el cas espanyol.

A continuació, s'exposen les 65 variables escollides per la BBDD (vegeu apèndix II):

Taula 1. Llistat de les variables recopilades i respectiu origen en el banc de dades.

Banc de dades	Variables extretes
"ECEPOVhogar_2021"	iden idq_pv idq_mun npv motivocambio sexo edad nacim provn munin pnacim pnac1 pnac2 nac nprogen reprogen1 hogp1 vivoprogen1 nacimp1 pnacimp1 pnac1p1 pnac2p1 estudiosp1 reprogen2 hogp2 vivoprogen2 nacimp2 pnacimp2 pnac1p2 pnac2p2 estudiosp2 nucleo tiponucleo sexo_pareja nacion_pareja union_pareja estudios_pareja sitlab_pareja situhogar hijos_nucleo tipohogar nachogar niveducahogar sitlabhogar
"ECEPOVadultos_2021"	iden npv ec estudios edadestud campo sitlab edadsl flexi lugtrab hijos nhijos pareja aposoci aposocihij aposocipama aposociher aposociotpar aposociotro lresmd
"ECEPOVvivienda_2021"	iden nresi regvi pagovi ingrehog

Font: Elaboració pròpia a partir de les variables extretes de l'ECEPOV (INE, 2023).

Malgrat tot, cal tenir en compte les següents consideracions:

- Amb el filtratge per la variable *motiu de canvi*, que indica la raó per la qual s'ha deixat la residència anterior, s'estarà exclouent de la BBDD aquells joves emancipats que han canviat dues o més vegades de llar i que, suposadament, tindrien altres raons per haver-se traslladat el moment que van contestar el qüestionari, com podria ser per motius laborals, per citar algun cas.
- Segons el full de metodologia de l'enquesta (INE, 2023), es va considerar l'estat emancipat a tot jove que hagués abandonat la llar familiar al marge de si tingués independència econòmica o no. És a dir, dins d'aquest grup de 1.248 joves emancipats, pot haver-hi que depenen econòmicament total o parcialment de les seves famílies i/o altres entitats per subsistir (dimensió no inclosa en la BBDD per la

falta d'una variable que ho determini). També pot donar-se que s'hagin traslladat a un habitatge que sigui de propietat dels pares o altres familiars (dimensió inclosa en la BBDD en el règim de *tenença de l'habitatge*), entre altres circumstàncies.

- Pel que fa a la representativitat poblacional de l'estudi, es pot afirmar que, amb una població total de 5.880.875 joves d'entre 18 a 29 anys l'any 2021 a Espanya (INE, s.d.a), 21.318 individus és una mostra representativa amb un 99% de nivell de confiança i un marge d'error del $\pm 0,88\%$. Cal puntualitzar que la majoria de les respostes de l'ECEPOV es van recollir el 2021, i és per això que s'ha escollit aquest període pel càlcul en comptes del 2022.

4.2. Confecció del model predictiu

Un cop la base de dades ha estat ajustada, s'ha procedit amb el desenvolupament d'un model predictiu que permeti detectar, per una banda, les variables que més determinen l'emancipació i, per l'altra, l'impacte relatiu de cadascuna d'aquestes. Històricament, la regressió ha sigut una de les metodologies més escollides per aquest camp d'estudi (Maeso & Méndez, 2008; Moreno, 2018; Donat & Martín, 2020), així com en àrees de recerca afins (Perugini & Signorelli, 2009; Baranowska & Gebel, 2010), abastant diverses modalitats, com són la regressió lineal, logística i polinòmica. Particularment en aquests casos, l'ús d'aquest mètode ha permès obtenir una major comprensió dels fenòmens estudiats, exposant el grau de relació de les variables independents respecte a la dependent.

Ara bé, com afirmen Aronow & Samii (2016), la regressió, tot i ser útil per comprendre quin pes té cada una de les variables independents respecte a la dependent, pot tenir dificultats a capturar la seva correlació quan les relacions són complexes o hi ha molta dispersió entre elles (com és el cas de la BBDD treballada). A tal efecte, una àmplia dispersió de les dades pot afectar negativament la precisió, la fiabilitat i la capacitat de generalització d'un model basat en la regressió lineal o similars, generant prediccions menys fiables en conseqüència.

Per aquests motius, s'ha optat per la metodologia *Random forest*, que és una eina "de machine-learning per a la classificació i regressió...[amb] un procediment inherent per a produir la importància de les variables" (Grömping, 2012, p. 2).

La raó d'aquesta aplicació metodològica recau en la capacitat del mètode de gestionar grans quantitats de dades, així com en la seva robustesa davant de la presència de variables correlacionades o la no-linealitat de les relacions entre les variables independents i dependent (ibid).

A més, la metodologia *Random Forest* permet avaluar l'eficàcia del model i evitar el sobreajustament (*overfitting*) mitjançant la tècnica de *train-test split*. En aquest procés, les dades es divideixen en dues parts: un conjunt de dades d'entrenament, que s'utilitza per construir el model, i un conjunt de dades de prova, que es reserva per avaluar el rendiment del model. La divisió permet validar el model amb dades que no ha vist abans, ajudant a confirmar així la seva capacitat de generalització i predicció (Salazar et al., 2022).

Com es desprèn de Hartshorn (2016), el procés del *Random Forest* consisteix en la construcció de múltiples arbres de decisió independents, cadascun d'ells agrupant una mostra aleatòria del conjunt de dades originals, incloent-hi tant les files (observacions) com les columnes (variables predictorres). En la construcció de cada arbre, el model avalua les característiques de les variables de la base de dades original i, amb l'objectiu de dividir el conjunt de dades en subgrups més homogenis, es fa una pregunta, que serà el valor d'una variable determinada. A partir d'aquí, el model va classificant en funció de si les observacions seleccionades compleixen la condició o no, creant, a efectes pràctics, un diagrama d'arbres de decisió.

A cada arbre consecutiu se li va assignant aleatòriament un nou subconjunt de les variables predictorres, augmentant així la diversitat del model. Per tant, cada arbre de decisió estarà basat en característiques úniques del conjunt de dades original. Alhora, cada ramificació representarà una bifurcació en el camí dels nodes o punts de decisió, separant les observacions del conjunt de dades en subgrups més petits en funció del valor de les variables (ibid).

Els hiperparàmetres són les regles que defineixen la construcció de l'arbre, com ara la profunditat màxima de l'arbre, el nombre mínim d'observacions per node i el nombre d'arbres en el bosc (ibid). En aquest cas, s'ha decidit construir un bosc de 100 arbres, on cada arbre ha de tenir almenys 10 observacions per node, tot dividint-se el conjunt de dades en 10 estrats. De mode que, quan un node tingui menys de 10 observacions (individus), el model no continuarà expandint-se en aquella direcció.

A mesura que l'arbre es vagi desenvolupant, les observacions s'aniran segregant en subgrups cada vegada més específics, agrupant-se aquells perfils d'observacions més similars, fins a arribar a nodes que representin les prediccions finals (vegeu apèndix III).

Així doncs, un cop establerts els darrers arbres, els resultats d'aquests es combinen mitjançant el valor mitjà. La mitjana és essencial perquè les prediccions finals associades a cada subgrup poden ser molt diferents, ja que reflecteixen les característiques úniques de les dades que cauen en aquella branca específica. D'altra banda, combinant les prediccions de tots els arbres del *Random Forest*, es podran identificar les variables independents que tenen més pes, i que actuaran al mateix temps com les característiques més explicatives de l'estat d'emancipació (ibid).

En aquest punt, es traspasa la base de dades prèviament processada a Stata a l'aplicació Jupyter Notebook, que és amb la que s'ha treballat per l'aplicació metodològica. Dit això, per garantir una representació equilibrada de les dues categories i evitar un biaix en l'anàlisi, el primer pas ha consistit a balancejar el nombre d'observacions entre persones emancipades i no emancipades. És a dir, s'ha reduït – de manera aleatòria – el nombre de persones no emancipades per igualar-lo al d'emancipades (vegeu apèndix IV).

Consecutivament, s'han transformat els espais buits (NaN) de la BBDD a valors que puguin ésser processats pel model. Per tant, quan en alguna de les variables categòriques no hi hagi valor algun, s'assignarà un valor zero (és a dir, valor inexistent) o, sinó, per tal de no alterar el model resultant, la mitjana de la columna corresponent. *Grosso modo*, per les variables categòriques s'ha tendit a assignar el valor zero, mentre que per les variables numèriques s'ha computat la mitjana (vegeu apèndix IV).

A partir d'aquí s'ha procedit amb el desenvolupament dels diversos models de predicció (*Random Forest*), tenint en compte la inclusió i exclusió de les variables independents en un procés manual, a fi d'avaluar la seva contribució a la capacitat predictiva del model. El procés s'ha repetit iterativament, ajustant el model fins a assolir una configuració en la qual es maximitzi la precisió. En aquest sentit, s'han pogut identificar les característiques sociodemogràfiques que més estan correlacionades amb l'estat d'emancipació (tant si sí com si no), així com quantificar la seva importància relativa (vegeu apèndix IV).

Trobat el model amb la millor precisió, s'ha utilitzat l'eina SHAP per obtenir una explicació detallada de com, per a cada individu en particular, influeix el seu perfil sociodemogràfic (conjunt de les variables explicatives) en la predicció de la seva emancipació. *De facto*, això permet aprofundir en ambdós estats de l'emancipació; emancipat (estat *true*) i no emancipat (estat *false*) (vegeu apèndix V).

SHAP (SHapley Additive exPlanations) permet calcular la contribució de cada variable a la predicció final mitjançant un procés en què s'exploren totes les combinacions possibles de les variables independents respecte a la dependent, a través de la seva inclusió i retirada, per avaluar-ne l'impacte en el model. A tal propòsit, SHAP assigna un valor a cada variable en funció de la seva rellevància relativa en la predicció final (Nohara et al., 2022).

Per generalitzar aquestes prediccions generades per cada individu, caldrà, en darrer terme, calcular la mitjana de les prediccions individuals i valors de cadascuna de les variables explicatives, així com, si escau, les seves medianes. D'aquesta manera, el model resultant no només permetrà obtenir una predicció de la probabilitat d'emancipació de cada individu, sinó que també proporcionarà una visió global dels factors que hi contribueixen i de la seva importància en el procés d'emancipació.

5. Resultats obtinguts

En última instància, s'ha dut a terme una síntesi dels resultats obtinguts, inclòs una anàlisi exhaustiva dels diversos models generats, amb l'objecte de detectar patrons, tendències i correlacions significatives per corroborar o refutar les hipòtesis plantejades.

5.1. Model predictiu: aplicació del Random Forest.

Correlació no implica causalitat. En el cas de l'emancipació juvenil, on intervenen múltiples variables, és molt difícil, sinó impossible, determinar amb certesa la causalitat o no causalitat de les característiques sociodemogràfiques en el procés d'emancipació.

A través de l'aplicació metodològica *Random Forest*, s'ha pogut constatar, però, que cada una de les variables de la BBDD té un pes específic en relació amb l'estat d'emancipació.

Això comporta que, malgrat que no es pugui afirmar amb seguretat que una variable causa directament l'emancipació, sí que es pot observar quines variables actuen com a predictores de la independència - i no independència - dels joves.

La precisió del model, que registra un 91,20% (vegeu apèndix IV), demostra la seva capacitat per predir l'estat d'emancipació juvenil amb gran fiabilitat. Per consegüent, com a primera aproximació, es pot inferir que les característiques sociodemogràfiques analitzades, les quals han estat provades mitjançant la seva inclusió i exclusió en el model per verificar el respectiu impacte (vegeu apèndix IV), exerceixen una influència significativa en el procés d'emancipació juvenil.

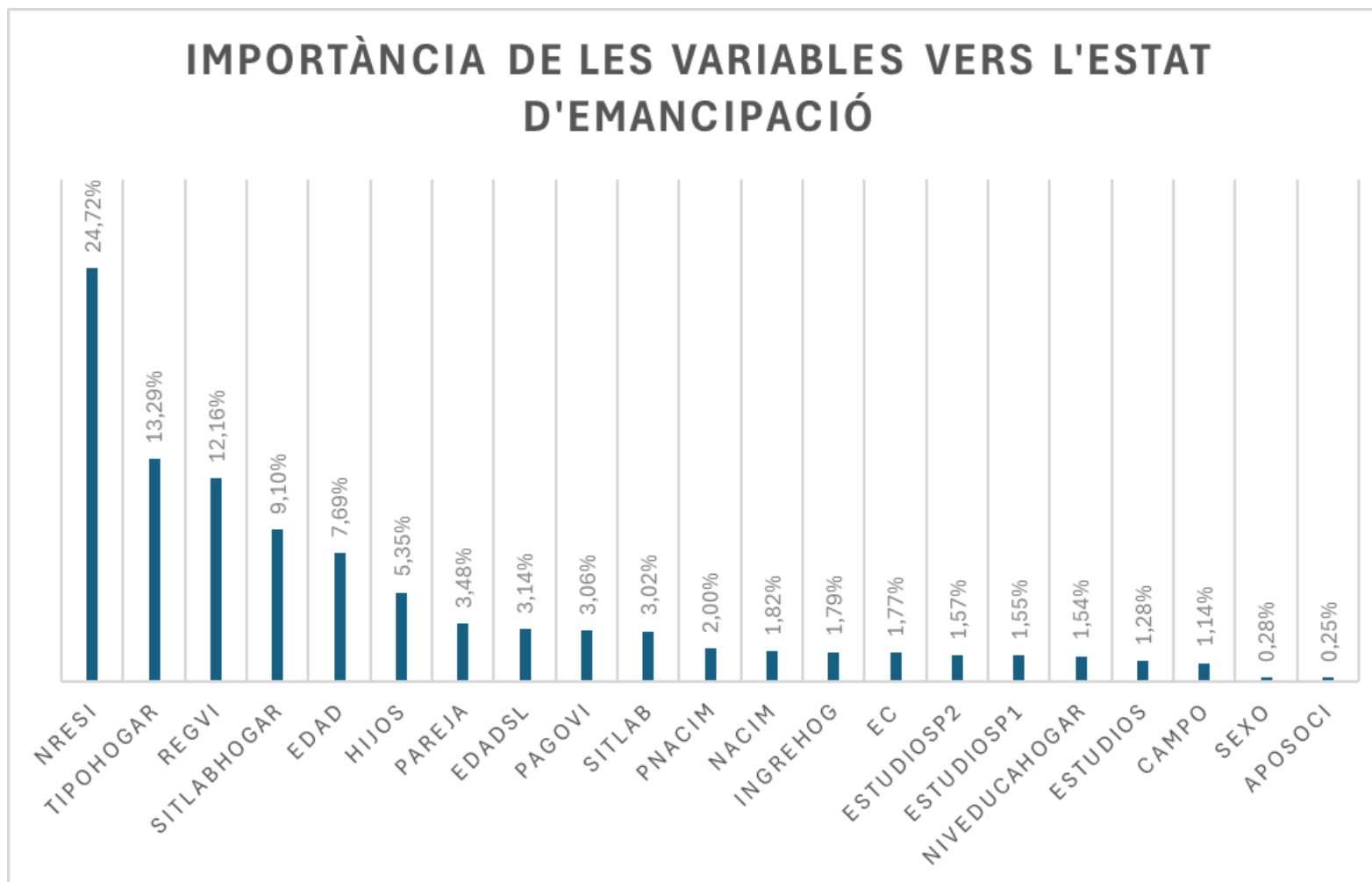
A continuació, es presenten els resultats obtinguts de la metodologia *Random Forest*:

Taula 2. Importància de les variables vers l'estat d'emancipació (tant si sí com si no).

	Variable	Importància
1	nresi	24,72%
2	tipohogar	13,29%
3	regvi	12,16%
4	sitlabhogar	9,10%
5	edad	7,69%
6	hijos	5,35%
7	pareja	3,48%
8	edadsl	3,14%
9	pagovi	3,06%
10	sitlab	3,02%
11	pnacim	2,00%
12	nacim	1,82%
13	ingrehog	1,79%
14	ec	1,77%
15	estudiosp2	1,57%
16	estudiosp1	1,55%
17	niveducahogar	1,54%
18	estudios	1,28%
19	campo	1,14%
20	sexo	0,28%
21	aposoci	0,25%

Font: Elaboració pròpia.

Figura 1. Gràfic de barres de la importància de les variables vers l'estat d'emancipació (tant si sí com si no).



Font: Elaboració pròpia

Recapitulant, els resultats obtinguts mitjançant la metodologia de *Random Forest*, exposen la importància relativa de les característiques sociodemogràfiques en l'emancipació juvenil (tant si es troben emancipats, com si no). Per aquest motiu, les variables es poden agrupar en tres categories segons el seu percentatge d'importància: aquelles amb major pes, aquelles amb pes intermedi, i aquelles amb poc pes.

En el grup de les variables amb major pes, destaca el nombre de residents de l'habitatge (nresi) amb un 24,72%. De totes les variables és la més influent en la predicció de l'emancipació, suggerint que el nombre de persones amb qui un jove conviu pot ser un factor determinant en la seva decisió o capacitat d'emancipar-se. La segona variable més rellevant és el tipus d'habitatge (tipohogar) amb un 13,29%, la qual reflecteix la composició i estructura del nucli familiar, des d'habitatges unipersonals fins a habitatges amb múltiples famílies convivint. També és destacable el règim de tinença (regivi) amb un 12,16%, que pot incloure propietat, lloguer, o altres formes d'ocupació de l'habitatge, repercutint significativament en l'emancipació juvenil. La situació laboral dels membres de l'habitatge (sitlabhogar) amb un 9,10%, i l'edat de l'individu (edad) amb un 7,69%, completen aquest grup, indicant que tant la situació laboral del conjunt de l'habitatge com l'edat són factors predictors de la transició residencial.

El segon grup, el de les variables amb pes intermedi, inclou característiques com tenir fills (hijos) amb un 5,35%, la presència d'una parella (pareja) amb un 3,48%, i l'edat des que es troba en la situació laboral actual (edadsl) amb un 3,14%. Aquestes variables apunten a aspectes més personals i relacionals, que també podrien tenir un impacte en el procés d'emancipació per se. La quota mensual que paga d'hipoteca o lloguer (pagovi) amb un 3,06%, la situació laboral de l'individu (sitlab) amb un 3,02%, i el país de naixement (pnacim) amb un 2,00%, o, per defecte, haver nascut a Espanya o en un altre país (nacim), amb un 1,82%, són altres factors rellevants que evoquen la importància de les responsabilitats financeres, la situació laboral i l'origen geogràfic en la independència o no independència dels joves.

En el grup de les variables amb poc pes, es troba el nivell d'ingressos de l'habitatge com a conjunt (ingrehog), amb un 1,79%, però també l'estat civil de l'individu (ec) amb un 1,77%. En efecte, tenen un impacte menor però encara rellevant en la predicció de l'emancipació. El nivell d'estudis assolit pels progenitors (estudiosp2 i estudiosp1) amb un 1,57% i 1,55% respectivament, el nivell educatiu del nucli familiar (niveducahogar) amb un 1,54%, i el nivell major assolit per l'individu (estudios) amb un 1,28%, així com l'àrea d'estudis d'aquest

darrer (campo) amb un 1,14%, també formen part d'aquest grup, posant en evidència la influència educacional en l'emancipació. Finalment, el gènere de l'individu (sexo) amb un 0,28%, i si té algú amb qui parlar dels seus problemes quotidians o sentiments (aposoci) amb un 0,25%, completen la llista, manifestant un impacte menor però present en la capacitat o decisió d'emancipar-se.

Resumint, mentre que les variables amb major pes, com el nombre de residents de l'habitatge i el tipus d'habitatge, tenen una influència destacada en la predicció de l'emancipació juvenil, les variables amb pes intermedi i poc pes aporten una mica més de clarividència del fenomen estudiat, més tenint en compte la diversitat de factors que hi intervenen. Així, tot i que la correlació no implica causalitat, es pot observar quines variables actuen com a predictores de l'emancipació juvenil.

5.2. Model predictiu: mètode SHAP.

Aplicant el mètode SHAP al *dataset* (vegeu apèndix V), prèviament adequat amb la tècnica *Random Forest*, s'obtenen els següents resultats:

Taula 3. Mitjanes i medianes de les variables més explicatives en la predicció d'emancipació *true*; subgrup d'individus emancipats - I.

	prob_motiv ocambio	hijos	estudios	nresi	regvi	sitlab	pareja	pago vi	edad	naci m	pnaci m	ingre hog	ec	apos oci	sexo	tipoh ogar	sitlab hogar	niveduca hogar
VALOR - MITJANA	82,61%	5,21	6,84	2,20	3,39	2,02	2,82	4,88	25,94	1,08	119,6	3,81	1,11	1,32	1,55	3,09	4,02	1,85
VALOR - MEDIANA	-	6	6	2	3	1	1	5,04	27	1	108	4	1	1	2	3	4	2
Mitjana % - pes en la predicció d'emancipació	-	1,73 %	0,005 %	10,26 %	4,10 %	0,71 %	1,56 %	0,74 %	2,44 %	0,68 %	0,55 %	0,31 %	0,38 %	0,02 %	-0,02 %	5,17 %	2,65%	0,53%

Font: Elaboració pròpia.

Taula 4. Mitjanes i medianes de les variables més explicatives en la predicció d'emancipació *true*; subgrup d'individus emancipats - II.

	campo	edadsl	estudiosp1	estudiosp2
MITJANA	3,69	16,19	4,77	4,83
MEDIANA	3	21	4	4
Mitjana % - pes en la predicció d'emancipació	0,08%	1,12%	0,12%	0,10%

	campo*	edadsl*	estudiosp1*	estudiosp2*
MITJANA²	5,22	22,55	4,97	4,98
MEDIANA²	4,5	23	4	4
Mitjana % - pes en la predicció d'emancipació²	-0,18%	2,46%	0,09%	0,08%

Font: Elaboració pròpia.

² Mitjana/Mediana/Pes en la predicció d'emancipació d'acord amb l'existència de feina, estudis i progenitors de les observacions de joves emancipats, respectivament (vegeu pestanya "CÀLCUL sgs existència variables" de la BBDD_final).

Taula 5. Mitjanes i medianes de les variables més explicatives en la predicció d'emancipació *true*; subgrup d'individus no emancipats - I.

	prob_motiv ocambio	hijos	estud ios	nresi	regvi	sitlab	parej a	pago vi	edad	naci m	pnaci m	ingre hog	ec	apos oci	sexo	tipoh ogar	sitlabh ogar	niveduca hogar
MITJANA	16,70%	5,94	6,49	3,74	2,52	2,66	4,35	5	22,65	1,00	108	4,69	1	1,45	1,49	3,83	4,93	1,72
MEDIANA	-	6	6	4	2	3	6	5,04	22	1	108	5	1	1	1	4	5	2
Mitjana % - pes en la predicció d'emancipació	-	-2,03 %	-0,10 %	-9,99 %	-3,61 %	-0,95 %	-1,05 %	-1,19 %	-2,43 %	-0,70 %	-0,78 %	-0,26 %	-0,62 %	-0,00 2%	-0,01 %	-4,06 %	-3,15%	-0,66%

Font: Elaboració pròpia.

Taula 6. Mitjanes i medianes de les variables més explicatives en la predicció d'emancipació *true*; subgrup d'individus no emancipats - II.

	campo	edadsl	estudiosp1	estudiosp2
MITJANA	3,76	7,99	5,49	5,49
MEDIANA	3	0	4	4
Mitjana % - pes en la predicció d'emancipació	-0,01%	-0,78%	-0,21%	-0,10%

	campo*	edadsl*	estudiosp1*	estudiosp2*
MITJANA³	4,74	22,53	5,56	5,61
MEDIANA³	4	23	5	5
Mitjana % - pes en la predicció d'emancipació³	-0,58%	1,82%	-0,26%	-0,14%

Font: Elaboració pròpia.

³ Mitjana/Mediana/Pes en la predicció d'emancipació d'acord amb l'existència de feina, estudis i progenitors de les observacions de joves emancipats, respectivament (vegeu pestanya "CÀLCUL sgs existència variables" de la BBDD_final).

Pel que fa a la presentació dels resultats de SHAP, aquestes es divideixen en dues seccions per a cada subgrup d'observacions (joves emancipats i no emancipats); una primera part (Taules 3 i 5) que inclou variables relacionades amb aspectes demogràfics i socioeconòmics generals, i una segona part (Taules 4 i 6) que conté variables específiques addicionals. En aquestes últimes, s'inclouen tant la mitjana i la mediana de tots els valors, així com la mitjana i la mediana si es compleix l'existència dels valors als quals s'associen. Per exemple, si un individu té feina, tindrà un valor assignat a la variable "edadsl" (que representa l'edat des de la qual es troba en la situació laboral actual), però si no té feina, aquest valor, òbviament, no existirà. D'aquesta manera, amb els càlculs pertinents, es pot analitzar amb més profunditat com una circumstància, com ara tenir feina i les característiques que es vinculen a aquesta, afecta en la predicció d'emancipació, oferint una visió més completa dels processos que porten a la independència juvenil.

Anàlogament, cada variable inclou el percentatge mitjà del seu pes en la predicció d'emancipació, extret del mètode SHAP, permetent així identificar quines variables exerceixen una influència major en les prediccions realitzades pel model. Tot seguit, es procedirà amb l'anàlisi dels valors que contribueixen a la predicció de l'emancipació juvenil (estat *true*), agrupant-los segons temàtica i, si escau, donant suport amb els resultats del subgrup de joves no emancipats (estat *false*). Sobre això, es verificaran o refutaran, quan sigui convenient, les hipòtesis plantejades.

Situació laboral

La influència de la situació laboral individual (sitlab), especialment l'ocupació a temps complet o parcial, emergeix com un factor relatiu en el procés d'emancipació juvenil. Segons els resultats del mètode SHAP, aquest factor té un impacte moderat en la predicció de l'emancipació juvenil, que se situa en el 0,71%. No obstant, es revela un valor mitjà de 2,02 i una mediana d'1, indicant que la distribució es concentra en les categories 1 i 2, que representen, respectivament, "Ocupat/da - A temps complet" i "Ocupat/da - A temps parcial", essent aquestes les categories que més repercuteixen en la transició residencial.

Simultàniament, l'anàlisi també destaca la variable "sitlabhogar", amb una mitjana de 4,02 i una mediana de 4. És a dir, hi ha una concentració predominant en la categoria que tots els membres de l'habitatge estan actius laboralment. Així doncs, la presència d'ocupació remunerada en tots els membres de l'habitatge presenta un impacte encara més notable en l'emancipació juvenil, amb un pes en la predicció del 2,65%. Això suggereix que, tot i que la

situació laboral individual és rellevant, la dinàmica laboral de tot l'habitatge exerceix una major influència en les decisions o capacitat d'independència dels joves.

De fet, aquesta constatació es reforça quan es compara amb el subgrup de joves no emancipats. En aquest darrer conjunt, la variable "sitlab" presenta un valor mitjà de 2,66 i una mediana de 3 en aquest grup (és a dir, ocupat/da a temps parcial i estudiant, respectivament), amb una contribució negativa del -0,95% en la predicció de l'emancipació juvenil. Així i tot, és important assenyalar que la majoria d'aquests joves no estan actius laboralment, sinó que més aviat es troben en situació d'estudiant. En altres paraules, en el context dels joves que no s'han emancipat, quan es troben en situacions com l'ocupació a temps parcial o són estudiants, emergeix com un punt crític en la tendència de la no emancipació juvenil. Tanmateix, pel que fa a "sitlabhogar" d'aquest subgrup, aquesta variable revela una mitjana de 4,93 i una mediana de 5, categoria que implica que algun membre de l'habitatge no té feina ni està jubilat i algun altre té feina, amb una contribució encara més negativa del -3,15%.

En aquest punt, podria afirmar-se que treballar a temps parcial pot ser tant un pas cap a l'emancipació com un obstacle, segons el context específic de cada individu i el seu entorn familiar. A diferència dels joves que no estan emancipats, els joves que si ho estan, en trobar-se en entorns on tots els membres de l'habitatge estan actius laboralment, treballar a temps parcial no representaria tant una barrera en la seva emancipació, puix que disposarien del suport de la resta de l'habitatge per aconseguir l'estabilitat econòmica.

D'altra banda, aquells joves emancipats que es troben ocupats laboralment, presenten una particularitat, que es visualitza amb la variable "edadsl*" (edat d'ençà que es troba en la situació laboral actual). Els resultats constaten un valor mitjà de 22,55 anys i una mediana de 23 anys en aquest grup, mentre que afavoreix en un 2,46% en la predicció de l'emancipació juvenil. Per tant, la majoria dels joves emancipats que treballen han aconseguit les seves primeres experiències laborals relativament d'hora, més si es considera que aquest subgrup té una mitjana d'edat de 26 anys (vegeu mitjana de la variable "edad"). Entre altres aspectes, això posa de manifest la rellevància de l'accés precoç al món laboral en el procés d'autonomia juvenil.

Després d'analitzar les dades obtingudes, es pot concloure que, des de la perspectiva dels joves emancipats, la situació laboral té un paper quelcom fonamental en el procés d'emancipació; treballar, sigui a temps complet o parcial, està vinculat amb una major

probabilitat d'emancipació. Un supòsit que es consolida pel fet que la majoria dels joves emancipats que ocupen una feina l'han aconseguit relativament aviat, com indica la mitjana de 22,55 anys en la variable "edadsl*". A més, aquests mateixos joves resideixen en habitatges on tots els membres es troben ocupats laboralment, corroborant així la relació entre la situació laboral i l'emancipació juvenil. En perspectiva d'aquests resultats, s'accepta la H1, que sosté que estar actiu laboralment afavoreix la probabilitat d'emancipació juvenil.

Situació educativa

La variable "estudios" representa el nivell d'estudis assolit pels joves, amb categories que abasten des de la falta d'alfabetització fins a nivells educatius més alts, com màsters i doctorats. Entre aquestes categories, la "segona etapa de l'educació secundària amb orientació professional (FP)" i "educació postsecundària no superior" (categories 6 i 7, respectivament) destaquen com les més freqüents entre els joves emancipats.

Tal com s'exposava en la taula d'importàncies del *Random Forest* (vegeu Taula 2), els resultats de l'anàlisi SHAP constaten que la contribució d'aquestes categories és extraordinàriament baixa, pràcticament negligible podria dir-se, amb un índex de tan sols 0,005%. Això suposa que, des del punt de vista de la predicció de l'emancipació residencial, el nivell educatiu té poc o cap impacte observable.

En la mateixa línia, la contribució de les variables "estudiosp1*" i "estudiosp2*", que representen el major nivell d'estudis dels progenitors, és mínima en la predicció de l'emancipació juvenil, amb percentatges del 0,09% i 0,08%, respectivament. Es pot concloure, doncs, que el nivell educatiu dels progenitors tampoc té cap impacte en les decisions d'emancipació dels seus fills. La mateixa regla s'aplica també en el camp d'especialització dels estudis ("campo*"), puix que el seu efecte en la transició residencial és només del -0,18%. És a dir, pràcticament nul.

No obstant, igual que el seu paral·lel en el bloc previ de la situació laboral, el nivell educatiu de tots els membres de l'habitatge sí que sembla tenir algun efecte, però continua sent molt baix igualment. La variable "niveducahogar" mostra un valor mitjà d'1,85 i una mediana de 2, amb una contribució del 0,53% en la predicció de l'emancipació juvenil. De manera que, tot i que el nivell educatiu dels membres de l'habitatge, en aquest cas que algun membre de l'habitatge tingui estudis superiors (valor 2), té un cert impacte, és un de molt limitat en el procés d'emancipació dels joves.

Per tot això, es refuta la hipòtesi H2, que afirma que "haver assolit estudis superiors disminueix la probabilitat d'emancipació", ja que l'anàlisi revela que ni el nivell educatiu individual ni el dels membres de l'habitatge comprenen un impacte significatiu, sigui positiu o negatiu, en l'emancipació dels joves espanyols.

Entorn familiar (estructura de la llar)

Quant al bloc d'entorn familiar, inclou quatre variables interrelacionades que ofereixen una visió completa de les dinàmiques familiars de la joventut espanyola. Per exemple, la variable "hijos" indaga sobre la presència de fills, amb una mitjana que oscil·la el 5,21 i una mediana de 6, cerciorant que una gran part d'aquest segment no en tenen (codificats com 1 per "Sí" i 6 per "No"). Amb tot, la seva contribució a la predicció de l'emancipació és de l'1,73%, plantejant interrogants sobre altres factors del present bloc que podrien incidir més en la independència residencial.

Així, el nombre de residents a l'habitatge ("nresi"), amb un valor mitjà de 2,20 i una mediana de 2, confirma que la majoria dels joves emancipats conviuen, com a mínim, amb una persona més. Es tracta de l'element més determinant en la predicció d'emancipació, amb un considerable 10,26%, destacant-se respecte a la resta de variables analitzades.

No es queda lluny el tipus d'habitatge on resideixen els joves emancipats ("tipohogar"), amb una mitjana de 3,09 i una contribució del 5,17% en la predicció, evidenciant que la majoria dels joves viuen en parella sense fills. De manera anàloga, el fet de tenir parella ja és per si mateix un dels altres predictors de la transició residencial, puix que la mediana es troba en 1 ("Sí") amb una aportació de l'1,56%, significant que un nombre important de joves emancipats tenen parella.

En contrast, els joves no emancipats tendeixen a no tenir parella i a viure en nuclis familiars on hi ha una parella amb fills (és a dir, els mateixos joves no emancipats), amb una mitjana de quatre membres per habitatge. A més, pràcticament no tenen fills propis, fet que actua com a mesura dissuasiva de la seva emancipació. És cert que els joves emancipats tampoc en tenen, però si es comparen les dades, el valor mitjà dels emancipats és de 5,21, mentre que el dels no emancipats és de 5,94. Això posa en evidència que, tot i que la tendència general dels dos grups és no engendrar fills entre els 18 i 29 anys, els joves emancipats són lleugerament més propensos a tenir-ne, fet que acaba incrementant, encara que sigui mínimament, les seves probabilitats d'assolir la plena adultesa.

En conjunt, aquestes variables apunten al fet que l'entorn familiar, particularment el nombre de residents a l'habitatge, és una dimensió rellevant en la decisió d'emancipació dels joves. El nombre predilecte de convivents per a l'emancipació juvenil sembla ser dos, tot i que altres aspectes com tenir fills o parella i el tipus d'habitatge també comprenen un efecte a tenir en compte.

Clar està, en aquesta anàlisi hi ha cert biaix, ja que, al viure en la llar familiar, els joves no emancipats normalment cohabituen amb els pares i/o altres familiars, augmentant així el nombre de membres de l'habitatge i reduint la seva probabilitat d'emancipació en el procés, tal com constata el mètode SHAP. No obstant, aquest biaix no desmenteix la conclusió que l'entorn familiar, concretament el nombre de residents a l'habitatge, és un factor rellevant en la decisió d'independitzar-se. A contraposició dels no emancipats, gran part de joves que sí que ho estan tenen parella, i certament conviuen amb ella en el nou habitatge, posant de manifest que la convivència amb una altra persona és un element clau per a l'emancipació juvenil. Per consegüent, es verifica la H3 perquè queda demostrat que conviure amb una o més persones augmenta la probabilitat d'abandonar la llar familiar.

Condicions residencials i econòmiques

La variable "regvi" descriu el règim de tinença de l'habitatge, amb un valor mitjà de 3,39 i una mediana de 3, exposant que els joves espanyols que no es troben en la llar d'origen s'emancipen en habitatges propis amb pagaments pendents (hipoteca), però també llogats, tot i que la balança s'inclina cap a la primera opció. A més, amb una contribució del 4,10% en la predicció de l'emancipació, aquest punt denota un pes prou rellevant en la decisió dels joves de viure pel seu compte.

És sorprenent que aquesta sigui la tendència entre els joves espanyols emancipats, ja que, per regla general, es podria esperar que optessin per lloguers a causa de la seva situació econòmica, sovint menys estable. Podria teoritzar-se que, entre aquest grup, hi ha certa disposició o capacitat per assumir compromisos financers a llarg termini, possiblement secundats per suport familiar o per una perspectiva més optimista de l'estabilitat econòmica futura. Podria considerar-se també un canvi de paradigma, en el qual els joves no s'emancipen fins que no accedeixen a una hipoteca. Tot i això, cal incidir que les observacions, al tractar-se de tot el país, podrien no reflectir la realitat de certes comunitats autònomes, però sí del còmput global.

Pel que fa a els ingressos mensuals nets de l'habitatge ("ingrehog"), amb un valor mitjà de 3,81 i una mediana de 4, s'extreu que una gran proporció dels habitatges amb joves emancipats es troben en una franja d'ingressos de 1.500 € a 2.000 € al mes. En certa manera l'interval remet a una estabilitat econòmica en aquest segment de la població, permetent-los assumir el compromís de l'emancipació. Malgrat tot, la contribució d'aquesta variable a la predicció de l'emancipació és només del 0,31%, reflectint un impacte moderat en les seves decisions.

En relació amb la variable "pagovi", que correspon a la quota mensual pagada d'hipoteca o lloguer, els valors mitjans indiquen que molts joves paguen entre 400 € i 500 € al mes per mantenir el seu habitatge. Una informació que podria suggerir una càrrega financera significativa, però no desproporcionada, si es compara amb la mitjana dels ingressos de l'habitatge. En efecte, però, amb una contribució del 0,74% en la predicció, la variable també revela exercir un efecte limitat en l'establiment de l'habitatge propi.

Característiques idiosincràtiques

En els condicionants idiosincràtics, destaquen aquells relacionats amb la nacionalitat, amb una contribució del 0,68% i del 0,55%, depenent la variable, en la predicció de la seva emancipació. Tant la variable "nacim", que especifica si han nascut a Espanya o a un altre país, com "pnacim", que precisa el país de naixement, amb una mediana que correspon a Espanya (valor 108), subratllen la predominança del naixement a Espanya entre els joves emancipats. Tanmateix, cal destacar que aquest descobriment podria tenir un biaix inherent en la recollida de dades de l'ECEPOV, ja que potser no s'ha tingut en compte una proporció equilibrada d'espanyols i estrangers.

Per contrapart, els resultats assenyalen que els joves, quan s'emancipen, ho fan amb l'estat civil solter/a ("ec" - valor 1, amb una contribució del 0,38% en la predicció), fet que no és d'estranyar amb les evidències extretes de la revisió literatura respecte al descens del nombre de casaments. Alhora, tot i que compten amb una persona amb qui parlar dels seus problemes quotidians o sentiments ("aposoci"), sembla que no és un factor prou determinant per a la seva emancipació (0,02%).

L'edat de l'individu, no obstant, juga un rol més prominent, ja que els joves emancipats se situen en la franja dels 26 als 27 anys, amb una contribució significativa a la seva emancipació del 2,44%. Paral·lelament, els joves no emancipats es troben entre els 22 i els 23 anys, fet que repercuteix negativament en la seva probabilitat d'emancipació amb un

-2,43%. Així doncs, aquests resultats corroboren les elevades edats d'emancipació d'Espanya, que ascendien als 30,03 anys en el transcurs del 2022 (Eurostat, 2023).

Finalment, el gènere per si mateix no sembla contribuir a l'emancipació, ja que tant pels joves emancipats, amb un valor mitjà d'1,55 i una contribució de -0,02%, com pels no emancipats, amb una mitjana d'1,49 i una contribució de -0,01%, no presenta un impacte substancial. Sí que és cert, però, que hi ha una petita variació entre els joves emancipats, que són lleugerament més dones que homes, i viceversa amb els joves no emancipats. No obstant això, la contribució en la predicció d'emancipació és tan subtil, que es pot afirmar que el gènere per se no és un dels determinants en la transició residencial. En vista d'aquests resultats, es refuta la H4, puix que el gènere en si, almenys per aquest conjunt de dades, no sembla modificar la probabilitat d'emancipació.

6. Conclusions de la recerca

Els resultats d'aquesta investigació permeten aprofundir en el paradigma de l'emancipació juvenil a Espanya i els seus determinants. Segons el model predictiu elaborat, es constata una clara correlació entre la situació laboral dels joves i la seva probabilitat d'emancipació. Estar actiu laboralment, sigui a temps complet o parcial, s'associa amb una major probabilitat d'emancipació, tot i que l'entrada prematura al món laboral pot facilitar més la transició residencial. No obstant, la dinàmica laboral de tot l'habitatge, en la qual tots els membres es troben ocupats laboralment, exerceix una influència encara més significativa en la independència dels joves.

Comparant-ho amb la literatura existent, es comprova que altres acadèmics arriben a conclusions similars, especialment pel que fa a l'estat ocupacional (Bellart & Oller, 2005; Becker et al., 2010; Stone et al., 2011; Moreno, 2012; Colom & Molés, 2016; Donat & Martín, 2020). A més, ingressar abans al món laboral també s'identifica en la literatura com un factor que afavoreix l'emancipació juvenil, malgrat no indagar en les diferències de gènere com Miret (2006). D'altra banda, si bé el fet que tots els membres de l'habitatge estiguin ocupats laboralment també sembla augmentar les oportunitats d'emancipació, aquest aspecte no ha sigut documentat, almenys en el context espanyol. Així, aquesta troballa en relació amb la dinàmica laboral i l'emancipació juvenil aporta una nova perspectiva al camp d'estudi.

Es constata, a més, que la situació educacional, com el camp d'estudis en el que un s'especialitza o el nivell d'estudis assolit, no exerceixen cap efecte, ni positiu ni negatiu, en la transició residencial dels joves espanyols, que són uns resultats que coincideixen amb investigacions prèvies com les de Moreno (2018) o Miret (2006). No obstant això, es reafirma que l'entorn familiar, específicament el nombre de residents a l'habitatge, és un factor clau en la decisió d'independitzar-se, ja que viure amb una o més persones augmenta la probabilitat d'abandonar la llar familiar. En contrast amb els joves no emancipats, aquells que si ho estan sovint tenen parella i cohabiten amb ella en el nou habitatge, encara que sense tenir fills ni estar casats. Uns resultats que exemplifiquen les troballes de Molgat (2007), Moreno (2018) i South & Lei (2015) que varen determinar que la presència d'una parella fomenta l'establiment de l'habitatge propi.

En darrer terme, tot i l'existència d'estudis que demostren el contrari (Elder & Shanahan, 2007; Moreno, 2012), sembla que, en l'actualitat, el gènere per se no modifica el procés d'emancipació a Espanya. Altres factors sociodemogràfics que si propicien la transició residencial, però, seria l'edat; en el context espanyol, com més jove és una persona, menor és la probabilitat que es produeixi la seva emancipació, i viceversa. Una circumstància àmpliament contrastada en la literatura, com queda reflectit en investigacions com la de Billari & Liefbroer (2007).

Malauradament, a causa de limitacions de la BBDD emprada, no s'han pogut extreure suficients evidències de l'efecte de la renda ni tampoc de la nacionalitat vers l'emancipació espanyola, tot i haver-ne determinat la seva contribució en la predicció. Malgrat haver descobert que els joves emancipats gaudeixen de certa estabilitat econòmica en relació amb el cost de l'habitatge, la variable relativa als ingressos correspon al conjunt de l'habitatge. Com que no està presentada de manera individualitzada no es pot saber amb exactitud quina quantitat aporta cada individu sense entrar en biaixos, fet que estudiar-ne els efectes, tant en els joves emancipats com no emancipats, planteja diversos desafiaments que no s'han pogut abordar. En el que això respecte, tampoc s'ha pogut inquirir en el factor nacionalitat perquè, en la mostra recopilada, no es va procurar una representació suficient d'espanyols nascuts a Espanya i d'altres amb diferent origen.

Una altra limitació d'aquesta investigació és la falta d'una sèrie temporal en la base de dades utilitzada, la qual cosa impedeix contrastar els resultats i avaluar l'efectivitat del model predictiu plantejat. És més, l'absència de dades longitudinals dificulta l'anàlisi de les variacions i tendències en l'emancipació juvenil al llarg dels anys.

D'igual manera, es fa patent la dependència de la base de dades utilitzada, ja que, en ser de font secundària, ha impedit l'anàlisi de variables no presents en aquesta, però que han estat estudiades anteriorment (mencionades en l'apartat 2.1.) i podrien tenir un impacte significatiu en l'emancipació juvenil a Espanya.

Amb tot, es pot afirmar que aquesta investigació obra noves línies de recerca, sobretot des d'un punt de vista d'una nova aplicació metodològica (*Random Forest*). Tot i haver-se utilitzat en múltiples *papers* per la seva consistència en els resultats, el *Random Forest* no s'havia emprat, fins ara, en l'estudi de l'emancipació juvenil en l'àmbit d'Espanya. A part, tot i que no era el propòsit del treball, no s'ha explorat les diferències en els patrons de comportament de gènere o cohorts d'edat en relació amb l'emancipació, com van fer Miret (2006), Molgat (2007), Colom & Molés (2016) i Moreno (2018) en el primer cas, i Stone et al. (2011) en el segon. Un altre aspecte que roman obert és la pregunta sobre per què, en major proporció, els joves espanyols opten per emancipar-se en habitatges propis amb pagaments pendents (hipoteca), en lloc d'optar per habitatges de lloguer. Definitivament, això també encamina cap a noves vies de recerca potencials.

La present investigació posa de manifest la complexitat de la situació dels joves espanyols en relació amb la seva emancipació, la qual es troba en un moment crític. Davant d'aquest escenari, esdevé evident la necessitat d'intervenció pública amb polítiques que abordin les necessitats reals d'aquest segment de la població. Certament, la investigació contínua és l'única via per comprendre millor les barreres que poden trobar els joves en el procés d'emancipació. No obstant això, el primer exercici de reflexió que cal fer per abordar aquesta conjuntura és reconèixer que, fins al moment present, les mesures adoptades no han estat efectives, sinó més aviat tot el contrari.

7. Referències

- Alberdi, I. (2014). El retraso de la procreación y el mantenimiento de la juventud. *Metamorfosis. Revista del Centro Reina Sofía sobre Adolescencia y Juventud*, 9-25. <https://dialnet.unirioja.es/servlet/articulo?codigo=6163235>
- Aparicio, A. & Oppedisano, V. (2015). Fostering Household Formation: Evidence from a Spanish Rental Subsidy. *The B.E. Journal of Economic Analysis & Policy*, 15(1), 53-84. <https://doi.org/10.1515/bejeap-2014-0003>
- Aronow, P. M., & Samii, C. (2016). Does Regression Produce Representative Estimates of Causal Effects? *American Journal of Political Science*, 60(1), 250–267. <https://doi.org/10.1111/AJPS.12185>
- Ballesteros, J. C., Megías, I. & Rodríguez, E. (2012). Jóvenes y emancipación en España. *Fundación de Ayuda contra la Drogadicción*, 1-140. <https://www.centroreinasofia.org/publicacion/jovenes-y-emancipacion-en-espana/>
- Baranowska, A., & Gebel, M. (2010). The determinants of youth temporary employment in the enlarged europe. *European Societies*, 12(3), 367–390. <https://doi.org/10.1080/14616690903165434>
- Beck, U., Lash, S., Leiss, W., Ritter, M. & Wynne, B. (1994). Risk Society, Towards a New Modernity. *Canadian Journal of Sociology / Cahiers Canadiens de Sociologie*, 19(4), 544. <https://doi.org/10.2307/3341155>
- Becker, S. O., Bentolila, S., Fernandes, A. & Ichino, A. (2010). Youth emancipation and perceived job insecurity of parents and children. *Journal of Population Economics*, 23(3), 1047-1071. <https://doi.org/10.1007/s00148-008-0224-5>
- Bellart, C. & Oller, J. (2005). El acceso de los jóvenes a la vivienda: Una cuestión todavía no resuelta. *Documentación Social*, 138, 191-206. <https://dialnet.unirioja.es/servlet/articulo?codigo=1373197>
- Bernardi, F. (2007). Movilidad social y dinámicas familiares. Una aplicación al estudio de la emancipación familiar en España. *Revista Internacional De Sociología*, 65(48), 33–54. <https://doi.org/10.3989/ris.2007.i48.67>

Billari, F. C. & Liefbroer, A. C. (2007). Should I Stay or Should I Go?: The Impact of Age Norms on Leaving Home. *Demography*, 44(1), 181-198. <https://doi.org/10.1353/dem.2007.0000>

Billari, F. C. & Liefbroer, A. C. (2010). Towards a new pattern of transition to adulthood? *Advances in Life Course Research*, 15(2-3), 59-75. <https://doi.org/10.1016/J.ALCR.2010.10.003>

Blaauboer, M. & Mulder, C.H. (2010). Gender differences in the impact of family background on leaving the parental home. *Journal of Housing and the Built Environment*, 25(1), 53-71. <https://doi.org/10.1007/s10901-009-9166-9>

BOE (2007). *Real Decreto 1472/2007, de 2 de noviembre, por el que se regula la renta básica de emancipación de los jóvenes*. Agencia Estatal Boletín Oficial del Estado. Recuperat el 28/02/2024: <https://www.boe.es/eli/es/rd/2007/11/02/1472>

BOE (2022). *Real Decreto 42/2022, de 18 de enero, por el que se regula el Bono Alquiler Joven y el Plan Estatal para el acceso a la vivienda 2022-2025*. Agencia Estatal Boletín Oficial del Estado. Recuperat el 28/02/2024: <https://www.boe.es/eli/es/rd/2022/01/18/42>

Cantó, S. O. (2019). Desigualdad, redistribución y políticas públicas: ¿hay una brecha generacional? *ICE, Revista de Economía*, (908). <https://doi.org/10.32796/ICE.2019.908.6826>

Carballo-Cruz, F. (2011). Causes and consequences of the Spanish economic crisis: Why the recovery is taken so long? *Panoeconomicus*, 58(3), 309-328. <https://doi.org/10.2298/PAN1103309C>

CJE (2023). *2023-1SEM_Informe-estatal*. Consejo de la Juventud de España (CJE). Recuperat el 06/02/2024: <https://www.cje.org/observatorio-de-emancipacion/#268-269-primer-semester-2023>

Colom, M. C. & Molés, M. C. (2016). Emancipación familiar en España: Análisis del comportamiento de los jóvenes en 1990, 2000 y 2010. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 22, 120-138. <https://dialnet.unirioja.es/servlet/articulo?codigo=5778894>

Comas, D. (2015). La emancipación de personas jóvenes en España. El túnel del miedo. *Revista Metamorfosis: Revista del Centro Reina Sofía sobre Adolescencia y Juventud*, 2, 7-24. <https://dialnet.unirioja.es/servlet/articulo?codigo=6163225>

Cordón, J. (1997). Youth residential independence and autonomy: A comparative study. *Journal of Family Issues*, 18, 576–607. <https://doi:10.1177/019251397018006002>

Delgado, J. M. (2018). Más allá del tópico de la España vacía: una geografía de la despoblación. *Informe España 2018*, 232–295. <https://uvadoc.uva.es/handle/10324/38457>

de Miguel, H. (2024). *Solo el 16,3% de los jóvenes se va de casa: ¿por qué es casi imposible emanciparse en España?* https://www.larazon.es/sociedad/solo-163-jovenes-casa-que-casi-imposible-emanciparse-es-pana_2024011665a68ffe872b820001a75712.html

Donat, M. A, & Martín, M. D (2020). Juventud andaluza y emancipación. ¿Independencia residencial y/o autosuficiencia económica? *Papers. Revista de Sociologia*, 105(1), 57–81. <https://doi.org/10.5565/rev/papers.2616>

Duncan, S. & Phillips, M. (2008). *New families? Tradition and change in partnering and relationships*. British Social Attitudes: the 24th Report. London: SAGE Publications Ltd. <http://hdl.handle.net/10454/2239>

Echaves, A. G. (2017). El difícil acceso de los jóvenes al mercado de vivienda en España: precios, regímenes de tenencia y esfuerzos. *Cuadernos de Relaciones Laborales*, 35(1), 127–149. <https://doi.org/10.5209/CRLA.54986>

Elder, G. H., & Shanahan, M. J. (2007). The Life Course and Human Development. *Handbook of Child Psychology: Volume I. Theoretical Models of Human Development*, 665-715. <https://doi.org/10.1002/9780470147658.CHPSY0112>

Esping-Andersen, G. (1999). *Social Foundations of Postindustrial Economies*. Oxford University Press. <https://doi.org/10.1093/0198742002.001.0001>

Eurostat (2023). *Estimated average age of young people leaving the parental household by sex*. Eurostat. Recuperat el 09/12/2023: https://ec.europa.eu/eurostat/databrowser/view/YTH_DEMO_030__custom_7245840/bookmark/table?lang=en&bookmarkId=8187eeda-6c5b-48fc-ba39-4e8cf6eee71b

Gentile A. (2015). Rental subsidy and the emancipation of young adults in Spain. *International Journal of Housing Policy*, 16(2), 243-254. <https://doi.org/10.1080/14616718.2015.1119359>

Ghidoni, M. (2002). Determinants of young Europeans' decision to leave the parental household. *Royal Economic Society Annual Conference 2002*, 85. <https://ideas.repec.org/p/ecj/ac2002/85.html>

Gil Calvo, E. (2002). Emancipación tardía y estrategia familiar. *Revista de Estudios de Juventud*, 9-18. <http://www.svi.injuve.es/sites/default/files/articulo1.pdf>

Gil Calvo, E. (2014). El síndrome de dependencia familiar como encuadre (framing) de la agenda juvenil. *Revista Metamorfosis: Revista del Centro Reina Sofía sobre Adolescencia y Juventud*, 0, 47-64. <https://dialnet.unirioja.es/servlet/articulo?codigo=6163237>

González-Sanguino, C., Ausín, B., Castellanos, M. Á., Saiz, J., López-Gómez, A., Ugidos, C., & Muñoz, M. (2020). Mental health consequences during the initial stage of the 2020 Coronavirus pandemic (COVID-19) in Spain. *Brain, Behavior, and Immunity*, 87, 172–176. <https://doi.org/10.1016/J.BBI.2020.05.040>

Grömping, U. (2012). Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *The American Statistician*, 63(4), 308–319. <https://doi.org/10.1198/tast.2009.08199>

Hartshorn, S. (2016). *Machine Learning With Random Forests And Decision Trees: A Visual Guide For Beginners* [Kindle version].

INE (2023). *Encuesta de Características Esenciales de la Población y Viviendas - Año 2021*. Instituto nacional de estadística (INE). Recuperat el 02/02/2024: https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177092&menu=ultiDatos&idp=1254735572981

INE (s.d.a) *Principales series de población desde 1998*. Instituto nacional de estadística (INE). Recuperat el 04/02/2024: <https://www.ine.es/jaxi/Tabla.htm?path=/t20/e245/p08/l0/&file=01003.px&L=0>

INE (s.d.b) *Nivel educativo alcanzado de la población joven (18 a 24 años) por nivel de educación, CCAA y periodo*. Instituto nacional de estadística (INE). Recuperat el 23/04/2024: <https://www.ine.es/jaxi/Datos.htm?path=/t00/ICV/dim4/I0/&file=41301.px>

Julià, A. C., Moreno, A. M., López, A. P., Crespo, E. B. & Segado, S. C. (2012). *La transició dels joves a la vida adulta: Crisi econòmica i retard en l'emancipació*. Obra social "la Caixa". Recuperat el 07/12/2023: <https://prensa.fundacionlacaixa.org/wp-content/uploads/2019/09/25829.pdf>

Kneale, D. & Joshi, H. (2008). Postponement and childlessness: Evidence from two British cohorts. *Demographic Research*, 19(58), 1935-1968. <https://doi.org/10.4054/DemRes.2008.19.58>

Kuhar, M., & Reiter, H. (2014). Leaving home in Slovenia: A quantitative exploration of residential independence among young adults. *Journal of Adolescence*, 1–11, <https://doi:10.1016/j.adolescence.2014.05.011>

Levin, I. (2004). Living Apart Together: A New Family Form. *Current Sociology*, 52(2), 223–240. <https://doi.org/10.1177/0011392104041809>

Maeso, F., & Mendez, I. (2008). The role of partnership status and expectations on the emancipation behaviour of spanish graduates. *Documentos de Trabajo (CEMFI)*, 12, 1. <https://dialnet.unirioja.es/servlet/articulo?codigo=2865259>

Manacorda, M., Moretti, E., Manacorda, M., & Moretti, E. (2006). Why do Most Italian Youths Live with Their Parents? Intergenerational Transfers and Household Structure. *Journal of the European Economic Association*, 4(4), 800–829. <https://econpapers.repec.org/RePEc:tpr:jeurec:v:4:y:2006:i:4:p:800-829>

Merino, R., & Garcia, M. (2006). Emancipation enlargement and the acquisition of autonomy by young people in Catalonia. *YOUNG*, 14(1), 33-47. <https://doi.org/10.1177/1103308806059812>

Miret, P. (2006). Escolarización, mercado de trabajo y emancipación familiar en España: Un análisis longitudinal a escala de Comunidad Autónoma. *Papeles de Geografía*, 43, 73-92. <https://revistas.um.es/geografia/article/view/43591>

Molgat, M. (2007). Do transitions and social structures matter? How 'emerging adults' define themselves as adults. *Journal of Youth Studies*, 10(5), 495–516. <https://doi.org/10.1080/13676260701580769>

Moreno, A. (2012). The Transition to Adulthood in Spain in a Comparative Perspective. *YOUNG*, 20(1), 19–48. <https://doi.org/10.1177/110330881102000102>

Moreno, A. (2018). The youth emancipation in Spain: a socio-demographic analysis. *International Journal of Adolescence and Youth*, 23(4), 496-510. <https://doi.org/10.1080/02673843.2018.1438299>

Nohara, Y., Matsumoto, K., Soejima, H., & Nakashima, N. (2022). Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Computer Methods and Programs in Biomedicine*, 214, 106584. <https://doi.org/10.1016/j.cmpb.2021.106584>

Noriega, C. (2023). *Emancipación tardía: ¿qué retiene a los jóvenes en casa de sus padres?* El Debate. Recuperat el 23/01/2024: https://www.eldebate.com/familia/20231208/emancipacion-tardia-retiene-jovenes-casa-padres_157543.html

OECD (2024). *Unemployment rate by age group (indicator)*. Organització per a la Cooperació i el Desenvolupament Econòmic (OECD). Recuperat el 04/03/2024: <https://data.oecd.org/unemp/unemployment-rate-by-age-group.htm#indicator-chart>

Perugini, C., & Signorelli, M. (2010). Youth labour market performance in European regions. *Economic Change and Restructuring*, 43(2), 151–185. <https://doi.org/10.1007/S10644-009-9082-8/TABLES/8>

Requena, M. (2006a). Familia, convivencia y dependencia entre los jóvenes españoles. *Panorama Social*, 3, 64–77. <https://portalcientifico.uned.es/documentos/5f88eab329995259ef2944f1>

Requena, M. (2006b). Juventud y dependencia familiar en España. *Revista de Juventud*, 58, 10–23. https://www.academia.edu/40405019/Juventud_y_dependencia_familiar_en_Espa%C3%B1a_a_Juventud_y_dependencia_familiar_en_Espa%C3%B1a?auto=download

Rosas, M., Chacín, F., García, J., Ascanio, M. & Cobo, M. (2006). Modelos de regresión lineal múltiple en presencia de variables cuantitativas y cualitativas para predecir el rendimiento estudiantil. *Revista de la Facultad de Agronomía*, 23(2), 197-214. http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0378-78182006000200007&lng=es&lng=es.

Royo, S. (2020). Responding to COVID-19: The Case of Spain. *European Policy Analysis*, 6(2), 180–190. <https://doi.org/10.1002/EPA2.1099>

Sala i Roca, C. (2022). Una breu mirada jurídica del “Bono Alquiler Joven” a través de la doctrina del Tribunal Constitucional. *Housing: Revista de La Cátedra de Vivienda de La Universidad Rovira i Virgili*, 17, 20–24. <https://dialnet.unirioja.es/servlet/articulo?codigo=8648830>

Salazar, J. J., Garland, L., Ochoa, J., & Pyrcz, M. J. (2022). Fair train-test split in machine learning: Mitigating spatial autocorrelation for improved prediction accuracy. *Journal of Petroleum Science and Engineering*, 209, 109885. <https://doi.org/10.1016/j.petrol.2021.109885>

Soler, A. M., Torres-Tellez, J., & Ayala García, I. H. (2023). Youth Emancipation and the Labour Market in Spain. *Panoeconomicus*, 70(3), 429–454. <https://doi.org/10.2298/PAN191125016S>

Soler, P., Planas, A. & Feixa, C. (2014). Young people and youth policies in Spain in times of austerity: between juggling and the trapeze. *International Journal of Adolescence and Youth*, 19(1), 62-78. <https://doi.org/10.1080/02673843.2013.823552>

South, S. J. & Lei, L. (2015). Failures-to-launch and boomerang kids: Contemporary determinants of leaving and returning to the parental home. *Social Forces*, 94(2), 863-890. <https://doi.org/10.1093/sf/sov064>

Spatscheck, C. (2007). Youth Cultures between Control and Emancipation—Interpretations from a Western European Perspective. *International Social Work: Social Problems, Cultural Issues and Social Work Education*, 119. <https://books.google.es/books?hl=es&lr=&id=DfeZDwAAQBAJ&oi=fnd&pg=PA119&dq=youth+emancipation+europe&ots=tQL97YJryl&sig=mFoddgefB6sfKnRk7wh2IDdT2oY#v=onepage&q=youth%20emancipation%20europe&f=false>

Stone, J., Berrington, A. & Falkingham, J. (2011). The changing determinants of UK young adults' living arrangements. *Demographic Research*, 25, 629–666. <https://www.demographic-research.org/volumes/vol25/20/25-20.pdf>

Tobío, C. (2001). Marriage, cohabitation and the residential independence of young people in Spain. *International Journal of Law, Policy and the Family*, 15(1), 68–87. <https://doi.org/10.1093/LAWFAM/15.1.68>

Torres, A. (2023). *La OCDE alerta de que la difícil transición de los jóvenes a una vida adulta independiente lastra el potencial de España*. El País. Recuperat el 14/02/2024: <https://elpais.com/sociedad/2023-10-25/la-ocde-alerta-de-que-la-dificil-transicion-de-los-jovenes-a-una-vida-adulta-independiente-lastra-el-potencial-de-espana.html>

Toulemon, L. (1997). Cohabitation is Here to Stay. *Population: An English Selection*, 9, 11–46. <http://www.jstor.org/stable/2953824>

Úbeda, M., Cabasés, M. À., Sabaté, M., & Strecker, T. (2020). The Deterioration of the Spanish Youth Labour Market (1985–2015): An Interdisciplinary Case Study. *YOUNG*, 28(5), 544-563. <https://doi.org/10.1177/1103308820914838>

Yates, T. M., & Grey, I. K. (2012). Adapting to aging out: Profiles of risk and resilience among emancipated foster youth. *Development and Psychopathology*, 24(2), 475–492. <https://doi.org/10.1017/S0954579412000107>

8. Apèndixs

8.1. Apèndix I

Figura X. Tractament de les dades de la BBDD inicial (ECEPOV) mitjançant Stata - I.

```
. import delimited "C:\Users\Jordi\Desktop\BBDD\Encuesta de características esenciales de la población y las viviendas
(ECEPOV)\datos_2021\ECEPOVhogar_2021.csv", clear

keep iden idq_pv idq_mun npv motivocambio sexo edad nacim provn munin pncim pnc1 pnc2 nac nprogen reprogen1 hogp1 vivoprogen1 nacimp1 pncimp1
pnc1p1 pnc2p1 estudiosp1 reprogen2 hogp2 vivoprogen2 nacimp2 pncimp2 pnc1p2 pnc2p2 estudiosp2 nucleo tiponucleo sexo_pareja nacion_pareja
union_pareja estudios_pareja sitlab_pareja situhogar hijos_nucleo tipohogar nachogar niveducahogar sitlabhogar

keep if motivocambio==" " | motivocambio=="04"
tab motivocambio
replace motivocambio="emancipado" if motivocambio=="04"
replace motivocambio="no emancipado" if motivocambio==" "
tab motivocambio
summarize edad
keep if edad>17 & edad<30
tab edad
egen localizador = concat(iden edad sexo)
save ECEVOPhogar.dta, replace
```

Font: Elaboració pròpia.

Figura X. Tractament de les dades de la BBDD inicial (ECEPOV) mitjançant Stata - II.

```
. import delimited "C:\Users\Jordi\Desktop\BBDD\Encuesta de características esenciales de la población y las viviendas (ECEPOV)\datos_2021\ECEPOVadultos_2021.csv", clear

keep iden npv ec estudios edadestud campo sitlab edadsl flexi lugtrab hijos nhijos pareja aposoci aposocihij aposocipama aposociher aposociotpar aposociotro lresmd

merge m:m iden npv using "ECEVOPhogar.dta"

keep if _merge==3

save BASEBB.dta

use BASEBB.dta, clear

drop _merge

save BASEBB.dta, replace
```

Font: Elaboració pròpia.

Figura X. Tractament de les dades de la BBDD inicial (ECEPOV) mitjançant Stata - III.

```
. import delimited "C:\Users\Jordi\Desktop\BBDD\Encuesta de características esenciales de la población y las viviendas (ECEPOV)\datos_2021\ECEPOVvivienda_2021.csv", clear

keep iden nresi regvi pagovi ingrehog

merge m:m iden using "BASEBB.dta", force

keep if _merge==3

save BASEBBdefinitive.dta
```

Font: Elaboració pròpia.

8.2. Apèndix II

Taula X. Relació de les 65 variables i corresponents categories de la BBDD de la recerca.

Variable	Descripció	Categoria assignada	
iden	Número d'identificador del qüestionari	-	
idq_pv	Codi de la província de residència	51 codis de totes les províncies + 1 codi NS/NC	
idq_mun	Codi del municipi de residència de més de 50.000 habitants o capital de província de residència	2 codis principals; un per municipi de <50.000 hab. i un altre per NS/NC +1 codi per municipi	
npv	Número de l'ordre de la persona en l'habitatge	-	
motivocambio	Motiu de l'últim canvi de residència	0	No emancipat
		1	Emancipat (a la BBDD de l'ECEPOV tenia inicialment assignada la categoria 4)
sexo	Gènere de l'individu enquestat	1	Home
		2	Dona
edad	Edat de l'individu enquestat	-	
nacim	Lloc de naixement	1	Espanya
		2	Un altre país
provsn	Província espanyola de naixement	51 codis de totes les províncies + 1 codi NS/NC	
munin	Municipi espanyol de naixement	2 codis principals; un per municipi de <50.000 hab. i un altre per NS/NC +1 codi per municipi	
pnacim	País de naixement	201 codis de tots els països + 1 codi NS/NC	
pnac1	País de nacionalitat 1	201 codis de tots els països + 1 codi NS/NC	
pnac2	País de nacionalitat 2	201 codis de tots els països + 1 codi NS/NC	
nac	Nacionalitat: espanyola o estrangera	1	Espanyola

		2	Estrangera
nprogen	Número de progenitors que té o ha tingut	0	NS/NC
		1	No ha tingut
		2	Un
		3	Dos
		0	NS/NC
reprogen1	Relació de parentesc amb el seu progenitor 1	1	Pare
		2	Mare
		0	NS/NC
hogp1	El progenitor 1 resideix en l'habitatge	1	Sí
		6	No
		0	NS/NC
vivoprogen1	El progenitor 1 continua viu	1	Mort
		2	Viu
		0	NS/NC
nacimp1	Lloc de naixement del progenitor 1	1	Espanya
		2	Un altre país
		0	NS/NC
pnacim1	País de naixement del progenitor 1	201 codis de tots els països + 1 codi NS/NC	
estudiosp1	Major nivell d'estudis assolit pel progenitor 1	0	NS/NC
		1	No sap llegir o escriure
		2	Sap llegir i escriure però va anar menys de 5 anys a l'escola
		3	Educació primària completa o va anar més de 5 anys a l'escola
		4	Primera etapa de l'educació secundària

		5	Segona etapa de l'educació secundària amb orientació general (Batxiller)
		6	Segona etapa de l'educació secundària amb orientació professional (FP)
		7	Educació postsecundària no superior
		8	Ensenyament de formació professional, arts plàstiques i disseny i esportistes de grau superior i equivalents
		9	Graus universitaris de fins a 240 crèdits, diplomats universitaris, títols propis d'universitat d'especialització i similars.
		10	Graus universitaris de més de 240 crèdits i llicenciats
		11	Màster
		12	Doctorat universitari
pnac1p1	País de nacionalitat 1 del progenitor 1	201 codis de tots els països + 1 codi NS/NC	
pnac2p1	País de nacionalitat 2 del progenitor 1	201 codis de tots els països + 1 codi NS/NC	
reprogen2	Relació de parentesc amb el seu progenitor 2	0	NS/NC
		1	Pare
		2	Mare
hogp2	El progenitor 2 resideix en l'habitatge	0	NS/NC
		1	Sí
		6	No
vivoprogen2	El progenitor 2 continua viu	0	NS/NC
		1	Mort

		2	Viu
nacimp2	Lloc de naixement del progenitor 2	0	NS/NC
		1	Espanya
		2	Un altre país
pnacim2	País de naixement del progenitor 2	201 codis de tots els països + 1 codi NS/NC	
estudiosp2	Major nivell d'estudis assolit pel progenitor 2	0	NS/NC
		1	No sap llegir o escriure
		2	Sap llegir i escriure però va anar menys de 5 anys a l'escola
		3	Educació primària completa o va anar més de 5 anys a l'escola
		4	Primera etapa de l'educació secundària
		5	Segona etapa de l'educació secundària amb orientació general (Batxiller)
		6	Segona etapa de l'educació secundària amb orientació professional (FP)
		7	Educació postsecundària no superior
		8	Ensenyament de formació professional, arts plàstiques i disseny i esportistes de grau superior i equivalents
		9	Graus universitaris de fins a 240 crèdits, diplomats universitaris, títols propis d'universitat d'especialització i similars.
		10	Graus universitaris de més de 240 crèdits i llicenciats
		11	Màster

		12	Doctorat universitari
pnac1p2	País de nacionalitat 1 del progenitor 2	201 codis de tots els països + 1 codi NS/NC	
pnac2p2	País de nacionalitat 2 del progenitor 2	201 codis de tots els països + 1 codi NS/NC	
nucleo	Número d'ordre del nucli dins de l'habitatge que pertany cada NPV. Si valor = 0; no pertany a cap nucli	-	
tiponucleo	Tipus de nucli familiar	1	Matrimoni sense fills
		2	Matrimoni amb fills
		3	Parella no casada sense fills
		4	Parella no casada amb fills
		5	Pare només amb fills convivint
		6	Mare només amb fills convivint
sexo_pareja	Gènere dels membres de la parella	1	Parella de diferent gènere
		2	Parella del mateix gènere masculí
		3	Parella del mateix gènere femení
nacion_pareja	Nacionalitat dels membres de la parella	1	Ambdós espanyols
		2	Ambdós estrangers
		3	Espanyol amb parella estrangera
		4	Espanyola amb parella estrangera
union_pareja	Tipus d'unió de la parella	1	Parella casada
		2	Parella no casada, ambdós d'estat civil solters
		3	Altres tipus de parella no casada
estudios_pareja	Nivell d'estudis assolit pels membres de la parella	1	Cap membre de la parella té estudis superiors
		2	Un membre de la parella té est. sup. i l'altre no
		3	Ambdós membres de la parella tenen est. sup.

sitlab_pareja	Situació laboral dels membres de la parella	1	Ambdós membres de la parella tenen treball
		2	Un membre de la parella té feina i l'altre està jubilat i/o incapacitat
		3	Un membre de la parella té feina i l'altre no
		4	Ambdós membres de la parella estan jubilats i/o incapacitats
		5	Un membre de la parella no treballa i l'altre està jubilat i/o incapacitat
		6	Cap membre de la parella treballa
situhogar	Situació de la persona en l'habitatge	1	Persona que viu sola
		2	Pare/mare sol/a amb fills a l'habitatge
		3	Persona que viu en parella, s/s fills a l'habitatge
		4	Persona que viu en parella, amb fills a l'habitatge
		5	Fill/a
		6	Altres tipus de situacions
hijos_nucleo	Quantitat de fills que pertanyen al nucli familiar	-	
tipohogar	Tipus d'habitatge	1	Habitatge unipersonal
		2	Pare/mare sol/a amb fills que conviuen en l'habitatge
		3	Parella sense fills que conviuen en l'habitatge
		4	Parella amb fills que conviuen en l'habitatge
		5	Nucli familiar amb altres persones que no formen nucli familiar
		6	Persones que no formen cap tipus de nucli familiar entre si
		7	Dos o més nuclis familiars

nachohogar	Nacionalitat dels membres de l'habitatge	1	Habitatge exclusivament espanyol
		2	Habitatge mixte (amb espanyols i estrangers)
		3	Habitatge exclusivament estranger amb tots el seus membres de la mateixa nacionalitat
		4	Habitatge exclusivament estranger amb tots el seus membres de diferent nacionalitat
niveducahogar	Nivell d'estudis dels membres de l'habitatge	1	Cap membre de l'habitatge té estudis superiors
		2	Algun membre de l'habitatge té estudis superiors
		3	Tots els membres de l'habitatge tenen est. sup.
sitlabhogar	Situació laboral dels membres de l'habitatge	1	Cap membre de l'habitatge treballa ni està jubilat
		2	Tots els membres de l'habitatge estan jubilats
		3	Algun membre de l'habitatge està jubilat i algun altre no ho és ni té feina
		4	Tots els membres de l'habitatge tenen treball
		5	Algun membre de l'habitatge no té feina ni està jubilat i algun altre té feina
		6	Algun membre de l'habitatge té feina i algun altre està jubilat
		7	Algun membre de l'habitatge té feina, algun altre està jubilat i algun altre no treballa
ec	Estat civil legal	1	Solter/a
		2	Casat/da en primeres núpcies
		3	Casat/da en segones o més núpcies
		4	Vidu/a
		5	Separat/da
		6	Divorciat/da

estudios	Nivell d'estudis assolit	0	NS/NC
		1	No sap llegir o escriure
		2	Sap llegir i escriure però va anar menys de 5 anys a l'escola
		3	Educació primària completa o va anar més de 5 anys a l'escola
		4	Primera etapa de l'educació secundària
		5	Segona etapa de l'educació secundària amb orientació general (Batxiller)
		6	Segona etapa de l'educació secundària amb orientació professional (FP)
		7	Educació postsecundària no superior
		8	Ensenyament de formació professional, arts plàstiques i disseny i esportistes de grau superior i equivalents
		9	Graus universitaris de fins a 240 crèdits, diplomats universitaris, títols propis d'universitat d'especialitat i similars.
		10	Graus universitaris de més de 240 crèdits i llicenciats
		11	Màster
12	Doctorat universitari		
campo	Camp dels estudis	1	Educació
		2	Arts i Humanitats
		3	Ciències Socials
		4	Negocis i Dret
		5	Ciències

		6	Tecnologies de la informació i comunicació
		7	Arquitectura, Construcció, Formació Tècnica i Indústries
		8	Agricultura, Ramaderia, Pesca i Veterinària
		9	Salut i Serveis Socials
		10	Altres serveis
sitlab	Situació laboral principal durant la darrera setmana	1	Ocupat/da - A temps complet
		2	Ocupat/da - A temps parcial
		3	Estudiant
		4	Aturat/da - Ha treballat en anterioritat
		5	Aturat/da - No ha treballat en anterioritat
		6	Jubilat/da, prejubilat/da
		7	Incapacitat/da permanentment per a treballar
		8	Dedicat/da a les feines de l'habitatge
		9	Altres tipus d'inactivitat
edadsl	Edat des de que es troba en aquesta situació laboral	-	
flexi	Pot flexibilitzar, adaptar o acomodar la seva jornada laboral per conciliació	1	Sí
		6	No
lugtrab	On es troba el seu lloc de feina/estudis	1	En el propi domicili
		2	En varis municipis
		3	En el municipi que resideixo
		4	En un altre municipi de la mateixa província
		5	En un altre municipi d'una altra província
		6	En un altre país
hijos	Ha tingut fills	1	Sí
		6	No

nhijos	Número total de fills, visquin o no en el mateix habitatge, o estiguin vius o no en l'actualitat	-	
pareja	Té parella actualment	1	Sí
		6	No
aposoci	Té algú amb qui pugui parlar dels seus problemes quotidians/sentiments	1	Sí
		6	No
aposocihij	Tipus de parentesc amb les persones que li donen recolzament social: Fills/es	1	Sí
		6	No
aposocipama	Tipus de parentesc amb les persones que li donen recolzament social: Pares/Mares	1	Sí
		6	No
aposociher	Tipus de parentesc amb les persones que li donen recolzament social: Germans/es	1	Sí
		6	No
aposociotpar	Tipus de parentesc amb les persones que li donen recolzament social: Un altre parent	1	Sí
		6	No
aposociotro	Tipus de parentesc amb les persones que li donen recolzament social: Amics, Veïns, etcètera.	1	Sí
		6	No
lresmd	Lloc de residència de la persona que es troba a menor distància entre les persones que li brinden recolzament social	1	En el mateix municipi
		2	En un altre municipi de la mateixa província
		3	En una altra província de la mateixa CCAA
		4	En una altra CCAA dins d'Espanya
		5	En un altre país
nresi	Número de residents en l'habitatge	-	
		1	Pròpia per herència o donació
		2	Pròpia, per compra, totalment pagada

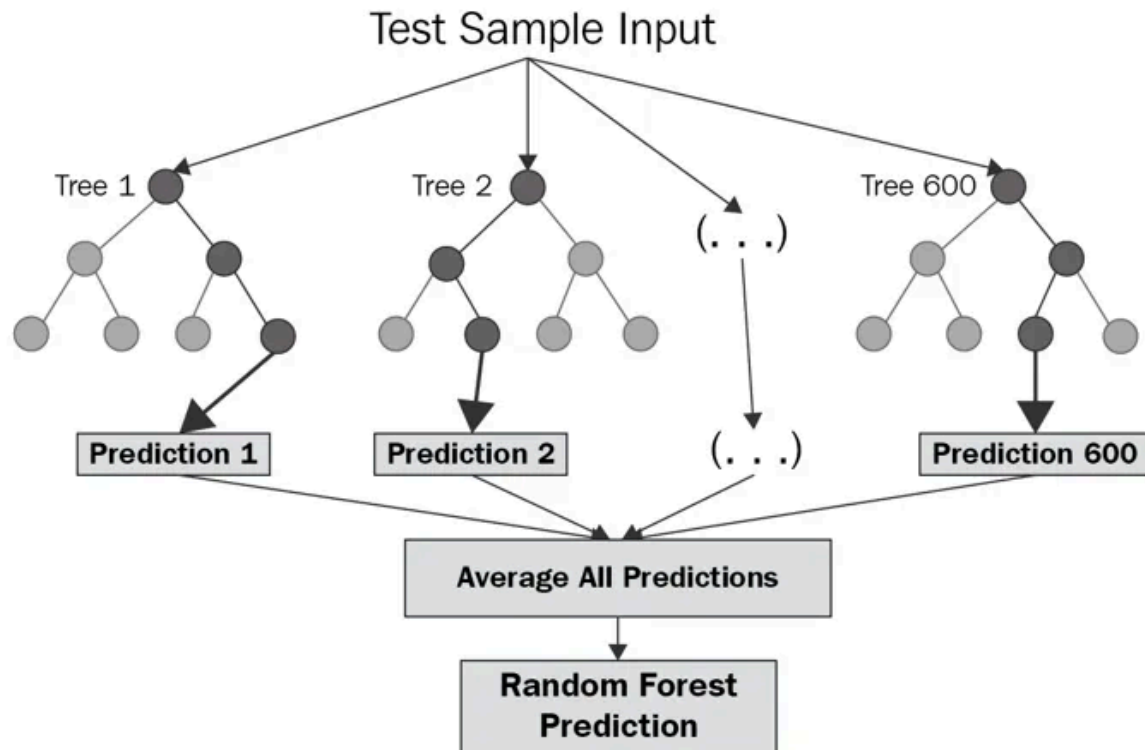
regvi	Règim de tinença de l'habitatge	3	Pròpia, per compra, amb pagaments pendents (hipoteca)
		4	Llogada
		5	Cedida gratuïtament o a baix cost (per un altra habitatge, pagada per l'empresa, etcètera.)
		6	Altra forma
pagovi	Quota mensual que paga d'hipoteca o lloguer intervalitzada	1	Menys de 100 €
		2	De 100 € a menys de 200 €
		3	De 200 € a menys de 300 €
		4	De 300 € a menys de 400 €
		5	De 400 € a menys de 500 €
		6	De 500 € a menys de 600 €
		7	De 600 € a menys de 700 €
		8	De 700 € a menys de 800 €
		9	De 800 € a menys de 900 €
		10	De 900 € a menys de 1.000 €
		11	De 1.000 € a menys de 1.250 €
		12	De 1.250 € a menys de 1.500 €
		13	De 1.500 € a menys de 2.000 €
		14	De 2.000 € a menys de 3.000 €
		15	De 3.000 € a menys de 5.000 €
		16	De 5.000 € o més
ingrehog	Ingressos mensuals nets de l'habitatge en intervals	1	Menys de 500 €
		2	De 500 € a menys de 1.000 €
		3	De 1.000 € a menys de 1.500 €
		4	De 1.500 € a menys de 2.000 €

		5	De 2.000 € a menys de 2.500 €
		6	De 2.500 € a menys de 3.000 €
		7	De 3.000 € a menys de 5.000 €
		8	De 5.000 € a més de 7.500 €
		9	7.500 € o més

Font: Elaboració pròpia a partir de la relació de les variables extretes de l'ECEPOV (INE, 2023).

8.3. Apèndix III

Figura X. Representació visual de la classificació i regressió de les dades mitjançant la metodologia *Random Forest*.



Font: Extret de <https://corporatefinanceinstitute.com/resources/data-science/random-forest/>

8.4. Apèndix IV

Figura X. Instal·lació dels prerequisits i llibreries necessàries per a dur terme el model predictiu. Jupyter Notebook - I.

```
[1]: !pip install scikit-learn
Requirement already satisfied: scikit-learn in c:\users\jordi\anaconda3\lib\site-packages (1.2.2)
Requirement already satisfied: numpy>=1.17.3 in c:\users\jordi\anaconda3\lib\site-packages (from scikit-learn) (1.26.4)
Requirement already satisfied: scipy>=1.3.2 in c:\users\jordi\anaconda3\lib\site-packages (from scikit-learn) (1.11.4)
Requirement already satisfied: joblib>=1.1.1 in c:\users\jordi\anaconda3\lib\site-packages (from scikit-learn) (1.2.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in c:\users\jordi\anaconda3\lib\site-packages (from scikit-learn) (2.2.0)

[2]: !pip install --upgrade openpyxl
Requirement already satisfied: openpyxl in c:\users\jordi\anaconda3\lib\site-packages (3.1.3)
Requirement already satisfied: et-xmlfile in c:\users\jordi\anaconda3\lib\site-packages (from openpyxl) (1.1.0)

[3]: !pip install shap
```

Font: Elaboració pròpia.

Figura X. Instal·lació dels prerequisits i llibreries necessàries per a dur terme el model predictiu. Jupyter Notebook - II.

```
[34]: from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.preprocessing import LabelEncoder
import numpy as np

import pandas as pd

import shap

import seaborn as sns
import matplotlib.pyplot as plt

[35]: pd.set_option('display.max_columns', None)
```

Font: Elaboració pròpia.

Figura X. Instal·lació de les funcions necessàries per a dur terme el model predictiu. Jupyter Notebook - III.

```
[37]: def seleccionar_columnas(dataset, columnas):  
    """  
    Esta función toma un DataFrame y una lista de nombres de columnas,  
    y devuelve un nuevo DataFrame que contiene solo las columnas especificadas.  
  
    Parámetros:  
    dataset (pd.DataFrame): El DataFrame original.  
    columnas (list): Lista de nombres de las columnas a seleccionar.  
  
    Retorna:  
    pd.DataFrame: Un nuevo DataFrame con las columnas especificadas.  
    """  
    # Verificar si todas las columnas están en el DataFrame original  
    for col in columnas:  
        if col not in dataset.columns:  
            raise ValueError(f"La columna '{col}' no existe en el DataFrame original.")  
  
    # Seleccionar las columnas especificadas  
    nuevo_dataset = dataset[columnas].copy()  
  
    return nuevo_dataset
```

Font: Elaboració pròpia.

Figura X. Instal·lació de les funcions necessàries per a dur terme el model predictiu. Jupyter Notebook - IV.

```
[38]: def balancear_dataset(dataset, columna_target):  
    """  
    Balancea un conjunto de datos binario dividiéndolo en dos clases según el valor de la columna objetivo y  
    muestreando aleatoriamente para igualar el tamaño de las clases.  
  
    Parámetros:  
    - dataset : DataFrame  
        El conjunto de datos que se desea balancear.  
    - columna_target : str  
        El nombre de la columna que se utilizará como objetivo para la clasificación binaria.  
  
    Retorna:  
    - DataFrame  
        Un nuevo DataFrame con el conjunto de datos balanceado.  
    """  
    # Separar las filas con 0s y 1s en diferentes DataFrames  
    df_clase_0 = dataset[dataset[columna_target] == 0]  
    df_clase_1 = dataset[dataset[columna_target] == 1]  
  
    # Obtener el tamaño de la clase menos frecuente  
    min_clase_size = min(len(df_clase_0), len(df_clase_1))  
  
    # Muestrear aleatoriamente de ambas clases para igualar el tamaño  
    df_clase_0_balanced = df_clase_0.sample(n=min_clase_size, replace=False, random_state=42)  
    df_clase_1_balanced = df_clase_1.sample(n=min_clase_size, replace=False, random_state=42)  
  
    # Concatenar Los DataFrames balanceados  
    df_balanced = pd.concat([df_clase_0_balanced, df_clase_1_balanced])  
  
    return df_balanced
```

Font: Elaboració pròpia.

Figura X. Instal·lació de les funcions necessàries per a dur terme el model predictiu. Jupyter Notebook - V.

```
[39]: def rellenar_nulos(df, columna, metodo):  
    """  
    Rellena los valores nulos de una columna en un DataFrame con la media o la mediana.  
  
    Parámetros:  
    df (pd.DataFrame): El DataFrame en el que se van a rellenar los valores nulos.  
    columna (str): El nombre de la columna en la que se van a rellenar los valores nulos.  
    metodo (str): El método a utilizar para rellenar los valores nulos ('media' o 'mediana').  
  
    Retorna:  
    pd.DataFrame: El DataFrame con los valores nulos de la columna especificada rellenados.  
    """  
    if metodo == 'media':  
        valor_relleno = df[columna].mean()  
    elif metodo == 'mediana':  
        valor_relleno = df[columna].median()  
    else:  
        raise ValueError("Método no reconocido. Use 'media' o 'mediana'.")  
  
    df[columna] = df[columna].fillna(valor_relleno)  
    return df
```

Font: Elaboració pròpia.

Figura X. Instal·lació de les funcions necessàries per a dur terme el model predictiu. Jupyter Notebook - VI.

```
[40]: def reemplazar_nulos_por_ceros(dataframe, column_name):  
    """  
    Reemplaza los valores NaN por 0 en la columna especificada de un DataFrame.  
  
    Parámetros:  
    - dataframe : DataFrame  
      El DataFrame en el que se realizará el reemplazo.  
    - column_name : str  
      El nombre de la columna en la que se reemplazarán los valores NaN por 0.  
  
    Retorna:  
    - DataFrame  
      El DataFrame modificado con los valores NaN reemplazados por 0 en la columna especificada.  
    """  
  
    dataframe[column_name] = dataframe[column_name].fillna(0)  
    return dataframe
```

Font: Elaboració pròpia.

Figura X. Importació de la BBDD original i balancejament de les dades (mateixa quantitat d'observacions) en funció de *motivocambio* de manera aleatòria. Jupyter Notebook - VII.

```
[48]: df = pd.read_excel('C:\\Users\\Jordi\\Documents\\BASEBDefinitive 1.xlsx')  
  
[94]: datos_balanceados = balancear_dataset(df, 'motivocambio')  
  
#Executar aquí si vull reiniciar el dataset
```

Font: Elaboració pròpia.

Figura X. Substitució dels valors NaN, per mitjana o valor zero, segons la tipologia de la variable. Jupyter Notebook - VIII.

```
[96]: datos_balancedados = rellenar_nulos(datos_balancedados, 'pagovi', metodo='media')
datos_balancedados = rellenar_nulos(datos_balancedados, 'nacimp1', metodo='media')
datos_balancedados = rellenar_nulos(datos_balancedados, 'estudiosp1', metodo='media')
datos_balancedados = rellenar_nulos(datos_balancedados, 'nacimp2', metodo='media')
datos_balancedados = rellenar_nulos(datos_balancedados, 'estudiosp2', metodo='media')
datos_balancedados = rellenar_nulos(datos_balancedados, 'edadestud', metodo='media')
datos_balancedados = rellenar_nulos(datos_balancedados, 'pnaci', metodo='media')

[97]: datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'estudios_pareja')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'sexo_pareja')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'nacion_pareja')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'union_pareja')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'sitlab_pareja')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'edads1')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'lugtrab')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'lresmd')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'tiponucleo')
datos_balancedados = reemplazar_nulos_por_ceros(datos_balancedados, 'campo')
```

Font: Elaboració pròpia.

Figura X. Maquetació del *train test split*⁴, fixació dels hiperparàmetres (regles de construcció) del *Random Forest* i execució d'aquest darrer⁵.

Jupyter Notebook - IX.

```
[315]: X_new = seleccionar_columnas(datos_balanceados, ['hijos', 'estudios', 'nresi', 'regvi', 'sitlab', 'pareja', 'pagovi',
          'edad', 'nacin', 'pnacin', 'ingrehog', 'ec', 'aposoci', 'sexo', 'campo', 'edads1',
          'estudiosp1', 'estudiosp2', 'tipohogar', 'sitlabhogar', 'niveducahogar'])
y_new = seleccionar_columnas(datos_balanceados, ['motivocambio'])
y_new = np.ravel(y_new)

[316]: X_train, X_test, y_train, y_test = train_test_split(X_new, y_new, test_size=0.20, random_state=42)

[317]: clf = RandomForestClassifier(n_estimators=100, min_samples_split=10, max_depth=10, random_state=42)

[318]: clf.fit(X_train, y_train)

[318]: * RandomForestClassifier
RandomForestClassifier(max_depth=10, min_samples_split=10, random_state=42)

[319]: # Hacer predicciones
y_pred = clf.predict(X_test)

# Evaluar el modelo
accuracy = accuracy_score(y_test, y_pred)
print(f"Precisión del modelo: {accuracy * 100:.2f}%")

Precisión del modelo: 91.20%
```

Font: Elaboració pròpia.

⁴ *Train test split*: 20% és el conjunt de dades d'entrenament seleccionades de manera aleatòria de la BBDD amb les dades ja balancejades.

⁵ Després d'haver provat múltiples combinacions de les 65 variables en el model, amb la inclusió i retirada d'aquestes, s'obté un amb el 91,20% de predicció.

Figura X. Procés per determinar la importància de les variables respecte l'estat d'emancipació (tant si sí com si no) integrades en el model.

Jupyter Notebook - X.

```
[320]: y_pred_proba = clf.predict_proba(X_test)

[321]: importancias = clf.feature_importances_

[322]: importancias_df = pd.DataFrame({'feature': X_train.columns, 'importance': importancias})
importancias_df = importancias_df.sort_values(by='importance', ascending=False)

[323]: importancias_df

#executar fins aquí
```

Font: Elaboració pròpia.

Figura X. Obtenció de la taula d'importàncies⁶ de les variables respecte l'estat d'emancipació (tant si sí com si no) integrades en el model.

Jupyter Notebook - XI.

	feature	importance			
2	nresi	0.247199	8	nacim	0.018207
18	tipohogar	0.132947	10	ingrehog	0.017893
3	regvi	0.121618	11	ec	0.017695
19	sitlabhogar	0.091012	17	estudiosp2	0.015682
7	edad	0.076936	16	estudiosp1	0.015549
0	hijos	0.053465	20	niveducahogar	0.015407
5	pareja	0.034830	1	estudios	0.012783
15	edadsl	0.031368	14	campo	0.011393
6	pagovi	0.030587	13	sexo	0.002771
4	sitlab	0.030163	12	aposoci	0.002526
9	pnacim	0.019969			

Font: Elaboració pròpia.

⁶ Ordenades de major a menor (* 100 %) - primera aproximació d'aquelles variables que determinaran l'estat d'emancipació *true* i *false*.

8.5. Apèndix V

Figura X. Maquetació i execució del mètode SHAP. Jupyter Notebook - XII.

```
[324]: explainer = shap.TreeExplainer(clf)
[325]: explainer
[325]: <shap.explainers._tree.TreeExplainer at 0x26eed1688d0>
[326]: shap_values = explainer(X_test)
[327]: X_test
[327]:
```

	hijos	estudios	nresi	regvi	sitlab	pareja	pagovi	edad	nacim	pnacim	ingrehog	ec	aposoci	sexo	campo	edadsl	estudiosp1	estudiosp2	tipohogar
92	6	8	1	3	1	6	5.000000	29	1	108	5	1	1	1	7.0	25.0	6.0	5.0	1
3343	6	11	1	3	1	6	5.000000	29	1	108	4	1	1	2	1.0	24.0	5.0	5.0	1
6602	6	5	5	2	3	1	5.041885	22	1	108	5	1	1	1	5.0	0.0	5.0	5.0	4
11965	6	6	2	4	4	1	4.000000	29	1	108	3	1	1	1	7.0	0.0	4.0	4.0	3
4100	1	4	5	4	1	1	1.000000	28	1	108	3	1	6	1	0.0	18.0	4.0	4.0	4
...
18950	6	11	4	3	3	1	4.000000	23	1	108	5	1	1	2	2.0	0.0	4.0	8.0	4
12472	6	8	1	2	1	6	5.041885	26	1	108	3	1	1	1	8.0	26.0	6.0	6.0	1
6095	1	4	3	3	4	1	3.000000	28	1	108	4	1	1	1	0.0	0.0	3.0	4.0	4
19728	6	6	4	6	1	1	5.041885	19	1	108	4	1	1	1	6.0	19.0	7.0	5.0	4
16436	6	5	4	3	1	6	8.000000	23	1	108	4	1	1	1	5.0	23.0	4.0	4.0	4

500 rows × 21 columns

Font: Elaboració pròpia.

Figura X. Creació de les noves columnes/variables en la BBDD: importància % de cadascuna de les variables en la predicció d'emancipació *true* (=motivocambio), probabilitat de 'motivocambio' i predicció de 'motivocambio' per individu. Jupyter Notebook - XIII.

```
[329]: lista_nombres = X_test.columns

[330]: lista_nombres = [nombre + '_shap' for nombre in lista_nombres]

[331]: lista_nombres

[331]: ['hijos_shap',
'estudios_shap',
'nresi_shap',
'regvi_shap',
'sitlab_shap',
'pareja_shap',
'pagovi_shap',
'edad_shap',
'nacim_shap',
'pnacim_shap',
'ingrehog_shap',
'ec_shap',
'aposoci_shap',
'sexo_shap',
'campo_shap',
'edads1_shap',
'estudiosp1_shap',
'estudiosp2_shap',
'tipohogar_shap',
'sitlabhogar_shap',
'niveducahogar_shap']

[332]: shap_df = pd.DataFrame(shap_values[:, :, 1].values, columns = lista_nombres)

[334]: df1_reset = X_test.reset_index(drop=True)
df2_reset = shap_df.reset_index(drop=True)
df3 = pd.DataFrame(y_test, columns=['motivocambio'])
df4 = pd.DataFrame(y_pred, columns=['motivocambio_pred'])
df5 = pd.DataFrame(y_pred_proba[:, 1], columns=['prob_motivocambio'])

result = pd.concat([df3, df5, df4, df1_reset, df2_reset], axis=1)
```

Font: Elaboració pròpia.

Figura X. Exportació de la BBDD final. Jupyter Notebook - XIV.

```
[337]: ruta_archivo_csv = 'C:\\Users\\Jordi\\Documents\\archivo_final2.csv'

# Exporta el DataFrame a un archivo CSV
result.to_csv(ruta_archivo_csv, sep=';', index=False, decimal=',')
```

Font: Elaboració pròpia.