

Grau en Estadística

Títol: Anàlisi dels riscos psicosocials al treball i el seu efecte sobre la salut

Autor: Aniol Garriga Reixach

Director: David Moríña Soler

Departament: Economia, Estadística i Economia Aplicada

Convocatòria: Gener 2022



RESUM

Aquest és un estudi de la segona enquesta catalana de les condicions de treball (II ECCT) realitzada l'any 2010. L'enquesta es basa en el mètode COPSOQ desenvolupat a Dinamarca l'any 2000.

En el treball, es busca, per una banda, confirmar que les dades observades s'ajusten al model de dimensions teoritzat, per fer-ho s'ha utilitzat una anàlisi factorial confirmatòria (CFA). Amb l'ajuda d'aquest mètode s'ha pogut confirmar el model teoritzat a partir de les dades observades. També ha permès obviar una qüestió que no aportava nova informació, ja que les respostes eren molt similars a l'anterior.

D'altra banda, amb aquest estudi es vol brindar una eina que ajudi a predir quan un individu pot estar en risc de patir una malaltia mental en funció al que hagi respost. Per fer-ho s'ha aplicat a les dades l'algoritme "Random forest" que prediu quan un individu està en risc de patir una malaltia de salut mental pràcticament en un 75% dels casos.

Paraules clau:

Random forest, Arbres de decisió/regressió, bagging, CFA, riscos psicosocials, COPSOQ, Salut mental.

ABSTRACT

This is a study of the second Catalan survey on working conditions (II ECCT) carried out in 2010. The survey is based on the COPSOQ method developed in Denmark in 2000.

The work seeks to confirm that the observed data conform to the theorized model of dimensions, using a confirmatory factor analysis (CFA) to do so. Using this method, it has been possible to confirm the theorized model from the observed data. It also made it possible to ignore a question that did not provide new information, as the answers were very similar to the previous one.

On the other hand, this study aims to provide a tool to help predict when an individual may be at risk for mental illness depending on how has responded. The random forest was the chosen algorithm because apart from predict it shows which variables helps to validate the prediction. The results are that almost 75% of cases are well predicted.

Key words:

Random Forest, decision/regression trees, bagging, CFA, psychosocial risks, COPSOQ, mental health

CLASSIFICACIÓ AMS

- 62H25 Factor analysis and principal components; correspondence analysis
- 62P25 Applications of statistics to social sciences
- 62H30 Classification and discrimination; cluster analysis (statistical aspects)
- 91C20 Clustering in the social and behavioral sciences

Índex

1.	Introducció	4
1.1	Objectius	5
1.2	Agraïments	5
2.	Mètodes	6
2.1	Qüestionari	6
2.2	Variables	8
2.3	Anàlisi factorial confirmatori	10
2.4	Model predictiu	11
2.5	Software empleat	12
3.	Resultats	13
3.1	Descriptiva bivariant:	13
3.2	Estudi del qüestionari	17
3.2.1	Relació entre les qüestions	17
3.2.2	Anàlisi factorial confirmatori.....	18
3.3	Model predictiu	22
3.3.1	Salut Mental	22
3.3.2	Salut auto percebuda	23
3.3.3	Percepció de l'afecte de la feina sobre la salut.....	25
4.	Conclusions.....	27
5.	Bibliografia	29
6.	Annex.....	30

1. Introducció

Actualment el motor econòmic de Catalunya consta de més de 600 mil empreses que donen feina a aproximadament 3,5 milions de persones. S'ha avaluat en diverses ocasions la competitivitat d'aquestes empreses però no és tant comú estudiar com són les condicions d'aquestes feines i com afecten a la salut de les persones. (*Idescat. Demografia i Societat, n.d.*)

Aquest fet, tot i que no és un problema que hagi aparegut en aquests moments, s'ha vist reflectit de forma clara en aquests últims anys on, per desgràcia i degut a la pandèmia que patim, moltes persones els han modificat les seves condicions laborals. Per aquest motiu és de vital importància tenir en compte les condicions laborals de les persones perquè afecten a la salut mental de tothom.

Per estudiar-ho i entendre-ho cal saber que s'entén per Salut. El concepte de salut és més ampli del que sembla a primera vista; durant la constitució de la OMS l'any 1946 a la Conferència Internacional de Salut celebrada a Nova York es definia la salut de l'individu com "*Un estat de benestar físic, mental i social, i no solament l'absència d'afeccions i malalties*". (*World Health Organization. Health Definition, n.d.*)

Per actuar en conseqüència l'any 2000 es va crear la metodologia COPSQQ per a la investigació, la avaluació i la prevenció dels riscos psicosocials mitjançant una enquesta. La primera versió va ser creada per el "National Research Center for the Working Environment" de Dinamarca. Aquesta fou traduïda i adaptada a la nostra comunitat per primer cop l'any 2005 sota el nom de COPSQQ-Istas. Amb aquesta metodologia s'ha estat enquestant una mostra de la població activa de Catalunya cada cinc anys. (Nübling M, 2013)

Els riscos psicosocials als que està exposada la població activa estan descrits per les següents cinc dimensions; Doble Presència (DP), que s'entén com la necessitat de pensar en els problemes de la llar mentre s'està treballant; Exigències psicològiques a la feina (EPF), que compren diversos tipus d'exigència que pot tenir una feina; Control sobre la feina (CSF), és a dir el control que te una persona sobre la feina que realitza; Suport Social i Qualitat del lideratge (SSQL), com el propi nom indica aquesta dimensió contempla la relació de l'individu amb el seu entorn personal a la feina; finalment, Compensacions per la feina (CF), que també s'enten pel nom de la dimensió que en aquest cas s'explica quan la persona enquestada se sent valorada i compensada pel treball realitzat. Cadascuna d'elles esta formada per diferents categories que estan representades al qüestionari per tres o menys preguntes. (Instituto Sindical del Trabajo Ambiente y Salud, 2010)

Recentment, en la recerca de salut ocupacions, i mitjançant aquestes enquestes, s' ha demostrat que l'exposició a factors de risc comporta problemes de salut física com malalties cardiovasculars o problemes musculoesquelètic. També es constaten problemes de salut mental com absentisme. Per desgràcia, estudis mes concrets apunten la DP com una causa d'estrés i falta de vitalitat o que una alta exposició a les EPF ajuden a patir depressió. (Utzet et al., 2021)

1.1 Objectius

Primerament es busca corroborar que les dimensions estan construïdes segons el model a priori, alhora que es busca reduir el nombre de preguntes ja que en enquestar un individu es tarda pràcticament mitja hora concretament 27,4 minuts de mitjana. Seguidament es vol proporcionar una eina amb la que a partir d'aquesta enquesta es pugui detectar quan un individu esta exposat a un tipus concret de risc social i poder predir si aquest te un efecte negatiu a la seva salut mental.

1.2 Agraïments

Primerament, és oportú agrair el tutor d'aquest projecte, David Moríña Soler per tota l'ajuda prestada durant tota la realització del treball. També es precís agrair a la Marta Franch Pujol per la correcció gramatical ortogràfica i de forma en la redacció d'alguns apartats.

2. Mètodes

L'enquesta es va aplicar a 3601 individus. Tot i així, per tal d'evitar biaix s'han exclòs de l'anàlisi tots aquells individus que treballen de forma autònoma o ho fan sense contracte. Així que la mostra amb la que treballa aquest estudi és de 2566 persones.

2.1 Qüestionari

El qüestionari en el que es basa aquest treball esta organitzat de la següent manera; la primera part descriu el perfil sociodemogràfic de l'individu. Seguidament es pregunta sobre el tipus de feina i les condicions que aquesta comporta. Després s'entra a la part de les condicions físiques de la feina amb les preguntes 28 29 i 30. Les dimensions dels riscos psicosocials descrites anteriorment estan distribuïdes de la següent forma entre les preguntes 31 i 35 amb l'excepció de la doble presència que debat amb les preguntes 64 i 65;

Subdimensio	Pregunta	Opcions de resposta
Doble presència		
Doble Presència (DP)	Quina part del treball familiar i domèstic fa vostè? (P64)	B
	Si falta algun dia de casa, les feines domèstiques que realitza es queden sense fer? (P65A)	A
	Quan és a l'empresa, pensa en les feines domèstiques i familiars? (P65B)	A
	Hi ha moments que necessitaria ser a l'empresa i a casa alhora? (P65C)	A
Exigències psicològiques a la feina		
Exigències psicològiques quantitatives (EPF)	Ha de treballar molt de pressa (P31A)	A
	La distribució de tasques és irregular, cosa que provoca que se li acumuli la feina (P31B)	A
	Té temps de portar la feina al dia (P31C)	A invertida
Exigències emocionals (EE)	Li costa oblidar els problemes de la feina (P31D)	A
	La seva feina, en general, desgasta emocionalment (P31E)	A
Exigència d'amagar	La seva feina requereix que amagui les seves emocions (P31F)	A

emocions (EAE)		
Control sobre la feina		
Influència (In)	Té influència sobre la quantitat de treball que se li assigna (P32A)	A
	Es té en compte la seva opinió quan se li assignen les tasques (P32B)	A
	Té influència sobre l'ordre en què realitza les feines (P32C)	A
Possibilitats de desenvolupament al treball (PDT)	La seva feina requereix que tingui iniciativa (P32F)	A
	La seva feina permet que aprengui coses noves (P32G)	A
Control del temps (CdT)	Pot decidir quan fa un descans (P32D)	A
	Si té algun assumpte personal o familiar, pot deixar el seu lloc de treball almenys una hora, sense haver de demanar un permís especial (P32E)	A
Sentit de la feina (SF)	Se sent compromès amb la seva professió (P32H)	A
	Tenen sentit les seves feines (P32I)	A
Integració en l'empresa (IE)	Parla amb entusiasme de la seva empresa a altres persones (P32J)	A
Suport social i qualitat del lideratge		
Previsibilitat (Pre)	A la seva empresa, se li informa amb prou antelació dels canvis que poden afectar el seu futur? (P34C)	A
	Rep tota la informació que necessita per fer bé la seva feina? (P34D)	A
Claredat de rol (CR)	Sap exactament quin marge d'autonomia té a la seva feina? (P34A)	A
	Sap exactament quines tasques són responsabilitat seva? (P34B)	A
Qualitat del lideratge (QL)	Els seus/les seves caps immediats planifiquen bé la feina? (P34I)	A
	Els seus/les seves caps immediats es comuniquen bé amb els treballadors i treballadores? (P34J)	A
Suport social a la feina (SSF)	Rep ajut i suport de les seves companyes o companys? (P34E)	A
	Rep ajut i suport del seu superior immediat/a? (P34F)	A
Possibilitats de relació social (PRS)	El seu lloc de treball es troba aïllat dels seus companys/es? (P34G)	A invertida
Sentiment de Grup (SG)	A la feina, sent que forma part d'un grup? (P34H)	A
Compensacions a la feina		
Inseguretat a la feina (IF)	Com seria de difícil trobar una altra feina en el cas que es quedés a l'atur (P33A)	A

	Si li canvien les tasques contra la seva voluntat (P33B)	A
	Si li canvien l'horari (torn, dies de la setmana, hores d'entrada i sortida) contra la seva voluntat (P33C)	A
	Si li varien el salari (que no l'actualitzin, que l'abaixin, que introdueixin el salari variable, que paguin en espècie, etc.) (P33D)	A
Estima (Est)	Els seus superiors li donen el reconeixement que mereix? (P35A)	A
	En les situacions difícils a la feina rep el suport necessari? (P35B)	A
	A la feina el tracten injustament. (P35C)	A invertida
	Si pensa en tota la feina i l'esforç que ha realitzat, el reconeixement que rep a la seva feina li sembla adequat? (P35D)	A
<i>Taula 1. Qüestions que construeixen les dimensions de riscos psicosocials.</i>		

L'opció de resposta A codifica amb els valors de 0 a 4 les següents respostes: Sempre, Moltes vegades, De vegades, Només alguna vegada, Mai.

L'opció de resposta A invertida codifica amb els mateixos valors que l'anterior però a la inversa.

L'opció de resposta B també pren valors de 0 a 4 per les següents respostes: Soc el/la responsable principal i faig la major part de les feines familiars i domèstiques, Faig aproximadament la meitat de les feines familiars i domèstiques, Faig més o menys una quarta part de les feines familiars i domèstiques, Només faig feines puntuals, No faig cap d'aquestes feines o gairebé cap. (Moncada et al., 2005)

2.2 Variables

Podem parlar de tres tipus de variables; les sociodemogràfiques, que descriuen aspectes bàsics de l'individu; les dimensions de riscos psicosocials de les quals s'ha parlat ja abans en aquest estudi amb l'afegit de les condicions ambientals construïda a partir de les preguntes 28, 29 i 30; finalment les variables resposta que són tres per una banda la salut mental de l'individu construïda amb la pregunta 44, la salut auto percebuda del pacient que és la resposta a la pregunta 41 i finalment la resposta a la pregunta 42A que li demana a la persona enquestada si sent que la feina afecta directament a la seva salut.

Les variables sociodemogràfiques estan construïdes de la mateixa manera que presenta el qüestionari hi ha més informació sobre les seves distribucions i valors a l'annex. Aquestes són el sexe de l'individu, la seva edat, la nacionalitat, el lloc de

procedència, el sector laboral on treballa, el tipus de contracte que té (si és o no temporal), si treballa a temps complert, i finalment el salari que rep.

Les variables utilitzades per representar les condicions laborals de cada individu són les cinc dimensions ja esmentades de riscos psicosocials sempre que els anàlisis aplicats en aquest estudi ho han permès s'ha emprat la variable ja inclosa al set de dades.

La variable dicotòmica referent a cada dimensió, és a dir la variable que indica quan un individu està exposat a una certa dimensió de risc, es calcula la suma de totes les qüestions incloses en la pregunta o preguntes que defineixen aquesta dimensió. Seguidament en funció de l'ordre de les respostes s'escull el valor que cau en el terç amb puntuacions més altes o més baixes com discriminant per saber quan un individu esta o no exposat a aquesta dimensió.

Per exemple la pregunta 34 que conté deu qüestions i descriu la dimensió SSIQL, per com estan formulades les preguntes, la resposta "Sempre" que proporciona un valor 1 indica un bon nivell de Suport social o qualitat de lideratge. En canvi, amb la pregunta 31 que conte sis qüestions i descriu la dimensió EPF, la formulació de les preguntes dona entendre que "Sempre" que també val 1, implica una alta exposició a Exigències Psicològiques a la feina.

De la mateixa forma esta creada la variable de Condicions ambientals que esta formada amb les preguntes 28, 29 i 30 amb 12, 12 i 11 qüestions respectivament. En aquestes preguntes se li demana a l'individu enquestat sobre com el propi lloc de treball.

Finalment les variables resposta que parlen concretament respecte **la salut** i que son tres per una banda la *salut mental de l'individu* construïda amb la pregunta 44, la *salut auto percebuda del pacient* que és la resposta a la pregunta 41 i finalment la resposta a la pregunta 42A que li demana a la persona enquestada si sent que la *feina afecta directament a la seva salut*.

Aquestes són les utilitzades com a resposta en el model predictiu. Dues d'elles, la qüestió 42A i la única qüestió de la pregunta 41 s'han empleat, tal com estaven codificades al qüestionari utilitzat. La restant, la 44 s'ha codificat de la mateixa manera que les variables explicatives referents a les condicions laborals. S'han sumat les respostes de les 12 qüestions incloses en la pregunta i s'ha considerat que les persones que sumen mes de 23 punts en totes les preguntes estan en risc de patir algun tipus de malaltia mental.

2.3 Anàlisi factorial confirmatori

Com a part de l'anàlisi descriptiu de les dades i per complir amb un dels objectius d'aquest estudi, s'ha estudiat la relació entre les preguntes que constitueixen les dimensions de riscos psicosocials explicades anteriorment. Per fer-ho s'ha aplicat un anàlisi factorial confirmatori (CFA).

Aquest és un mètode d'anàlisi de dades quantitatives que pertany a la família de tècniques de modelat d'equacions estructurals (SEM). Aquest és un mètode que permet avaluar l'ajustament de les dades observades en comparació amb un model conceptualitzat *a priori*. Així doncs aquest mètode s'ha de veure com una tècnica de refutació que facilita el rebuig estadístic o en el millor dels casos una retenció molt temptativa de una teoria específica respecte al factor o factors responsables de les relacions observades en les dades. (Alhija, 2010)

El model que avalua aquest mètode és un sistema d'equacions com el que es presenta a continuació on X_i és la i -essima variable observada, ξ_j el j -essim factor latent, δ_i és l' i -essim residu, i $\lambda_{i,j}$ el pes que té la variable i sobre el factor latent ξ_j i només pren valor diferent de zero quan la variable està associada al factor latent;

$$\begin{bmatrix} X_1 \\ \vdots \\ X_{k_i} \\ \vdots \\ X_i \end{bmatrix} = \begin{bmatrix} \lambda_{1,1} & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & \lambda_{k_i,k_j} & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & \lambda_{i,j} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_j \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_{k_i} \\ \vdots \\ \delta_i \end{bmatrix}$$

Que es pot representar com, $X = \Delta\xi + \delta$ on X representa el vector vertical de les variables observades, Δ la matriu dels pesos, ξ el vector de factors latents i δ el vector dels residus.

La implicació de les equacions estructurals que acompanyen el model es que la matriu poblacional de variàncies i covariàncies per les variables X , Σ , és una funció dels pesos λ_{ij} en la matriu Δ , les variàncies i covariàncies entre els factors latents a la matriu Φ i les variàncies i covariàncies en la matriu Θ_δ . Més específicament, si tots els paràmetres del model estan continguts en un vector d'una sola columna *theta* la matriu poblacional de variàncies i covariàncies de les variables observades s'explica pel model $\Sigma(\theta)$ donat per;

$$\Sigma(\theta) = \Delta\Phi\Delta' + \Theta_\delta$$

L'objectiu doncs del CFA és comprovar la discrepància entre el la funció de paràmetre estimat $\Sigma(\hat{\theta})$ i la matriu de variàncies i covariàncies poblacional S a partir del contrast de l'estadístic amb la χ^2 . (J. Boissevain, n.d.)

2.4 Model predictiu

Per tal de complir amb el segon objectiu de l'estudi s'empra un mètode predictiu basat en la classificació de les dades. En aquest cas l'algoritme escollit és el "*Random forest*" ja que permet saber quin efecte té cada variable sobre l'exactitud de la predicció. Aquest és un algoritme d'aprenentatge supervisat que construeix un "bosc" d'arbres de decisió o de regressió entrenats amb el mètode "bagging".

Abans de definir l'algoritme utilitzat cal fer una breu introducció als arbres de regressió o arbres de decisió que son en essència el mateix però es calculen per diferents mètodes en funció de la variable resposta. Els primers, els de regressió son emprats quan la variable resposta és continua mentre els de regressió s'usen per una variable resposta categòrica, que és el cas d'aquest estudi:

En un arbre de decisió hi ha tres tipus de node, el primer des del qual surt l'arbre s'anomena "*Root node*", el següent tipus són els nodes de decisió i finalment els nodes terminals o "*leaf nodes*". Amb tot el set de dades a l'arrel es busca l'atribut que millor descriu la resposta. S'entenen les divisions resultants com a nous nodes arrels des dels quals es torna a buscar una altra millor variable que defineixi millor la resposta fins a arribar a un node terminal on tots els individus formin part d'una mateixa categoria o no hi hagi una variable que sigui capaç de categoritzar millor que l'anterior. (StatQuest with Josh Starner, 2018)

Hi ha diversos criteris per escollir la variable que millor divideix les dades com l'error de classificació, l'entropia, o l'índex de gini, entre altres. Aquest últim mesura el grau de puresa d'un node mesurant la probabilitat que en un node només hi hagi individus d'una sola categoria de la resposta:

$$Gini(t) = 1 - \sum_{i=1}^n (p_i)^2 = 1$$

On p_i és la probabilitat que un individu sigui de la classe i . (StatQuest with Josh Starner, 2018)

Els arbres de decisió son fàcils de crear i fàcils d'interpretar però la tècnica esta subjecte a sobre ajustament i en conseqüència comporten un nivell elevat de biaix així que no prediuen amb exactitud noves observacions. És per això que l'algoritme escollit és el random forest que com s'ha esmentat anteriorment crea un bosc a partir d'arbres de decisió aplicant la tècnica "*bagging*". (StatQuest with Josh Starner, 2018)

El nom de "*bagging*" és un acrònim de les paraules angleses bootstrap i aggregation, és a dir que es s'escull un nombre determinat de mostres on no importa

la repetició d'observacions. Amb cadascuna d'aquestes mostres es crea un arbre de decisió però no s'aplica exactament l'algoritme explicat anteriorment ja que això seria computacionalment inviable. La variable a partir de la qual es crea un node és escollida aleatòriament. Sempre que la capacitat de l'equip ho permeti es poden incloure diverses categories sobre les quals construir el node i s'escull d'aquestes la millor de la mateixa manera que amb l'arbre de decisió. (StatQuest with Josh Starner, 2018)

Un cop construït el bosc per tal de predir una nova observació per cadascun dels arbres anteriorment creats s'observa quina categoria té la variable resposta al node terminal on acaba la nova observació. Es prediu la resposta d'aquesta nova variable segons quina categoria hagi aparegut més cops. (StatQuest with Josh Starner, 2018)

La principal raó per escollir aquest algoritme és el fet que permet determinar la importància de les variables en la predicció. Aquesta es calcula de dues formes, la primera calcula el decreixement mitjà en l'omissió d'una variable; més concretament, es comença calculant la fiabilitat de la predicció de cada arbre amb les dades que on s'han usat per construir-lo. Seguidament, els valors de la variable dins la mostra no usada per l'arbre en concret es barregen aleatòriament mantenint l'ordre en la resta de variables. Finalment, es mesura el decreixement de la predicció. El valor obtingut és la mitjana d'aquests decreixements en tots els arbres. (*How Is Variable Importance Calculated for a Random Forest?*, n.d.)

L'altra forma de calcular la importància d'una variable és a partir de l'índex de Gini. En aquest cas es calcula una mitjana del sumatori de disminució de l'índex cada cop que la variable en qüestió ha estat escollida per dividir un node. Aquest sumatori és dividit pel nombre total d'arbres. Així doncs l'escala no és rellevant sinó que ho són els valors relatius. (*How Is Variable Importance Calculated for a Random Forest?*, n.d.)

2.5 Software empleat

Per tal de realitzar aquest estudi s'ha utilitzat el llenguatge R en la interfície RStudio. Per la anàlisi del model de dimensions s'ha utilitzat el paquet "lavaan" (Rosseel, 2014) i les seves versions. Per l'altre objectiu, el model predictiu, el paquet que més s'ha utilitzat és el "randomForest" (*Using_random_forests_V3.1*, n.d.). La resta de paquets es poden consultar al principi del codi adjunt a l'annex.

3. Resultats

Dels 2566 individus enquestats un 86,06% treballen amb contracte fix, permanent o indefinit (treballant a temps complert un 72,72% i a temps parcial un 13,34%). Són dones un 46,38% de les enquestades, l'edat mitjana de la mostra és de $39,22 \pm 10,81$ anys, amb una representació majoritària de persones amb l'edat compresa entre 25 i 45 anys. Pel que fa la distribució geogràfica de la mostra, la majoria d'individus viu a l'àrea metropolitana de Barcelona, concretament un 40,53% i la resta estan distribuïts en sis zones diferents de la comunitat com es pot veure en la següent taula.

Pel que fa a les respostes, veiem com la gran majoria de les persones es considera en bon estat de salut, concretament un 90,45%. Però de tota aquesta gent fins a un 28,02% esta en risc de patir una malaltia mental. En canvi del 9,55% restant que considera tenir una salut regular o fins i tot dolenta fins a 23,65% no pensa que la feina tingui una influència en aquest estat de salut. Finalment cal destacar que de totes les persones en risc de patir una malaltia mental més de la meitat (en concret un 51,19%) considera que les condicions laborals no tenen efecte sobre la seva salut.

3.1 Descriptiva bivariant:

Per poder tenir una idea general de les dades amb les que s'ha treballat, seguidament es mostra un recull de la relació entre les variables explicatives i les respostes.

	Àmbit Metropolità	Com. gironines	Camp de Tarragona	Com. centrals	Terres de l'Ebre	Àmbit de Ponent	Alt Pirineu i Aran
Exigències Psicològiques	29.55%	22.60%	25.75%	22.84%	17.70%	30.69%	28.69%
Doble presència	36.12%	29.90%	33.00%	37.93%	23.36%	23.83%	34.43%
Control del Treball	29.97%	24.91%	30.18%	40.71%	30.64%	36.53%	25.44%
Suport Social	16.39%	20.56%	8.83%	24.36%	27.83%	18.70%	5.36%
Estima	28.69%	27.49%	18.03%	35.82%	25.42%	22.51%	29.41%
Salut Mental	33.14%	34.59%	39.46%	16.84%	6.97%	52.71%	10.66%
Salut auto percebuda	11.92%	6.85%	4.68%	7.90%	6.97%	10.47%	13.93%
Afectació de la feina en la Salut	47.29%	40.41%	38.51%	36.55%	22.13%	41.52%	40.50%
Total	40.53%	11.38%	11.69%	11.34%	9.51%	10.80%	4.75%

Taula 2. Relació de la variable Residència amb els riscos psicosocials i les respostes

En aquesta primera taula es veu quin percentatge de la mostra resident en les set àrees de Catalunya està exposada als diferents riscos en les cinc primeres files. Les tres següents files expliquen quin percentatge de cada àrea es considera que té una salut mental pobra a partir de les respostes al qüestionari; percep una que té una salut regular o dolenta, o considera que la seva feina té un afecte nociu cap a la seva pròpia persona.

És interessant com sembla que les persones residents a les comarques centrals són les més exposades a la majoria dels riscos i de les menys afectades en la variable resposta.

La següents taules mostren les mateixes files que l'anterior, però en aquest cas s'estratifica pel gènere i l'edat de la persona enquestada, respectivament.

	Homes	Dones
Exigències Psicològiques	24.31%	29.06%
Doble presencia	14.95%	53.09%
Control del Treball	30.30%	32.25%
Suport Social	16.70%	18.94%
Estima	26.32%	28.10%
Salut Mental	28.66%	33.11%
Salut auto percebuda	7.27%	12.10%
Afectació de la feina en la Salut	39.24%	42.83%
Total	53.62%	46.38%

Taula 3. Relació de la variable Sexe amb els riscos psicosocials i les respostes

	De 16 a 24	De 25 a 34	De 35 a 44	De 45 a 54	Mes de 54
Exigències Psicològiques	22.11%	28.35%	26.72%	27.12%	22.74%
Doble presencia	16.42%	30.09%	35.63%	39.02%	30.43%
Control del Treball	42.62%	31.02%	30.72%	31.46%	24.81%
Suport Social	16.15%	17.70%	17.98%	18.15%	17.51%
Estima	26.02%	28.55%	27.67%	26.59%	23.70%
Salut Mental	24.87%	28.68%	29.22%	36.26%	33.82%
Salut auto percebuda	4.48%	6.56%	7.42%	14.98%	16.25%
Afectació de la feina en la Salut	29.15%	40.13%	41.73%	44.38%	42.39%
Total	7.83%	30.32%	29.46%	21.59%	10.80%

Taula 4. Relació de la variable Edat amb els riscos psicosocials i les respostes

Com s'ha vist en molts altres estudis, per raons probablement culturals, les dones estan molt més exposades a la doble presencia que els homes. En la resta de dimensions també es denota més exposició per les dones però sense ser tant alarmant.

En la segona taula, veiem com les persones d'una edat més avançada, tenen més control sobre la seva feina o com la doble presencia afecta sobretot a les persones de mitjana edat.

Seguidament es mostra la taula estratificada pel lloc de procedència de les persones enquestades.

	Espanyola	Europea	Fora d'Europa
Exigències Psicològiques	27.26%	27.55%	21.52%
Doble presencia	34.21%	32.65%	21.78%
Control del Treball	27.97%	41.49%	50.00%
Suport Social	17.21%	15.62%	22.41%
Estima	26.98%	25.00%	29.02%
Salut Mental	30.51%	38.14%	29.79%
Salut auto percebuda	10.13%	6.12%	6.69%
Afectació de la feina en la Salut	42.20%	28.57%	36.28%
Total	83.29%	3.83%	12.89%
<i>Taula 5. Relació de la variable Nacionalitat amb els riscos psicosocials i les respostes</i>			

Les diferències més importants es troben en el control sobre el treball on les persones amb nacionalitat espanyola semblen les menys exposades.

Seguidament es mostra el mateix gràfic que els anteriors aquest cop estratificat per el sector de treball.

	Transport	Comerç i hosteleria	Resta sector serveis	Indústria	Construcció	Agricultura
Exigències Psicològiques	28.05%	29.53%	28.64%	21.83%	21.51%	24.29%
Doble presencia	17.07%	39.53%	38.79%	26.02%	15.05%	22.86%
Control del Treball	39.74%	37.90%	25.09%	33.60%	34.94%	52.24%
Suport Social	23.38%	20.91%	16.22%	16.23%	16.73%	32.84%
Estima	35.00%	28.61%	25.06%	26.40%	30.55%	36.23%
Salut Mental	45.68%	33.26%	31.89%	25.74%	26.52%	31.43%
Salut auto percebuda	9.76%	10.67%	9.59%	11.26%	6.12%	1.43%
Afectació de la feina en la Salut	41.98%	34.97%	41.29%	41.13%	48.38%	38.57%
Total	3.20%	16.80%	46.34%	20.07%	10.87%	2.73%
<i>Taula 6. Relació de la variable Sector amb els riscos psicosocials i les respostes</i>						

Els valors més extrems es troben en els sectors de transport i agricultura però cal tenir en compte que son els sectors amb menys representació per la qual cosa les dades poden no ser totalment representatives.

Aquests dos gràfics mostren l'estratificació a partir del tipus de contracte i la durada de la jornada laboral respectivament.

	Contracte fix/indefinit/permanent	Contracte Temporal
Exigències Psicològiques	27.12%	22.47%
Doble presencia	30.22%	47.19%
Control del Treball	30.66%	34.86%
Suport Social	17.65%	18.21%
Estima	27.35%	25.88%
Salut Mental	29.81%	36.72%
Salut auto percebuda	8.94%	13.17%
Afectació de la feina en la Salut	41.91%	34.93%
Total	86.07%	13.93%

Taula 7. Relació de la variable Contracte amb els riscos psicosocials i les respostes

	Temps Complet	Temps Parcial
Exigències Psicològiques	26.75%	25.47%
Doble presencia	33.30%	29.62%
Control del Treball	27.32%	49.77%
Suport Social	16.72%	22.40%
Estima	25.83%	33.04%
Salut Mental	29.77%	34.88%
Salut auto percebuda	9.63%	9.01%
Afectació de la feina en la Salut	41.46%	38.48%
Total	81.41%	18.59%

Taula 8. Relació de la variable Horari amb els riscos psicosocials i les respostes

Cal tenir en compte al veure aquests gràfics que la representació de persones amb un contracte temporal o que treballen a temps parcial és molt inferior que les altres respostes.

Finalment l'última taula mostra l'estratificació per sou.

	600 o menys	Entre 601 i 1000	Entre 1001 i 1500	Entre 1501 i 2000	Mes de 2000
Exigències Psicològiques	19.18%	24.12%	26.20%	29.75%	37.58%
Doble presencia	46.90%	40.18%	27.42%	32.60%	33.56%
Control del Treball	44.85%	42.37%	34.95%	13.24%	3.38%
Suport Social	17.29%	21.28%	21.19%	10.80%	7.43%
Estima	23.74%	32.02%	30.36%	19.89%	9.46%
Salut Mental	36.11%	33.19%	28.75%	30.28%	28.38%
Salut auto percebuda	17.12%	11.79%	8.37%	7.42%	6.04%
Afectació de la feina en la Salut	69.66%	58.86%	58.50%	54.95%	57.82%
Total	6.27%	29.52%	42.16%	15.64%	6.40%

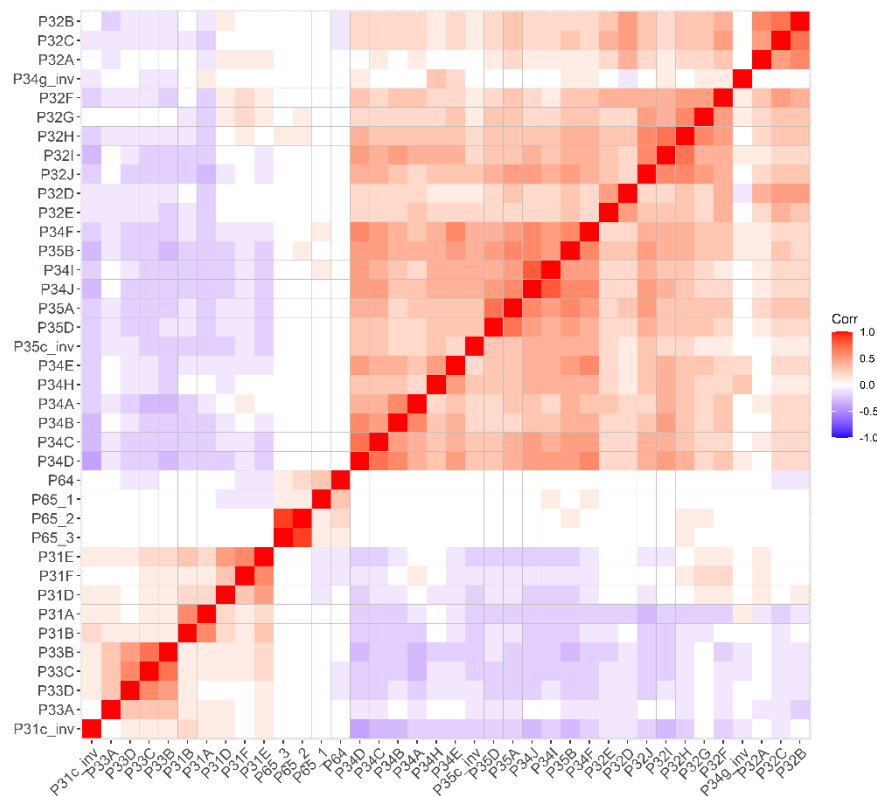
Taula 9. Relació de la variable Salari amb els riscos psicosocials i les respostes

És interessant comprovar que el control que te un individu sobre el treball que realitza augmenta amb el sou però també implica més exigència.

3.2 Estudi del qüestionari

3.2.1 Relació entre les qüestions

Un dels objectiu d'aquest treball és analitzar la pròpia enquesta, en concret les preguntes que ajuden a definir les dimensions de riscos psicosocials. Primerament s'ha estudiat la correlació entre aquestes preguntes i comprovar així quines podien resultar redundants, alhora que ajuda a veure si la construcció de les dimensions es correspon amb la relació de les preguntes.



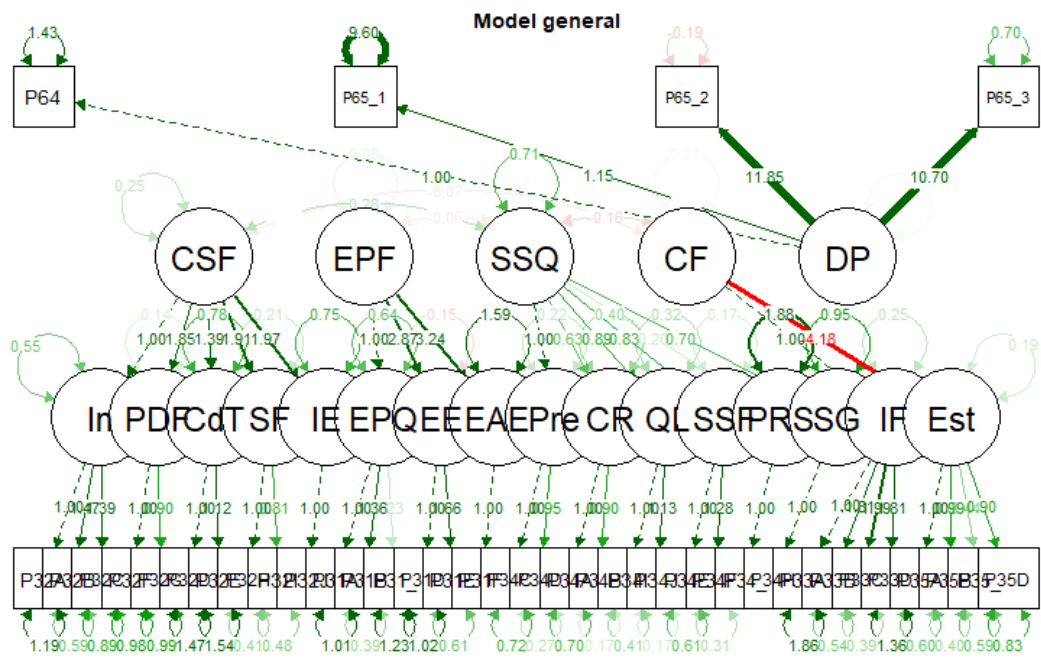
Gràfic 1. Correlació entre les qüestions que construeixen les dimensions de riscos psicosocials.

Abans de treure conclusions d'aquest gràfic cal tenir en compte que les preguntes 31 i 33 per tal com estan formulades donen una resposta inversa a la resta de preguntes, per tant no s'hauria de veure com la formació de tres dimensions sinó que el fet de que un individu pugui o no estar exposat a la risc de la doble presència no esta correlacionat amb el fet que aquest mateix individu pugui estar exposat a la resta de riscos.

També podem veure en el gràfic com les qüestions P65_2 i P65_3 estan fortament correlacionades, el que ens porta a pensar que podrien ser redundants. Les dues preguntes defineixen una part dels riscos d'estar exposat a la doble presència.

3.2.2 Anàlisi factorial confirmatori

Amb la intenció de confirmar la hipòtesi plantejada anteriorment sobre la redundància de les qüestions P65_2 i P65_3 o altres, s'aplica un anàlisi factorial confirmatori. L'estructura d'aquest esta explicada en detall a la Taula 1, tot i així cal remarcar que és un model complex ja que inclou l'estudi de la formació de les dimensions a partir de les subdimensions. Estudiant també la formació de les segones amb cada qüestió.

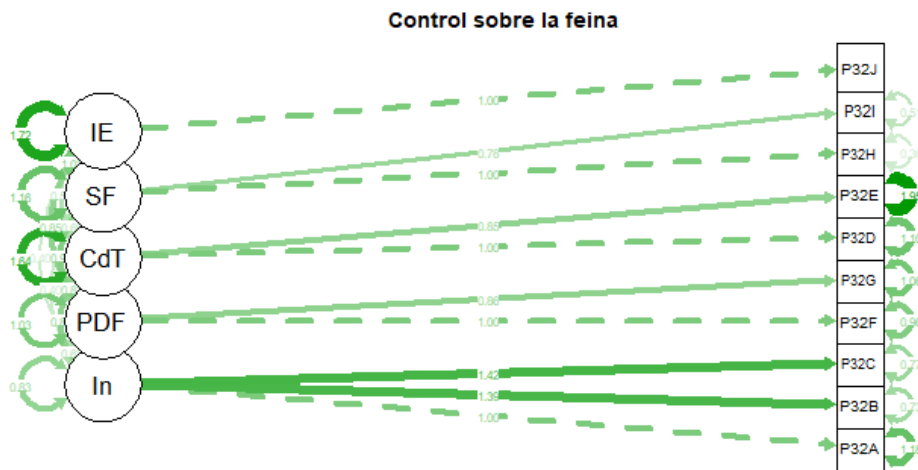


Gràfic 2. Anàlisi dimensional per CFA

Del gràfic anterior en destaca per una banda com l'estima (P35) i la inseguretat (P33) descriuen les compensacions a la feina de forma inversa. Això és fàcil d'entendre al veure com estan plantejades les qüestions de cada una de les preguntes ja que les respostes esperades són inverses.

D'altra banda podem comprovar com les dues qüestions mencionades anteriorment, la P65_2 i P65_3, descriuen molt considerablement la Doble presència. Aquest gràfic però, dona una visió molt general de com és el model i quines particularitats presenta, per això s'ha estudiat cada dimensió per separat;

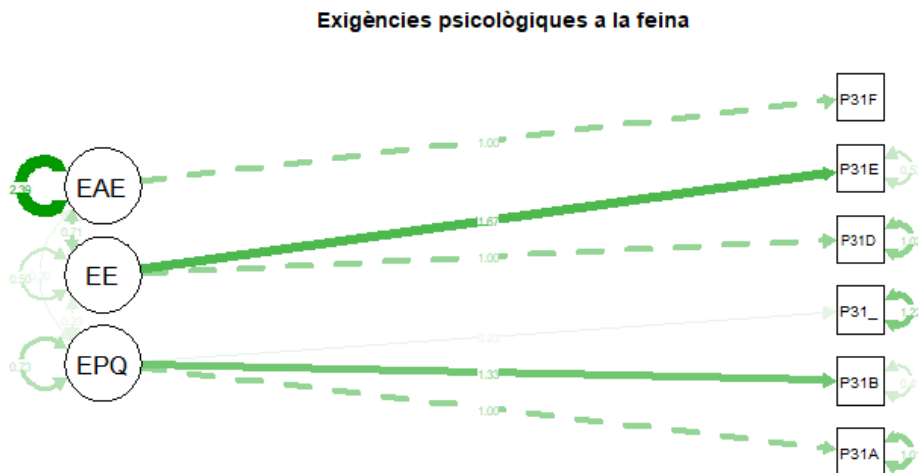
3.2.2.1 Control sobre la feina



Gràfic 3. Anàlisi de la dimensió CSF

Tal com s'intueix en el gràfic general aquesta dimensió no presenta cap tipus de dada alarmant i tenint en compte que el p-valor és pràcticament nul, podem corroborar que la construcció d'aquesta dimensió s'adequa als valors observats en la mostra.

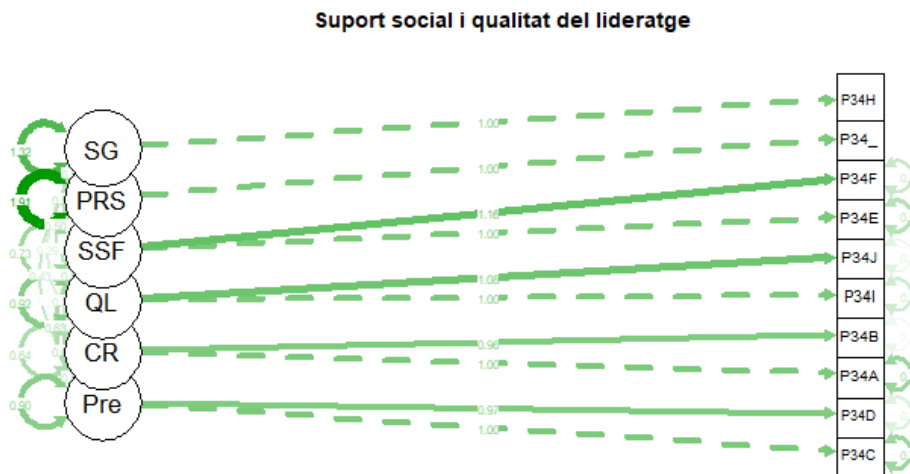
3.2.2.2 Exigències sobre la feina



Gràfic 4. Anàlisi de la dimensió EPF

En aquest gràfic podem destacar l'alta correlació en la subdimensió, exigències d'amagar emocions, tot i així el p-valor de la construcció d'aquesta dimensió igual que en l'anterior també és pràcticament nul per la qual cosa pensem que les dades observades corroboren el model teoritzat.

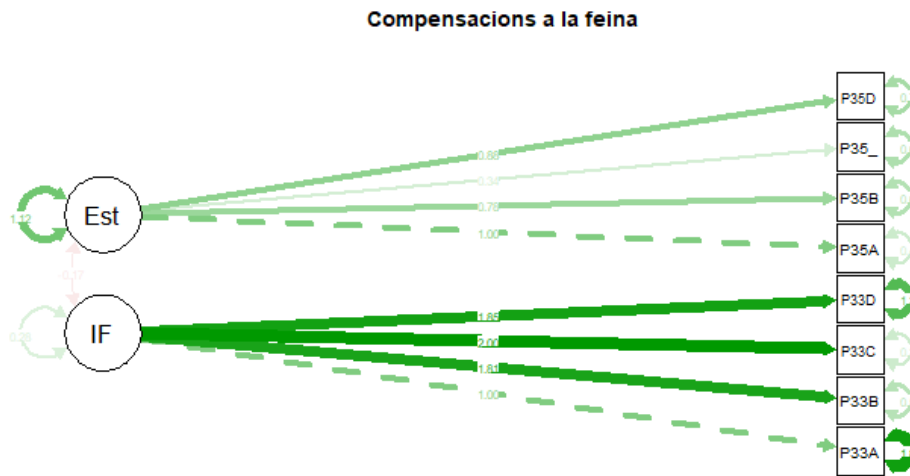
3.2.2.3 Suport social i qualitat del lideratge



Gràfic 5. Anàlisi de la dimensió SSIQL

Igual que en els dos gràfics anteriors no hi ha cap valor que ressalti i el p-valor també es pràcticament nul per això les dades corroboren el model teoritzat.

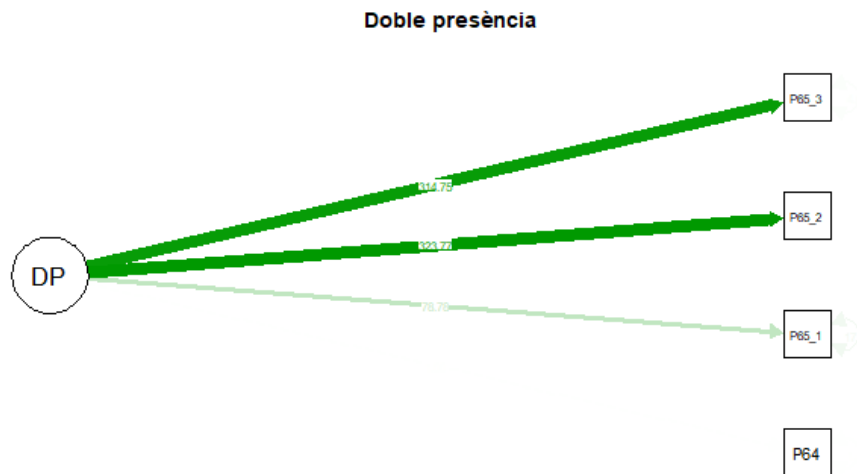
3.2.2.4 Compensacions a la feina



Gràfic 6. Anàlisi de la dimensió CF

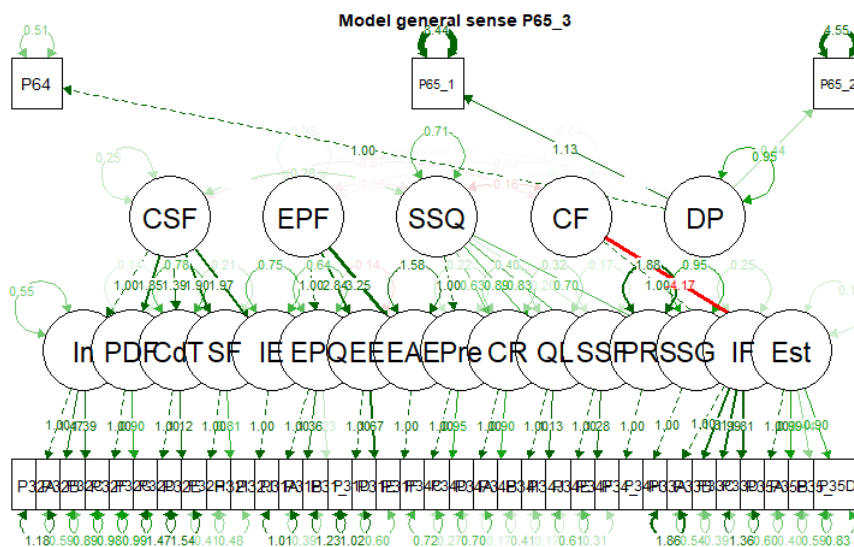
En aquest cas tal com s'ha comentat en el gràfic general veiem que la relació entre les dues subdimensions és negativa però sabent que es deu a la manera com estan fetes les preguntes no és alarmant ni afecta als resultats així que també posem pensar que el model teoritzat esta en acord amb les dades observades.

3.2.2.5 Doble presencia



Gràfic 7. Anàlisi de la dimensió DP

En aquest cas la doble presencia esta explicada per les dues qüestions que havien preocupat per la correlació que tenen. Veient doncs que la dimensió DP depèn bàsicament d'aquestes dues s'ha decidit eliminar la que menys pes té, la P65_3 (314,75).



Gràfic 8. Anàlisi dimensional per CFA sense la P65_3

Ometent les respostes de la qüestió redundant el model de riscos psicosocials queda de la següent forma.

És pràcticament igual que el primer amb la mateixa relació negativa en les compensacions a la feina però sense els dos valors extremadament alts en la dimensió de doble presencia.

Per la reconstrucció de la variable que indica quan un individu esta exposat a la doble presència s'ha utilitzat el mateix mètode que amb la resta de dimensions explicat anteriorment, ometent les respostes a la qüestió P65_3 que com bé s'ha dit s'han considerat redundants.

3.3 Model predictiu

Tot i que el model preveu el càlcul de l'"*out of bag error*" s'ha separat el set de dades en un 80% per l'entrenament del mètode i un 20% per la posterior avaluació.

En cada cas s'ha calculat un random forest de 500 arbres amb tres variables per escollir en cada node i cap límit màxim de nodes finals. Ja que no hi ha evidència científica que avaluï que per més arbres o més variables de decisió es millori la fiabilitat de la predicció.(Oshiro T.M., 2012)

Com s'ha explicat anteriorment el random forest permet veure quin efecte té cada variable en la predicció de la resposta. En la següent taula es mostren les dues mesures que proporciona el model. La primera, "MeanDecreaseAccuracy" indica com decreix l'exactitud de la predicció quan s'exclou la variable en qüestió. La segona mesura "MeanDecreaseGini" indica quin decreixement té la impuresa de Gini quan la variable s'utilitza per dividir el node.

3.3.1 Salut Mental

La primera variable a la que s'ha aplicat el model és una variable creada com s'explica a l'inici d'aquest apartat a partir de la pregunta 44. És una variable dicotòmica que indica si l'individu esta en risc de patir una malaltia mental.

Variable	No exposició	Exposició	MeanDecreaseAccuracy	MeanDecreaseGini
Residencia	0.0326378004	0.0590582685	0.0410831041	115.62916
Sexe	0.0044216100	0.0002387024	0.0030402778	20.22940
Edat	0.0040392078	-0.0002781505	0.0026313401	60.26211
Nacionalitat	0.0008718704	-0.0033109323	-0.0004779437	22.04173
Sector	0.0046392411	0.0080458267	0.0057255121	56.77800
Horari	0.0036920304	0.0001365811	0.0025674552	15.60594
Contracte	0.0020066786	0.0017119101	0.0019078979	15.45490
Salari	0.0039028189	0.0087000609	0.0054193423	26.16053
ExigPsico2k	0.0080528204	0.0126511507	0.0094879171	22.35228
DoblePres2k	0.0032927431	0.0066326973	0.0043543165	22.31321
ControlTreb2k	0.0224500170	-0.0065733378	0.0131585811	21.26289
SuportSocial2k	0.0525155946	0.0101869226	0.0389766909	34.31234
Estima2k	0.0197119583	0.0035106233	0.0145350371	22.85667
Cond_AmbientalsNou2k ¹	0.0049851450 ¹	0.0041053157 ¹	0.0046938230 ¹	21.60419 ¹

Taula 10. Importancia de les variables explicatives per la predicció de la salut mental

Gracies a aquesta taula veiem com les variables que descriuen les condicions laborals són el conjunt que més afecta tot i així sorprèn com la residència és la variable que més pes pren en l'exactitud de la predicció. Pel que fa a la interpretació de la segona mesura, el decreixement de l'índex de Gini cal tenir en compte que les tres variables amb un valor més alt; Residencia, Edat i Sector, són variables amb més de cinc categories mentre que la majoria de la resta son dicotòmiques.

Dades observades / Dades predites	No exposició	Exposició
No exposició predita	284	95
Exposició predita	22	40

Taula 11. Predicció per la variable Salut Mental

Per la resposta de Salut mental el model aconsegueix predir en el millor dels casos fins a un 74,88% dels casos quan un individu esta en risc de patir una malaltia mental. Això si amb un error tipus dos del 17,98%, tal i com es pot veure a la taula de la dreta.

3.3.2 Salut auto percebuda

Aquesta és la única resposta no dicotòmica, on la persona enquestada valora en una escala de l'1 al 5 la qualitat de la seva pròpia salut. El fet que la variable tingui mes

de dues categories com a resposta, pot implicar menys capacitat predictiva pel model. Com en el cas anterior primerament s'analitza la taula de resultats;

Variable	Excel·lent	Molt bona	Bona	Regular	Dolenta	MeanDecreaseAccuracy	MeanDecreaseGini
Residencia	0.003747702	0.03083894567	0.00964232214	0.00371043322	0.0014642857	0.0168614901	116.03764
Sexe	-0.009364138	0.00554592949	0.00564396284	0.00213222861	0.0046642857	0.0039567258	33.82800
Edat	0.011148038	0.01027259446	0.02661578518	0.00431771415	-0.0002865079	0.0167771869	90.38445
Nacionalitat	-0.001276970	-0.00258951906	0.00669759062	0.00005692705	0.0016666667	0.0017557178	31.58804
Sector	-0.007406567	0.00694771797	0.01477808003	-0.00170112058	-0.0075452381	0.0082138994	82.81186
Horari	0.002824797	-0.00009384787	0.00306418936	0.00003275912	-0.0017365079	0.0015190272	27.46847
Contracte	0.003433941	0.00049018455	0.00215103453	0.00078316909	0.0104904762	0.0015945443	23.20862
Salari	0.001882808	0.00623820414	0.00009885514	0.01908675724	0.0125873016	0.0042306076	33.03637
ExigPsico2k	0.003444550	-0.00079660887	0.00009050927	-0.00267579113	-0.0065944444	-0.0002367723	33.46149
DoblePres2k	-0.004424327	0.00146657944	0.00514780465	0.00804156266	0.0027849206	0.0030587520	34.29331
ControlTreb2k	-0.002636984	0.02259266590	-0.00642828676	0.00320538219	-0.0028334055	0.0060822976	33.08898
SuportSocial2k	-0.001086115	0.01804266942	-0.00190561377	0.00601962041	0.0032682540	0.0066392403	25.71611
Estima2k	-0.004977626	0.02419184196	-0.00467608279	0.00442582484	0.0010849206	0.0073165113	28.71349
Cond_AmbientalsNou2k	-0.003121381	0.00503087990	0.00210748096	-0.00085585034	0.0025920635	0.0025932749	32.15006

Taula 12. Importancia de les variables explicatives per la predicció de la salut autopercebuda

Per aquesta variable resposta les explicatives que més influeixen són les que tenen més de dues categories com l'Edat o la Residencia de l'individu.

Dades observades / Dades predites	Excel·lent	Molt bona	Bona	Regular	Dolenta
Excel·lent	0	2	2	0	0
Molt bona	18	102	62	8	1
Bona	18	77	118	28	4
Regular	0	2	1	3	0
Dolenta	0	0	0	0	0

Taula 13. Predicció per la variable Salut Autopercebuda

El que implica tenir una resposta amb varies possibles categories es un descens important de la capacitat predictiva, mentre que amb la ultima variable dicotomica estudiada predia el comportament de practicament tres de cada quatre persones, amb una variable categoria la capacitat predictiva és de només 49,5% i un 24,89% d'error tipus II. En aquest cas pero l'error de tipusII no és tant preocupant com en el cas anterior ja que com es pot veure en la taula la majoria dels errors es troba en les primeres tres categories de la resposta el que implica que la majoria dels individus considera que te una salut bona o fins i tot millor que bona.

3.3.3 Percepció de l'afecte de la feina sobre la salut

Finalment la última variable resposta correspon a l'evaluació que fa cada individu sobre l'efecte que té la seva feina en la seva salut. Ja que la qüestió només contempla SI/NO com a resposta la variable resultat és dicotòmica.

Variable	Afecta	No afecta	MeanDecreaseAccuracy	MeanDecreaseGini
Residencia	0.0108427610	0.0184176920	0.0153176193	91.28839
Sexe	0.0090812090	-0.0046423654	0.0009241515	21.50106
Edat	0.0012792971	0.0021284184	0.0017800143	66.96148
Nacionalitat	0.0013833158	0.0048181689	0.0034166991	25.45200
Sector	0.0104080658	-0.0020871719	0.0030331044	58.33396
Horari	-0.0005089500	0.0002313066	-0.0000603343	17.14335
Contracte	-0.0017566205	0.0021228762	0.0005370822	14.22246
Salari	-0.0001008407	0.0005383158	0.0002816314	21.57849
ExigPsico2k	0.0246878282	0.0260086538	0.0254570942	43.53123
DoblePres2k	0.0080075479	-0.0032593745	0.0013382772	21.31430
ControlTreb2k	-0.0053483213	0.0095806718	0.0034668756	21.55210
SuportSocial2k	-0.0010613844	0.0150357639	0.0084179553	18.25925
Estima2k	0.0026201516	0.0288252027	0.0180449797	30.66079
Cond_AmbientalsNou2k	0.0319281573	0.0285208801	0.0298270254	49.29926

Taula 14. Importància de les variables explicatives per la predicció de l'afecte percebut de la feina sobre la salut

Igual que en la primera resposta, les variables que més influeixen a predir són el lloc de residència dels individus i les dimensions dels riscos psicosocials. Apart de les que hem vist amb la primera resposta en aquest cas també són molt importants les condicions ambientals.

Dades observades / Dades predites	Afecta	No afecta
Afecta	86	49
No afecta	92	217

Taula 15. Predicció per la resposta de l'afecte percebut de la feina sobre la salut

Al tractar-se d'una variable dicotòmica s'entén que la capacitat predictiva s'assimili a la primera variable. En aquest cas és de 67,57% i un error tipus II de 12,16%.

4. Conclusions

Per finalitzar amb l'estudi realitzar cal recordar quins eren els dos objectius: el primer era avaluar la construcció de les dimensions dels riscos psicosocials que pateix la població treballadora de Catalunya. I el segon, en base als resultats del primer, és trobar un model predictiu que ajudi a detectar quan l'exposició excessiva a aquests riscos, posa en risc la salut mental dels individus.

En relació al primer objectiu, avaluar la construcció de les dimensions dels riscos psicosocials que pateix la població treballadora de Catalunya, cal tenir en compte que l'anàlisi s'ha enfocat amb dos punts de vista:

Per una banda, s'ha volgut comprovar si hi havia preguntes que podien ser redundants, per així reduir la llargada de l'enquesta (que recordem que té una mitjana de 27,4 minuts).

D'altra banda, s'ha buscat corroborar que la construcció del model a partir de dimensions i subdimensions es correspon amb la mostra observada.

Així doncs s'ha calculat la correlació entre les qüestions veient que dins la dimensió DP hi havia dues qüestions on es respon pràcticament el mateix, la P65_2 i la P65_3. Seguidament amb l'anàlisi factorial confirmatori s'ha pogut comprovar que degut a la correlació entre les dues prenen pràcticament tot el protagonisme dins la dimensió i que prescindint de la P65_3 els pesos dins el model s'equilibren amb les altres dues qüestions.

Amb la intenció d'assolir el segon objectiu s'ha utilitzat un algoritme de classificació, més concretament un algoritme d'aprenentatge supervisat que rep el nom de "*Random forest*". Amb aquest algoritme s'assoleix l'objectiu de saber quan una persona té un risc considerable de patir una malaltia de salut mental a partir de l'exposició als diferents riscos laborals explicats al treball.

L'assertivitat de les prediccions d'aquest mètode està entre el 70% i el 75%, és a dir, que de cada quatre persones enquestades, es pot predir l'efecte de les condicions laborals a la salut mental d'aproximadament tres. Tot i així cal remarcar que en les dades usades apareix un error de tipus II en el 22,91% dels casos, és a dir, que aquest mètode proporciona una quantitat elevada de falsos negatius.

El perquè d'aquest mètode?

Tal com s'ha explicat de forma més extensa en el treball, aquest mètode s'ha escollit ja que permet veure quines són les variables que donen més fiabilitat a les

prediccions. D'entre les cinc dimensions la més rellevant en tots els casos és el suport social i qualitat del lideratge. Tot i així sorprèn com variables de caire sociodemogràfic també prenen un pes important en la predicció. Com el cas de la residència dels individus.

Per últim, cal destacar dins la discussió sobre els resultats aconseguits, el tractament de les dades mancants. Per tal de facilitar els càlculs s'han omès totes aquelles observacions en les que hi havia respostes sense contestar. Aquests, juntament amb altres aspectes, poden marcar els següents passos a realitzar per millorar l'eina que brinda aquest estudi.

En definitiva els següents passos a seguir a partir d'aquest estudi són: per una banda els referents al model predictiu i per l'altra tot allò que té a veure amb la construcció de les dimensions i la relació d'aquestes amb les variables sociodemogràfiques:

En referència a la predicció és interessant comprovar a que es deu l'alt efecte de la variable residència a la fiabilitat de les prediccions que és totalment inesperat. Però sobretot es pensa que l'efecte de les dades mancants pot ser elevat. Una de les hipòtesis a estudiar en aquesta direcció és que els individus que eviten respondre segons quines qüestions ho fan per la incomoditat els genera.

De la relació entre les variables sociodemogràfiques i les dimensions de riscos psicosocials, el més destacable és la diferència entre gèneres en l'exposició a la doble presència. També és remarcable com el control sobre el treball realitzat és més baix entre les persones joves, la gent que treballa en el sector agrícola, i en les categories més baixes de sou.

Finalment, enfocant aquesta eina en l'actualitat, és interessant comprovar com ha pogut afectar la pandèmia tant a les condicions de treball com en la repercussió que tenen aquestes a la salut mental o general de les persones.

En conclusió, tal com s'havia marcat en els objectius aquest treball corrobora la construcció de l'enquesta COPSQ a partir de les dades recollides a Catalunya l'any 2010 (II ECTT). També brinda una eina que ajuda a les empreses a comprovar l'estat de salut de totes les persones que hi treballen, al mateix temps, que ajuda a aquestes persones a veure com les condicions laborals afecten a la seva salut.

5. Bibliografía

- Alhija, F. A. N. (2010). Factor analysis: An overview and some contemporary advances. *International Encyclopedia of Education*, 162–170. <https://doi.org/10.1016/B978-0-08-044894-7.01328-2>
- How is Variable Importance Calculated for a Random Forest?* (n.d.).
- Idescat. Demografia i Societat.* (n.d.).
- Instituto Sindical del Trabajo Ambiente y Salud. (2010). *Manual del método para la evaluación y prevención de los riesgos psicosociales para empresas con 25 o más trabajadores y trabajadoras.* 21(versión 2), 217.
- J. Boissevain. (n.d.). *Factor Analysis and Latent Structure, Confirmatory.*
- Moncada, S., Llorens, C., Navarro, A., Kristensen, T. S., & Moncada Lluís, S. (2005). *ISTAS21: the Spanish version of the Copenhagen psychosocial questionnaire (COPSOQ) ORIGINALES.* www.istas.net
- Nübling M, B. H. M. S. K. TS. (2013). COPSOQ International Network. *Co-Operation for Research and Assessment of Psychosocial Factors at Work.*
- Oshiro T.M., P. P. S. , B. J. A. (2012). How Many Trees in a Random Forest?. *Perner P. (Eds) Machine Learning and Data Mining in Pattern Recognition. MLDM 2012. Lecture Notes in Computer Science, 7376.*
- Rosseel, Y. (2014). The lavaan tutorial. *The Lavaan Tutorial*, 37. <http://lavaan.ugent.be/tutorial/tutorial.pdf>
- Using_random_forests_V3.1.* (n.d.).
- Utzet, M., Llorens, C., Morriña, D., & Moncada, S. (2021). Persistent inequality: evolution of psychosocial exposures at work among the salaried population in Spain between 2005 and 2016. *International Archives of Occupational and Environmental Health*, 94(4), 621–629. <https://doi.org/10.1007/s00420-020-01609-3>
- World Health Organization. Health definition.* (n.d.).
- Josh Starner. 2018 (<https://www.youtube.com/channel/UCtYLUtG3k1Fg4y5tAhLbw>)

6. Annex

Llibreries necessaries:

```
```{r}
#dades:
library(haven)
library(tidyverse)
library(naniar)
library(tidyr)
library(lubridate)

#CFA
library(lavaan)
library(eqs2lavaan)
```

```
#Grafiques
library(semPlot)
library(ggplot2)
library(GGally)
library(ggcorrplot)
library(ggraph)
library(tidygraph)
library(flextable)
library(webshot)
```

```
#RF
library(randomForest)
library(ranger)
library(tidymodels)
library(doParallel)
library(missForest)
library(tibble)
```
```

Entrada de dades:

```
```{r}
path <- file.path("C:/Users/comho/OneDrive/Escritorio/Uni", "TFG", "Treballadors_II-
ECCT_2010.sav")
fulldataset <- read_sav(path)
pathcarpeta <- "C:/Users/comho/OneDrive/Escritorio/UniTFG"
dd <- fulldataset[which(fulldataset$P1 == 1 & (fulldataset$P7 == 1 | fulldataset$P7 ==
2)),] #Asalariats i contracte fix, permanent o indefinit
```
```

Filtre ns/nc

```
```{r}
```

```

aux8i9 <- dd[,c(27,28,32,33,79:84,87:92,97:198,227:269,271,276:282)]
aux8i9[aux8i9 == 8] <- NA
aux8i9[aux8i9 == 9] <- NA
dd[,c(27,28,32,33,79:84,87:92,97:198,227:269,271,276:282)] <- aux8i9
aux98i99 <- dd[,c(37,41:58,60:78,85,86,93,94,96,199:227,270,272:275,283)]
aux98i99[aux98i99 == 98] <- NA
aux98i99[aux98i99 == 99] <- NA
dd[,c(37,41:58,60:78,85,86,93,94,96,199:227,270,272:275,283)] <- aux98i99
remove(aux8i9)
remove(aux98i99)
...

```

Descriptiva bàsica:

```

```{r}
aux <- dd[,c(13,16,18,21,23,59,65,283)]
names(aux) <- c("Residencia", "Sexe", "Edat", "Nacionalitat", "Sector", "Horari",
"Contracte", "Salari")
aux <- as.data.frame(lapply(aux, factor))

levels(aux$Residencia) <- c("Àmbit Metropolità", "Com. gironines", "Camp de
Tarragona", "Com. centrals", "Terres de l'Ebre", "Àmbit de Ponent", "Alt Pirineu i Aran")
levels(aux$Sexe) <- c("Home", "Dona")
levels(aux$Edat) <- c("De 16 a 24", "De 25 a 34", "De 35 a 44", "De 45 a 54", "Mes de
54")
levels(aux$Nacionalitat) <- c("Espanyola", "Europea", "Fora d'Europa")
levels(aux$Sector) <- c("Transport", "Comerç i hosteleria", "Resta sector serveis",
"Indústria", "Construcció", "Agricultura")
levels(aux$Contracte) <- c("Contracte fix/indefinit/permanent", "Contracte Temporal")
levels(aux$Horari) <- c("Temps Complert", "Temps Parcial")
levels(aux$Salari) <- c("600 o menys", "Entre 601 i 1000", "Entre 1001 i 1500", "Entre
1501 i 2000", "Mes de 2000")

percent(aux$Residencia)[1:7]
percent(aux$Sexe)[1:2]
percent(aux$Edat)[1:5]
percent(aux$Nacionalitat)[1:3]
percent(aux$Sector)[1:6]
percent(aux$Horari)[1:2]
percent(aux$Contracte)[1:2]
percent(aux$Salari)[1:5]

prop.table(table(aux$Horari))
prop.table(table(aux$Contracte))
prop.table(table(aux$Contracte, aux$Horari))
mean(dd$Q3_1); sd(dd$Q3_1)

```



```
inici <- dmy_hms(paste(dd$FECHAINI, dd$HORAINI))
final <- dmy_hms(paste(dd$FECHAFIN, dd$HORAFIN))
```

```
diff <- difftime(final, inici, units = "mins")
```

```
summary(as.numeric(diff))
mean(diff)
``
```

```
Corrplot
```

```
``{r}
corr <- round(cor(na.omit(dd[,rev(c(132,133,135:158,160:165,167,168,274,276,278,280))])), 1)
#p.mat <- cor_pmat(na.omit(dd[,rev(c(132,133,135:158,160:165,167,168,274,276,278,280))]))
pl <- ggcorrplot(corr,
  hc.order = T,
  type = "full",
#   p.mat = p.mat,
#   lab = T,
#   lab_size = 4,
#   method = "circle",
  ggtheme = theme_gray)+
  labs(caption = "Gràfic 1. Correlació entre les qüestions que construeixen les
dimensions de riscos psicosocials.") +
  theme(plot.caption = element_text(hjust = 0.7, face = "italic", size = 15))
#pl
ggsave("corplot.png", pl, width=10, height=10, dpi=700)
``
```

```
Grafic correlacions per dimensio
```

```
``{r}
#Exigencies psicològiques a la feina
cor.epf <- round(cor(na.omit(dd[,rev(c(132,133,135:138))])),2)
p.mat.epf <- cor_pmat(dd[,rev(c(132,133,135:138))])
ggcorrplot(cor.epf, hc.order = F,
  type = "upper",
#   p.mat = p.mat.epf,
  lab = T,
  lab_size = 4,
  method = "circle",
  ggtheme = theme_gray)
```

```
#Control sobre la feina
```

```
cor.csf <- round(cor(na.omit(dd[,rev(c(139:148))])),2)
p.mat.csf <- cor_pmat(dd[,rev(c(139:148))])
```

```

ggcorrplot(cor.csf, hc.order = F,
            type = "upper",
#           p.mat = p.mat.csf,
            lab = T,
            lab_size = 4,
            method = "circle",
            ggtheme = theme_gray)

#Exigencies psicologiques a la feina
cor.ssql <- round(cor(na.omit(dd[,rev(c(153:158,160:163))])),2)
p.mat.ssql <- cor_pmat(dd[,rev(c(153:158,160:163))])
ggcorrplot(cor.ssql, hc.order = F,
            type = "upper",
#           p.mat = p.mat.ssql,
            lab = T,
            lab_size = 4,
            method = "circle",
            ggtheme = theme_gray)

#Exigencies psicologiques a la feina
cor.cf <- round(cor(na.omit(dd[,rev(c(149:151,164,165,167,168))])),2)
p.mat.cf <- cor_pmat(dd[,rev(c(149:151,164,165,167,168))])
ggcorrplot(cor.cf, hc.order = F,
            type = "upper",
#           p.mat = p.mat.cf,
            lab = T,
            lab_size = 4,
            method = "circle",
            ggtheme = theme_gray)

#Exigencies psicologiques a la feina
cor.dp <- round(cor(na.omit(dd[,rev(c(274,276,278,280))])),2)
p.mat.dp <- cor_pmat(dd[,rev(c(274,276,278,280 ))])
ggcorrplot(cor.dp, hc.order = F,
            type = "upper",
#           p.mat = p.mat.dp,
            lab = T,
            lab_size = 4,
            method = "circle",
            ggtheme = theme_gray)
...

mfa
```{r}
model.CSF <- 'ln =~ P32A + P32B + P32C
 PDF =~ P32F + P32G

```

```

CdT =~ P32D + P32E
SF =~ P32H + P32I
IE =~ P32J'

model.EPF <- 'EPQ =~ P31A + P31B + P31c_inv
EE =~ P31D + P31E
EAE =~ P31F'

model.SSiQL <- 'Pre =~ P34C + P34D
CR =~ P34A + P34B
QL =~ P34I + P34J
SSF =~ P34E + P34F
PRS =~ P34g_inv
SG =~ P34H'

model.CF <- 'IF =~ P33A + P33B + P33C + P33D
Est =~ P35A + P35B + P35c_inv + P35D'

model.DP <- 'DP =~ P64 + P65_1 + P65_2 + P65_3'

model.total <- 'In =~ P32A + P32B + P32C
PDF =~ P32F + P32G
CdT =~ P32D + P32E
SF =~ P32H + P32I
IE =~ P32J
EPQ =~ P31A + P31B + P31c_inv
EE =~ P31D + P31E
EAE =~ P31F
Pre =~ P34C + P34D
CR =~ P34A + P34B
QL =~ P34I + P34J
SSF =~ P34E + P34F
PRS =~ P34g_inv
SG =~ P34H
IF =~ P33A + P33B + P33C + P33D
Est =~ P35A + P35B + P35c_inv + P35D
DP =~ P64 + P65_1 + P65_2 + P65_3'
...

CSF:
```{r}
fit.csf <- cfa(model.CSF, data = dd)
summary(fit.csf)

png("CSF.png", width = 750, height = 375)
semPaths(fit.csf, what = "par", rotation = 2)
title(main = "Control sobre la feina", sub = "Gràfic 3. Anàlisi de la dimensió CSF")

```

```
dev.off()
````
```

EPF:

```
````{r}
```

```
fit.epf <- cfa(model.EPF, data = dd)
summary(fit.epf)
```

```
png("EPF.png", width = 750, height = 375)
semPaths(fit.epf, what = "par", title = T, rotation = 2)
title(main = "Exigències psicològiques a la feina", sub = "Gràfic 4. Anàlisi de la dimensió EPF")
dev.off()
````
```

SSiQL:

```
````{r}
```

```
fit.ssql <- cfa(model.SSiQL, data = dd)
summary(fit.ssql)
```

```
png("SSiQL.png", width = 750, height = 375)
semPaths(fit.ssql, what = "par", rotation = 2, title = T)
title(main = "Suport social i qualitat del lideratge", sub = "Gràfic 5. Anàlisi de la dimensió SSiQL")
dev.off()
````
```

CF:

```
````{r}
```

```
fit.cf <- cfa(model.CF, data = dd)
summary(fit.cf)
```

```
png("CF.png", width = 750, height = 375)
semPaths(fit.cf, what = "par", rotation = 2, title = T)
title(main = "Compensacions a la feina", sub = "Gràfic 6. Anàlisi de la dimensió CF")
dev.off()
````
```

DP:

```
````{r}
```

```
fit.dp <- cfa(model.DP, data = dd)
summary(fit.dp, fit.measures = T)
```

```
png("DP.png", width = 750, height = 375)
semPaths(fit.dp, what = "par", rotation = 2, title = T)
title(main = "Doble presència", sub = "Gràfic 7. Anàlisi de la dimensió DP")
dev.off()
```

...

Total:

``{r}

```
fit <- cfa(model.total, data = dd)
```

```
summary(fit, fit.measures = T)
```

```
semPaths(fit, what = "par")
```

...

Nou model sense P65C

``{r}

```
model.totalnou <- 'ln =~ P32A + P32B + P32C
```

```
    PDF =~ P32F + P32G
```

```
    CdT =~ P32D + P32E
```

```
    SF =~ P32H + P32I
```

```
    IE =~ P32J
```

```
    EPQ =~ P31A + P31B + P31c_inv
```

```
    EE =~ P31D + P31E
```

```
    EAE =~ P31F
```

```
    Pre =~ P34C + P34D
```

```
    CR =~ P34A + P34B
```

```
    QL =~ P34I + P34J
```

```
    SSF =~ P34E + P34F
```

```
    PRS =~ P34g_inv
```

```
    SG =~ P34H
```

```
    IF =~ P33A + P33B + P33C + P33D
```

```
    Est =~ P35A + P35B + P35c_inv + P35D
```

```
    DP =~ P64 + P65_1 + P65_2'
```

```
modelsubdimencions <- 'ln =~ P32A + P32B + P32C
```

```
    PDF =~ P32F + P32G
```

```
    CdT =~ P32D + P32E
```

```
    SF =~ P32H + P32I
```

```
    IE =~ P32J
```

```
    EPQ =~ P31A + P31B + P31c_inv
```

```
    EE =~ P31D + P31E
```

```
    EAE =~ P31F
```

```
    Pre =~ P34C + P34D
```

```
    CR =~ P34A + P34B
```

```
    QL =~ P34I + P34J
```

```
    SSF =~ P34E + P34F
```

```
    PRS =~ P34g_inv
```

```
    SG =~ P34H
```

```
    IF =~ P33A + P33B + P33C + P33D
```

```
    Est =~ P35A + P35B + P35c_inv + P35D
```

```
    CSF =~ ln + PDF + CdT + SF + IE
```

```
EPF =~ EPQ + EE + EAE
SSiQL =~ Pre + CR + QL + SSF + PRS + SG
CF =~ IF + Est
DP =~ P64 + P65_1 + P65_2 + P65_3'
```

...

model guai: posarlo el primer de tots

```
``{r}
fit.modelsubdimencions <- cfa(modelsubdimencions, data = dd)
summary(fit.modelsubdimencions, fit.measures = T)
semPaths(fit.modelsubdimencions, what = "par")
title(main = "Model general", sub = "Gràfic 2. Anàlisi dimencional per CFA")

png("Modelguai1.png", width = 1000, height = 500)
semPaths(fit.modelsubdimencions, what = "par")
title(main = "Model general", sub = "Gràfic 2. Anàlisi dimencional per CFA")
dev.off()
...

```

Dp:

```
``{r}
model.DPnou <- 'DP =~ P64 + P65_1 + P65_2'
fit.dpnou <- cfa(model.DPnou, data = dd)
summary(fit.dpnou, fit.measures = T)
semPaths(fit.dpnou, what = "par", rotation = 2, title = T)
title(main = "Doble presència sense P65_3")
...

```

Total:

```
``{r}
modelsubdimencions2 <- 'ln =~ P32A + P32B + P32C
    PDF =~ P32F + P32G
    CdT =~ P32D + P32E
    SF =~ P32H + P32I
    IE =~ P32J
    EPQ =~ P31A + P31B + P31c_inv
    EE =~ P31D + P31E
    EAE =~ P31F
    Pre =~ P34C + P34D
    CR =~ P34A + P34B
    QL =~ P34I + P34J
    SSF =~ P34E + P34F
    PRS =~ P34g_inv
    SG =~ P34H
    IF =~ P33A + P33B + P33C + P33D

```

```

Est =~ P35A + P35B + P35c_inv + P35D
CSF =~ In + PDF + CdT + SF + IE
EPF =~ EPQ + EE + EAE
SSiQL =~ Pre + CR + QL + SSF + PRS + SG
CF =~ IF + Est
DP =~ P64 + P65_1 + P65_2'
fit.modelsubdimencions2 <- cfa(modelsubdimencions2, data = dd)
summary(fit.modelsubdimencions2, fit.measures = T)

png("Modelguai2.png", width = 1000, height = 500)
semPaths(fit.modelsubdimencions2, what = "par")
title(main = "Model general sense P65_3", sub = "Gràfic 8. Anàlisi dimencional per CFA
sense la P65_3")
dev.off()
...

```

Corr sense P65C

```

```{r}
corr <- round(cor(na.omit(dd[,rev(c(132,133,135:158,160:165,167,168,274,276,278))])),
1)
p.mat <- cor_pmat(na.omit(dd[,rev(c(132,133,135:158,160:165,167,168,274,276,278))]))
ggcorrplot(corr,
 hc.order = T,
 type = "full",
p.mat = p.mat,
 lab = T,
 lab_size = 4,
 method = "circle",
 ggtheme = theme_gray)
...

```

Reconstruccio DP

```

```{r}
dd$DoblePres <- rowSums(dd[,c(275,277,279)])
quantile(dd$DoblePres, c(0.33,0.66), na.rm = T)
dd$DoblePres_C <- ifelse(dd$DoblePres<3, 1, ifelse(dd$DoblePres<=5, 2, 3))
dd$DoblePres2k <- ifelse(dd$DoblePres_C < 3, 0, 1)
...

```

Quantils P44 (Salut Mental)

```

```{r}
dd$SalutMental <- rowSums(dd[,227:238])
quantile(dd$SalutMental, c(0.33,0.66), na.rm = T)
dd$SalutMental_C <- ifelse(dd$SalutMental<=18, 1, ifelse(dd$SalutMental<=22, 2, 3))
dd$SalutMental2k <- ifelse(dd$SalutMental_C < 3, 0, 1)

```

```
table(dd$SalutMental);table(dd$SalutMental_C);table(dd$SalutMental2k)
```
```

Exposició riscos ambientals/físics:

```
```{r}
summary(dd[,97:132])
corr <- round(cor(na.omit(dd[,rev(97:132)])), 1)
p.mat <- cor_pmat(na.omit(dd[,rev(97:132)]))
ggcorrplot(corr,
 hc.order = F,
 type = "full",
p.mat = p.mat,
 lab = T,
 lab_size = 2,
 method = "circle",
 ggtheme = theme_gray)
dd <- add_column(dd, P29b_inv = ((dd$P29B - 2)*-1)+2, .after = 110)

dd$Cond_AmbientalsNou <- rowSums(dd[,c(97:109,111:132)])
quantile(dd$Cond_AmbientalsNou, c(0.33,0.66), na.rm = T)
dd$Cond_AmbientalsNou_C <- ifelse(dd$Cond_AmbientalsNou<=12, 1,
ifelse(dd$Cond_AmbientalsNou<=28, 2, 3))
dd$Cond_AmbientalsNou2k <- ifelse(dd$Cond_AmbientalsNou_C < 3, 0, 1)
table(dd$Cond_AmbientalsNou);table(dd$Cond_AmbientalsNou_C);table(dd$Cond_AmbientalsNou2k)
summary(dd[,c(97:109,111:132)])
```
```

- Edat Q3_2
- Sexe Q2
- Contracte fixe/temporal P7
- Temps parcial/complet P12
- Nacionalitat (Spain vs UE vs Resta) Q4Cat
- Sector ocupacional Q6
- Territori Q1
- Salari P66
- Exposició a riscos ambientals/físics (P28 (expo a fenomens ambientals), P29(expo a riscos físics), P30(expo a riscos varios)) alt no ens agrada
- Dimensions de riscos psicosocials

80%:

```
```{r}
set.seed(1234)
ddrf <- cbind(aux, dd[,c(317:321,332,198,199,328,329)])
```



```

ddrf[,15:18] <- as.data.frame(lapply(ddrf[,15:18], factor))

levels(ddrf[,15]) <- c("Excel·lent", "Molt bona", "Bona", "Regular", "Dolenta")
levels(ddrf[,16]) <- c("Afecta", "No afecta")
levels(ddrf[,17]) <- c("Gens exposat", "Poc exposat", "Exposat")
levels(ddrf[,18]) <- c("No exposició", "Exposició")

colnames(ddrf[,15:18]) = c("Salut autopercebuda", "Afectació de la feina
autopercebuda", "Salut mentals en quantils", "Salut mental binaria")

train <- sample(1:nrow(dd), size = round(0.8*nrow(dd)))

ddtrain <- na.omit(ddrf[train,-c(15:17)])
ddtest <- na.omit(ddrf[-train,-c(15:17)])

ddtrainsalmen <- na.omit(ddrf[train,-c(15, 16,18)])
ddtestsalmen <- na.omit(ddrf[-train,-c(15, 16,18)])

ddtrainsalgen <- na.omit(ddrf[train,-c(15, 17,18)])
ddtestsalgen <- na.omit(ddrf[-train,-c(15, 17,18)])

ddtrainfeina <- na.omit(ddrf[train,-c(16:18)])
ddtestfeina <- na.omit(ddrf[-train,-c(16:18)])

summary(ddrf)
```

Salut mental 2k
```{r}
rf <- randomForest(factor(SalutMental2k) ~ ., importance = T, data = ddtrain,
keep.forest = T)
#bosc <- rf$forest
#bosc$nodepred
pred <- predict(rf, ddtest)
levels(pred) <- c("No exposició predita", "Exposició predita")
(res <- table(pred , ddtest$SalutMental2k))
sum(diag(res))/nrow(ddtest)
93/nrow(ddtest)
as.data.frame.matrix(res)
ftres <- flextable(as.data.frame.matrix(res) %>% rownames_to_column("Dades
observades / Dades predites"))
ftres <- add_footer(x = ftres, `Dades observades / Dades predites` = "Taula 11.
Predicció per la variable Salut Mental")
ftres <- merge_at(x = ftres, j = 1:3, part = "footer")
ftres <- theme_box(ftres)
ftres

```

```

ftrf <- flextable(cbind(Variable =
colnames(ddrf[1:14]),as.data.frame.matrix(rf$importance, row.names = T)))
ftrf <- footnote(ftrf, i = 14, j = 1:5, value = as_paragraph("iokese"))
ftrf <- add_footer(x = ftrf, Variable = "Taula 10. Importancia de les variables
explicatives per la predicció de la salut mental")
ftrf <- merge_at(x = ftrf, j = 1:5, part = "footer")
ftrf <- theme_box(ftrf)
ftrf
save_as_image(ftres, path = "Predicció RF Salut Mental.png")
save_as_image(ftrf, path = "Resultat RF Salut Mental.png")
...

```

```

```{r}
#resultats <- c()
#for (i in 1:14) {
# for (j in c(100,250,500,750,1000)) {
#   rf <- randomForest(factor(SalutMental2k) ~ ., importance = T, data = ddtrain,
keep.forest = T, mtry = i, ntree = j)
#   pred <- predict(rf, ddtest)
#   res <- table(pred , ddtest$SalutMental2k)
#   encert <- sum(diag(res))/nrow(ddtest)
#   error1 <- res[1,2]/nrow(ddtest)
#   error2 <- res[2,1]/nrow(ddtest)
#   resultats <- c(resultats, encert, error1, error2)
# }
# print(i)
#}
#comparacio <- matrix(resultats, ncol = 15, byrow = T)
#which.max(comparacio); comparacio
...

```

```

3/250: 0.7487685 0.1995074 0.05172414
4/500: 0.7487685 0.1798030 0.07142857
3/750: 0.7487685 0.2068966 0.04433498
4/750: 0.7487685 0.1822660 0.06896552

```

```

5/100: 0.7189873 0.1924051 0.08860759
5/250: 0.7189873 0.1974684 0.08354430
6/500: 0.7189873 0.1949367 0.08607595

```

Afectació feina cap a la Salut

```

```{r}
rf3 <- randomForest(factor(P42A) ~ ., importance = T, data = ddtrainsalgen)
rf3$importance
pred3 <- predict(rf3, ddtestsalgen)
res3 <- table(pred3 , ddtestsalgen$P42A)
(capacitatpredictiva3 <- sum(diag(res3))/nrow(ddtestsalgen))

```

```

(errortipusll3 <- sum(res3[upper.tri(res3, diag = F)])/nrow(ddtestsalgen))
ftres3 <- flextable(as.data.frame.matrix(res3) %>% rownames_to_column("Dades
observades / Dades predites"))
ftres3 <- add_footer(x = ftres3, `Dades observades / Dades predites` = "Taula 15.
Predicció per la resposta de l'afecte percebut de la feina sobre la salut")
ftres3 <- merge_at(x = ftres3, j = 1:3, part = "footer")
ftres3 <- theme_box(ftres3)
ftres3
ftrf3 <- flextable(cbind(Variable =
colnames(ddrf[1:14]),as.data.frame.matrix(rf3$importance, row.names = T)))
ftrf3 <- add_footer(x = ftrf3, Variable = "Taula 14. Importancia de les variables
explicatives per la predicció de l'afecte percebut de la feina sobre la salut")
ftrf3 <- merge_at(x = ftrf3, j = 1:5, part = "footer")
ftrf3 <- theme_box(ftrf3)
ftrf3
save_as_image(ftres3, path = "Predicció RF Feina vs Salut.png")
save_as_image(ftrf3, path = "Resultat RF Feina vs Salut.png")
...

```

```

```{r}
#resultats <- c()
#for (i in 1:14) {
# for (j in c(100,250,500,750,1000)) {
#   rf <- randomForest(factor(P42A) ~ ., importance = T, data = ddtrainsalgen,
keep.forest = T, mtry = i, ntree = j)
#   pred <- predict(rf, ddtestsalgen)
#   res <- table(pred , ddtestsalgen$P42A)
#   encert <- sum(diag(res))/nrow(ddtestsalgen)
#   error1 <- res[1,2]/nrow(ddtestsalgen)
#   error2 <- res[2,1]/nrow(ddtestsalgen)
#   resultats <- c(resultats, encert, error1, error2)
# }
# print(i)
#}
#comparacio <- matrix(resultats, ncol = 15, byrow = T)
#which.max(comparacio); comparacio
...

```

Salut autopercebuda

```

```{r}
rf4 <- randomForest(factor(P41) ~ ., importance = T, data = ddtrainfeina)
pred4 <- predict(rf4, ddtestfeina)
res4 <- table(pred4 , ddtestfeina$P41)
sum(diag(res4))/nrow(ddtestfeina)
sum(res4[upper.tri(res4, diag = F)])/nrow(ddtestfeina)

```

```

sum(res4[lower.tri(res4, diag = F)]/nrow(ddtestfeina)
ftres4 <- flextable(as.data.frame.matrix(res4) %>% rownames_to_column("Dades
observades / Dades predites"))
ftres4 <- add_footer(x = ftres4, `Dades observades / Dades predites` = "Taula 13.
Predicció per la variable Salut Autopercebuda")
ftres4 <- merge_at(x = ftres4, j = 1:6, part = "footer")
ftres4 <- theme_box(ftres4)
ftres4
ftrf4 <- flextable(cbind(Variable =
colnames(ddrf[1:14]),as.data.frame.matrix(rf4$importance, row.names = T)))
ftrf4 <- add_footer(x = ftrf4, Variable = "Taula 12. Importancia de les variables
explicatives per la predicció de la salut autopercebuda")
ftrf4 <- merge_at(x = ftrf4, j = 1:8, part = "footer")
ftrf4 <- theme_box(ftrf4)
ftrf4
save_as_image(ftres4, path = "Predicció RF Salut autoper.png")
save_as_image(ftrf4, path = "Resultat RF Salut autoper.png")
...

```

taules sexe vs riscos psicosocials

```

```{r}
#taux <- (matrix(rep(0,4), nrow = 2))
#
#tsex1 <- rbind(matrix(sprintf("%.2f%%",
round(prop.table(table(ddrf$ExigPsico2k,ddrf$Sexe))*100,2)),nrow = 2), taux)
#tsex2 <- rbind(matrix(sprintf("%.2f%%",
round(prop.table(table(ddrf$DoblePres2k,ddrf$Sexe))*100,2)),nrow = 2), taux)
#tsex3 <- rbind(matrix(sprintf("%.2f%%",
round(prop.table(table(ddrf$ControlTreb2k,ddrf$Sexe))*100,2)),nrow = 2), taux)
#tsex4 <- rbind(matrix(sprintf("%.2f%%",
round(prop.table(table(ddrf$SuportSocial2k,ddrf$Sexe))*100,2)),nrow = 2), taux)
#tsex5 <- rbind(matrix(sprintf("%.2f%%",
round(prop.table(table(ddrf$Estima2k,ddrf$Sexe))*100,2)),nrow = 2), taux)
#
#colnames(tsex1) <- levels(ddrf$Sexe)
#colnames(tsex2) <- levels(ddrf$Sexe)
#colnames(tsex3) <- levels(ddrf$Sexe)
#colnames(tsex4) <- levels(ddrf$Sexe)
#colnames(tsex5) <- levels(ddrf$Sexe)
#
#for (i in 1:2) {
# tsex1[3:4,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Sexe
==
levels(ddrf$Sexe)[i])]))*100, 2))

```

```

#           tsexe2[3:4,i]           <-           sprintf("%.2f%%",
round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Sexe
levels(ddrf$Sexe)[i]])))*100, 2))           ==
#           tsexe3[3:4,i]           <-           sprintf("%.2f%%",
round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Sexe
levels(ddrf$Sexe)[i]])))*100, 2))           ==
#           tsexe4[3:4,i]           <-           sprintf("%.2f%%",
round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Sexe
levels(ddrf$Sexe)[i]])))*100, 2))           ==
#           tsexe5[3:4,i]           <-           sprintf("%.2f%%",
round(prop.table(table(ddrf$Estima2k[which(ddrf$Sexe == levels(ddrf$Sexe)[i]])))*100,
2))
#}
#
#b <- c("No exposicio Total", "Exposicio Total", "No exposicio Categorica", "Exposicio
Categorica")
#tfootsexex <- rbind(rep("Exigences Psicologiques",2), tsexe1, rep("Doble presencia", 2),
tsexe2, rep("Control del Treball", 2), tsexe3, rep("Suport #Social", 2), tsexe4,
rep("Estima", 2), tsexe5)
#tfootsexex <- cbind(a = c("Exigences Psicologiques", b, "Doble presencia", b, "Control
del Treball", b, "Suport Social", b, "Estima", #b),as.data.frame.matrix(ttfootsexex))
#tfootsexex <- flextable(ttfootsexex)
#tfootsexex <- delete_part(ttfootsexex, part = "header")
#tfootsexex <- add_header(x = ttfootsexex, Home = "Homes", Dona ="Dones", top = T)
#tfootsexex <- merge_h(x = ttfootsexex, i = c(1,6,11,16,21))
#tfootsexex <- theme_box(ttfootsexex)
#tfootsexex
#save_as_image(ttfootsexex, path = "Taula Sexe.png")
...

```

Descriptives

residencia

```
``{r}
```

```
taux1 <- matrix(rep(c("Exigències Psicològiques", "Doble presencia", "Control del
Trebball", "Suport Social", "Estima", "Salut Mental", "Salut auto percebuda", "Afectació
de la feina en la Salut", "Total"), each = length(levels(ddrf$Residencia))+1), nrow = 9,
byrow = T)
```

```
tresidencia <- taux1
```

```
i <- 1
```

```
for (i in 2:8) {
```

```
  tresidencia[1,i]           <-           sprintf("%.2f%%",
round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Residencia
levels(ddrf$Residencia)[i-1]])))*100, 2))[2]           ==
```

```
  tresidencia[2,i]           <-           sprintf("%.2f%%",
round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Residencia
levels(ddrf$Residencia)[i-1]])))*100, 2))[2]           ==
```

```

tresidencia[3,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Residencia
levels(ddrf$Residencia)[i-1]]))*100, 2))[2]
tresidencia[4,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Residencia
levels(ddrf$Residencia)[i-1]]))*100, 2))[2]
tresidencia[5,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$Estima2k[which(ddrf$Residencia
levels(ddrf$Residencia)[i-1]]))*100, 2))[2]
tresidencia[6,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SalutMental2k[which(ddrf$Residencia
levels(ddrf$Residencia)[i-1]]))*100, 2))[2]
tresidencia[7,i] <- sprintf("%.2f%%",
sum(prop.table(table(ddrf$P41[which(ddrf$Residencia == levels(ddrf$Residencia)[i-
1]])))[4:5]*100))
tresidencia[8,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$P42A[which(ddrf$Residencia == levels(ddrf$Residencia)[i-
1]]))*100, 2))[1]
tresidencia[9,i] <- sprintf("%.2f%%", prop.table(table(ddrf$Residencia))*100)[i-1]
}

```

```

colnames(tresidencia) <- c("Variable", levels(ddrf$Residencia))
ftresidencia <- flextable(as.data.frame.matrix(tresidencia))
ftresidencia <- delete_part(ftresidencia, part = "header")
ftresidencia <- add_header(ftresidencia, `Àmbit Metropolità` = "Àmbit Metropolità",
`Com. gironines` = "Com. gironines", `Camp de Tarragona` = "Camp de Tarragona",
`Com. centrals` = "Com. centrals", `Terres de l'Ebre` = "Terres de l'Ebre", `Àmbit de
Ponent` = "Àmbit de Ponent", `Alt Pirineu i Aran` = "Alt Pirineu i Aran", top = T)
ftresidencia <- bold(x = ftresidencia, part = "body", i = 1:9, j = 1)
ftresidencia <- add_footer(ftresidencia, Variable = "Taula 2. Relació de la variable
Residencia amb els riscos psicosocials i les respostes")
ftresidencia <- merge_at(x = ftresidencia, j = 1:8, part = "footer")
ftresidencia <- theme_box(ftresidencia)
ftresidencia
save_as_image(ftresidencia, path = "Taula Residencia.png")
...

```

Sexe

```

```{r}
taux1 <- matrix(rep(c("Exigències Psicològiques", "Doble presència", "Control del
Trebball", "Suport Social", "Estima", "Salut Mental", "Salut auto percebuda", "Afectació
de la feina en la Salut", "Total"), each = length(levels(ddrf$Sexe))+1, nrow = 9, byrow =
T)
tsexe <- taux1
i <- 1
for (i in 2:3) {

```

```

tsexex[1,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Sexe ==
levels(ddrf$Sexe)[i-1]])))*100, 2))[2]
tsexex[2,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Sexe ==
levels(ddrf$Sexe)[i-1]])))*100, 2))[2]
tsexex[3,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Sexe ==
levels(ddrf$Sexe)[i-1]])))*100, 2))[2]
tsexex[4,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Sexe ==
levels(ddrf$Sexe)[i-1]])))*100, 2))[2]
tsexex[5,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$Estima2k[which(ddrf$Sexe ==
levels(ddrf$Sexe)[i-1]])))*100, 2))[2]
tsexex[6,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SalutMental2k[which(ddrf$Sexe ==
levels(ddrf$Sexe)[i-1]])))*100, 2))[2]
tsexex[7,i] <- sprintf("%.2f%%", sum(prop.table(table(ddrf$P41[which(ddrf$Sexe ==
levels(ddrf$Sexe)[i-1]])))[4:5]*100))
tsexex[8,i] <- sprintf("%.2f%%", round(prop.table(table(ddrf$P42A[which(ddrf$Sexe ==
levels(ddrf$Sexe)[i-1]])))*100, 2))[1]
tsexex[9,i] <- sprintf("%.2f%%", prop.table(table(ddrf$Sexe))*100)[i-1]
}

```

```

colnames(tsexex) <- c("Variable", levels(ddrf$Sexe))
ftsexex <- ftable(as.data.frame.matrix(tsexex))
ftsexex <- delete_part(ftsexex, part = "header")
ftsexex <- add_header(ftsexex, Home = "Homes", Dona = "Dones", top = T)
ftsexex <- bold(x = ftsexex, part = "body", i = 1:9, j = 1)
ftsexex <- add_footer(x = ftsexex, Variable = "Taula 3. Relació de la variable Sexe amb els
riscos psicosocials i les respostes")
ftsexex <- merge_at(x = ftsexex, j = 1:3, part = "footer")
ftsexex <- theme_box(ftsexex)
ftsexex
save_as_image(ftsexex, path = "Taula Sexe.png")
``

```

Estat

```
``{r}
```

```

taux1 <- matrix(rep(c("Exigències Psicològiques", "Doble presència", "Control del
T treball", "Suport Social", "Estima", "Salut Mental", "Salut auto percebuda", "Afectació
de la feina en la Salut", "Total"), each = length(levels(ddrf$Estat))+1), nrow = 9, byrow =
T)
tedat <- taux1
i <- 1
for (i in 2:6) {

```

```

tedat[1,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Edat ==
levels(ddrf$Edat)[i-1]])))*100, 2))[2]
tedat[2,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Edat ==
levels(ddrf$Edat)[i-1]])))*100, 2))[2]
tedat[3,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Edat ==
levels(ddrf$Edat)[i-1]])))*100, 2))[2]
tedat[4,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Edat ==
levels(ddrf$Edat)[i-1]])))*100, 2))[2]
tedat[5,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$Estima2k[which(ddrf$Edat ==
levels(ddrf$Edat)[i-1]])))*100, 2))[2]
tedat[6,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SalutMental2k[which(ddrf$Edat ==
levels(ddrf$Edat)[i-1]])))*100, 2))[2]
tedat[7,i] <- sprintf("%.2f%%", sum(prop.table(table(ddrf$P41[which(ddrf$Edat ==
levels(ddrf$Edat)[i-1]])))[4:5]*100))
tedat[8,i] <- sprintf("%.2f%%", round(prop.table(table(ddrf$P42A[which(ddrf$Edat ==
levels(ddrf$Edat)[i-1]])))*100, 2))[1]
tedat[9,i] <- sprintf("%.2f%%", prop.table(table(ddrf$Edat))*100)[i-1]
}

```

```

colnames(tedat) <- c("Variable", levels(ddrf$Edat))
ftedat <- flextable(as.data.frame.matrix(tedat))
ftedat <- delete_part(ftedat, part = "header")
ftedat <- add_header(ftedat, `De 16 a 24` = "De 16 a 24", `De 25 a 34` = "De 25 a 34",
`De 35 a 44` = "De 35 a 44", `De 45 a 54` = "De 45 a 54", `Mes de 54` = "Mes de 54",
top = T)
ftedat <- bold(x = ftedat, part = "body", i = 1:9, j = 1)
ftedat <- add_footer(x = ftedat, Variable = "Taula 4. Relació de la variable Edat amb els
riscos psicosocials i les respostes")
ftedat <- merge_at(x = ftedat, j = 1:6, part = "footer")
ftedat <- theme_box(ftedat)
ftedat
save_as_image(ftedat, path = "Taula Edat.png")
```

```

Nacionalitat

```
```{r}
```

```

taux1 <- matrix(rep(c("Exigències Psicològiques", "Doble presencia", "Control del
T treball", "Suport Social", "Estima", "Salut Mental", "Salut auto percebuda", "Afectació
de la feina en la Salut", "Total"), each = length(levels(ddrf$Nacionalitat))+1), nrow = 9,
byrow = T)
tnacionalitat <- taux1

```



```

i <- 1
for (i in 2:4) {
 tnacionalitat[1,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Nacionalitat
levels(ddrf$Nacionalitat)[i-1])))*100, 2))[2]
 tnacionalitat[2,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Nacionalitat
levels(ddrf$Nacionalitat)[i-1])))*100, 2))[2]
 tnacionalitat[3,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Nacionalitat
levels(ddrf$Nacionalitat)[i-1])))*100, 2))[2]
 tnacionalitat[4,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Nacionalitat
levels(ddrf$Nacionalitat)[i-1])))*100, 2))[2]
 tnacionalitat[5,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$Estima2k[which(ddrf$Nacionalitat
levels(ddrf$Nacionalitat)[i-1])))*100, 2))[2]
 tnacionalitat[6,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SalutMental2k[which(ddrf$Nacionalitat
levels(ddrf$Nacionalitat)[i-1])))*100, 2))[2]
 tnacionalitat[7,i] <- sprintf("%.2f%%",
sum(prop.table(table(ddrf$P41[which(ddrf$Nacionalitat == levels(ddrf$Nacionalitat)[i-
1]))))[4:5]*100))
 tnacionalitat[8,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$P42A[which(ddrf$Nacionalitat
levels(ddrf$Nacionalitat)[i-1])))*100, 2))[1]
 tnacionalitat[9,i] <- sprintf("%.2f%%", prop.table(table(ddrf$Nacionalitat))*100)[i-1]
}

```

```

colnames(tnacionalitat) <- c("Variable", levels(ddrf$Nacionalitat))
ftnacionalitat <- flextable(as.data.frame.matrix(tnacionalitat))
ftnacionalitat <- delete_part(ftnacionalitat, part = "header")
ftnacionalitat <- add_header(ftnacionalitat, Espanyola = "Espanyola", Europea =
"Europea", `Fora d'Europa` = "Fora d'Europa", top = T)
ftnacionalitat <- bold(x = ftnacionalitat, part = "body", i = 1:9, j = 1)
ftnacionalitat <- add_footer(x = ftnacionalitat, Variable = "Taula 5. Relació de la
variable Nacionalitat amb els riscos psicosocials i les respostes")
ftnacionalitat <- merge_at(x = ftnacionalitat, j = 1:4, part = "footer")
ftnacionalitat <- theme_box(ftnacionalitat)
ftnacionalitat
save_as_image(ftnacionalitat, path = "Taula Nacionalitat.png")
```

```

Sector

```
```{r}
```

```

taux1 <- matrix(rep(c("Exigències Psicològiques", "Doble presencia", "Control del
Trell", "Suport Social", "Estima", "Salut Mental", "Salut auto percebuda", "Afectació

```

```
de la feina en la Salut", "Total"), each = length(levels(ddrf$Sector))+1), nrow = 9, byrow = T)
```

```
tsector <- taux1
```

```
i <- 1
```

```
for (i in 2:7) {
```

```
 tsector[1,i] <- sprintf("%.2f%%",
 round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Sector == levels(ddrf$Sector)[i-1])]) * 100, 2))[2])
```

```
 tsector[2,i] <- sprintf("%.2f%%",
 round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Sector == levels(ddrf$Sector)[i-1])]) * 100, 2))[2])
```

```
 tsector[3,i] <- sprintf("%.2f%%",
 round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Sector == levels(ddrf$Sector)[i-1])]) * 100, 2))[2])
```

```
 tsector[4,i] <- sprintf("%.2f%%",
 round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Sector == levels(ddrf$Sector)[i-1])]) * 100, 2))[2])
```

```
 tsector[5,i] <- sprintf("%.2f%%",
 round(prop.table(table(ddrf$Estima2k[which(ddrf$Sector == levels(ddrf$Sector)[i-1])]) * 100, 2))[2])
```

```
 tsector[6,i] <- sprintf("%.2f%%",
 round(prop.table(table(ddrf$SalutMental2k[which(ddrf$Sector == levels(ddrf$Sector)[i-1])]) * 100, 2))[2])
```

```
 tsector[7,i] <- sprintf("%.2f%%", sum(prop.table(table(ddrf$P41[which(ddrf$Sector == levels(ddrf$Sector)[i-1])]) * 100))
```

```
 tsector[8,i] <- sprintf("%.2f%%",
 round(prop.table(table(ddrf$P42A[which(ddrf$Sector == levels(ddrf$Sector)[i-1])]) * 100, 2))[1])
```

```
 tsector[9,i] <- sprintf("%.2f%%", prop.table(table(ddrf$Sector)) * 100)[i-1]
```

```
}
```

```
colnames(tsector) <- c("Variable", levels(ddrf$Sector))
```

```
ftsector <- flextable(as.data.frame.matrix(tsector))
```

```
ftsector <- delete_part(ftsector, part = "header")
```

```
ftsector <- add_header(ftsector, `Transport` = "Transport", `Comerç i hosteleria` = "Comerç i hosteleria", `Resta sector serveis` = "Resta sector serveis", `Indústria` = "Indústria", `Construcció` = "Construcció", `Agricultura` = "Agricultura", top = T)
```

```
ftsector <- bold(x = ftsector, part = "body", i = 1:9, j = 1)
```

```
ftsector <- add_footer(x = ftsector, Variable = "Taula 6. Relació de la variable Sector amb els riscos psicosocials i les respostes")
```

```
ftsector <- merge_at(x = ftsector, j = 1:7, part = "footer")
```

```
ftsector <- theme_box(ftsector)
```

```
ftsector
```

```
save_as_image(ftsector, path = "Taula Sector.png")
```

```
...
```

Contracte

```

```{r}
taux1 <- matrix(rep(c("Exigències Psicològiques", "Doble presencia", "Control del
Trellall", "Suport Social", "Estima", "Salut Mental", "Salut auto percebuda", "Afectació
de la feina en la Salut", "Total"), each = length(levels(ddrf$Contracte))+1), nrow = 9,
byrow = T)
tcontracte <- taux1
i <- 1
for (i in 2:3) {
  tcontracte[1,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Contracte
levels(ddrf$Contracte)[i-1])))*100, 2))[2]
  tcontracte[2,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Contracte
levels(ddrf$Contracte)[i-1])))*100, 2))[2]
  tcontracte[3,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Contracte
levels(ddrf$Contracte)[i-1])))*100, 2))[2]
  tcontracte[4,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Contracte
levels(ddrf$Contracte)[i-1])))*100, 2))[2]
  tcontracte[5,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$Estima2k[which(ddrf$Contracte
levels(ddrf$Contracte)[i-1])))*100, 2))[2]
  tcontracte[6,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SalutMental2k[which(ddrf$Contracte
levels(ddrf$Contracte)[i-1])))*100, 2))[2]
  tcontracte[7,i] <- sprintf("%.2f%%",
sum(prop.table(table(ddrf$P41[which(ddrf$Contracte == levels(ddrf$Contracte)[i-
1])))[4:5]*100))
  tcontracte[8,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$P42A[which(ddrf$Contracte == levels(ddrf$Contracte)[i-
1])))*100, 2))[1]
  tcontracte[9,i] <- sprintf("%.2f%%", prop.table(table(ddrf$Contracte))*100)[i-1]
}

```

```

colnames(tcontracte) <- c("Variable", levels(ddrf$Contracte))
ftcontracte <- flextable(as.data.frame.matrix(tcontracte))
ftcontracte <- delete_part(ftcontracte, part = "header")
ftcontracte <- add_header(ftcontracte, `Contracte fix/indefinit/permanent` =
"Contracte fix/indefinit/permanent", `Contracte Temporal` = "Contracte Temporal",
top = T)
ftcontracte <- bold(x = ftcontracte, part = "body", i = 1:9, j = 1)
ftcontracte <- add_footer(x = ftcontracte, Variable = "Taula 7. Relació de la variable
Contracte amb els riscos psicosocials i les respostes")
ftcontracte <- merge_at(x = ftcontracte, j = 1:3, part = "footer")
ftcontracte <- theme_box(ftcontracte)
ftcontracte

```

```
save_as_image(ftcontracte, path = "Taula Contracte.png")
```

```
...
```

```
Horari
```

```
``{r}
```

```
taux1 <- matrix(rep(c("Exigències Psicològiques", "Doble presencia", "Control del Treball", "Suport Social", "Estima", "Salut Mental", "Salut auto percebuda", "Afectació de la feina en la Salut", "Total"), each = length(levels(ddrf$Horari))+1), nrow = 9, byrow = T)
```

```
thorari <- taux1
```

```
i <- 1
```

```
for (i in 2:3) {
```

```
  thorari[1,i] <- sprintf("%.2f%%",  
round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Horari == levels(ddrf$Horari)[i-1]))])*100, 2))[2]
```

```
  thorari[2,i] <- sprintf("%.2f%%",  
round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Horari == levels(ddrf$Horari)[i-1]))])*100, 2))[2]
```

```
  thorari[3,i] <- sprintf("%.2f%%",  
round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Horari == levels(ddrf$Horari)[i-1]))])*100, 2))[2]
```

```
  thorari[4,i] <- sprintf("%.2f%%",  
round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Horari == levels(ddrf$Horari)[i-1]))])*100, 2))[2]
```

```
  thorari[5,i] <- sprintf("%.2f%%",  
round(prop.table(table(ddrf$Estima2k[which(ddrf$Horari == levels(ddrf$Horari)[i-1]))])*100, 2))[2]
```

```
  thorari[6,i] <- sprintf("%.2f%%",  
round(prop.table(table(ddrf$SalutMental2k[which(ddrf$Horari == levels(ddrf$Horari)[i-1]))])*100, 2))[2]
```

```
  thorari[7,i] <- sprintf("%.2f%%", sum(prop.table(table(ddrf$P41[which(ddrf$Horari == levels(ddrf$Horari)[i-1]))])[4:5]*100))
```

```
  thorari[8,i] <- sprintf("%.2f%%", round(prop.table(table(ddrf$P42A[which(ddrf$Horari == levels(ddrf$Horari)[i-1]))])*100, 2))[1]
```

```
  thorari[9,i] <- sprintf("%.2f%%", prop.table(table(ddrf$Horari))*100)[i-1]
```

```
}
```

```
colnames(thorari) <- c("Variable", levels(ddrf$Horari))
```

```
fthorari <- flextable(as.data.frame.matrix(thorari))
```

```
fthorari <- delete_part(fthorari, part = "header")
```

```
fthorari <- add_header(fthorari, `Temps Complert` = "Temps Complert", `Temps Parcial` = "Temps Parcial", top = T)
```

```
fthorari <- bold(x = fthorari, part = "body", i = 1:9, j = 1)
```

```
fthorari <- add_footer(x = fthorari, Variable = "Taula 8. Relació de la variable Horari amb els riscos psicosocials i les respostes")
```

```
fthorari <- merge_at(x = fthorari, j = 1:3, part = "footer")
```

```
fthorari <- theme_box(fthorari)
```

```
fthorari
save_as_image(fthorari, path = "Taula Horari.png")
``
```

Salari

```
``{r}
taux1 <- matrix(rep(c("Exigències Psicològiques", "Doble presencia", "Control del
Trellall", "Suport Social", "Estima", "Salut Mental", "Salut auto percebuda", "Afectació
de la feina en la Salut", "Total"), each = length(levels(ddrf$Salari))+1), nrow = 9, byrow
= T)
tsalari <- taux1
i <- 1
for (i in 2:6) {
  tsalari[1,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ExigPsico2k[which(ddrf$Salari == levels(ddrf$Salari)[i-
1])))*100, 2))[2]
  tsalari[2,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$DoblePres2k[which(ddrf$Salari == levels(ddrf$Salari)[i-
1])))*100, 2))[2]
  tsalari[3,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$ControlTreb2k[which(ddrf$Salari == levels(ddrf$Salari)[i-
1])))*100, 2))[2]
  tsalari[4,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SuportSocial2k[which(ddrf$Salari == levels(ddrf$Salari)[i-
1])))*100, 2))[2]
  tsalari[5,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$Estima2k[which(ddrf$Salari == levels(ddrf$Salari)[i-
1])))*100, 2))[2]
  tsalari[6,i] <- sprintf("%.2f%%",
round(prop.table(table(ddrf$SalutMental2k[which(ddrf$Salari == levels(ddrf$Salari)[i-
1])))*100, 2))[2]
  tsalari[7,i] <- sprintf("%.2f%%", sum(prop.table(table(ddrf$P41[which(ddrf$Salari ==
levels(ddrf$Salari)[i-1])))[4:5]*100))
  tsalari[8,i] <- sprintf("%.2f%%", round(prop.table(table(ddrf$P42A[which(ddrf$Salari
== levels(ddrf$Salari)[i-1])))*100, 2))[2]
  tsalari[9,i] <- sprintf("%.2f%%", prop.table(table(ddrf$Salari))*100)[i-1]
}
```

```
colnames(tsalari) <- c("Variable", levels(ddrf$Salari))
ftsalari <- flextable(as.data.frame.matrix(tsalari))
ftsalari <- delete_part(ftsalari, part = "header")
ftsalari <- add_header(ftsalari, `600 o menys` = "600 o menys", `Entre 601 i 1000` =
"Entre 601 i 1000", `Entre 1001 i 1500` = "Entre 1001 i 1500", `Entre 1501 i 2000` =
"Entre 1501 i 2000", `Mes de 2000` = "Mes de 2000", top = T)
ftsalari <- bold(x = ftsalari, part = "body", i = 1:9, j = 1)
ftsalari <- add_footer(x = ftsalari, Variable = "Taula 9. Relació de la variable Salari amb
els riscos psicosocials i les respostes")
```

```

ftsalari <- merge_at(x = ftsalari, j = 1:6, part = "footer")
ftsalari <- theme_box(ftsalari)
ftsalari
save_as_image(ftsalari, path = "Taula Salari.png")
```



```

```{r}
table(ddrf$SalutMental2k); table(ddrf$P42A); table(ddrf$P41)
table(ddrf$SalutMental2k, ddrf$P42A, ddrf$P41)
table(ddrf$P42A[ddrf$P41 %in% c("Excel·lent", "Molt bona", "Bona")])
table(ddrf$P42A[ddrf$P41 %in% c("Regular", "Dolenta")])
644/(644 + 1654)
860/(860 + 1451)
57/(57+184)
(2566-sum(ddrf$P41 %in% c("Excel·lent", "Molt bona", "Bona")))/2566
table(ddrf$SalutMental2k, ddrf$P42A)
397/(397+380)
table(ddrf$P42A)
```

```


```