



ANÁLISIS DE CRÉDITOS BANCARIOS

Análisis de datos categóricos - 2025
Miguel Bande Rodríguez



UNIVERSIDAD
COMPLUTENSE
MADRID



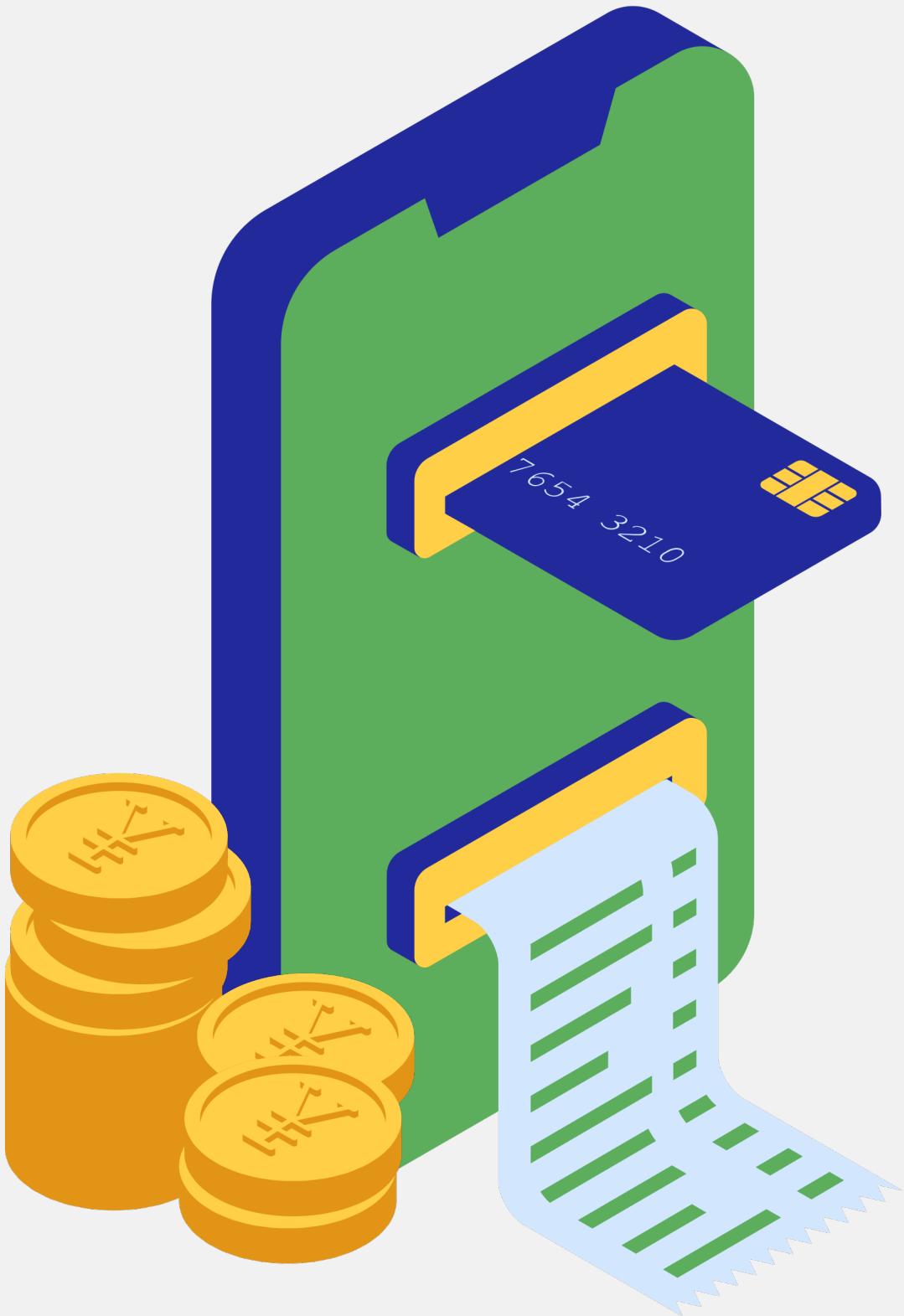
POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID

Índice

- 1.**Introducción y preprocesado de los datos
- 2.**Tarea 1. Tablas de contingencia.
- 3.**Tarea 2. Regresión logística
- 4.**Tarea 3. Regresión logística ordinal
- 5.**Aplicación





Introducción

- Base de datos con información sobre **créditos bancarios**:
 - Información sobre la persona solicitante
 - Información sobre el crédito
 - Información sobre si se ha pagado o no el crédito
- 4454 entradas en la base de datos



Introducción

#	Status	Seniority	Home	Time	Age	Marital	Records	Job	Expenses	Income	Assets	Debt	Amount	Price
5	good	0	rent	36	26	single	no	fixed	46	107	0	0	310	910
6	good	1	owner	60	36	married	no	fixed	75	214	3500	0	650	1645
7	good	29	owner	60	44	married	no	fixed	75	125	10000	0	1600	1800
8	good	9	parents	12	27	single	no	fixed	35	80	0	0	200	1093
9	good	0	owner	60	32	married	no	freelance	90	107	15000	0	1200	1957
10	bad	0	parents	48	41	married	no	partime	90	80	0	0	1200	1468
11	good	6	owner	48	34	married	no	freelance	60	125	4000	0	1150	1577
12	good	7	owner	36	29	married	no	fixed	60	121	3000	0	650	915
13	good	8	owner	60	30	married	no	fixed	75	199	5000	2500	1500	1650
14	good	19	priv	36	37	married	no	fixed	75	170	3500	260	600	940
15	bad	0	other	18	21	single	yes	partime	35	50	0	0	400	500
16	good	0	owner	24	68	married	no	fixed	75	131	4162	0	900	1186
17	good	15	priv	24	52	single	no	freelance	35	330	16500	0	1500	2201
18	good	33	rent	24	68	married	no	freelance	65	200	5000	2000	600	1350
19	bad	0	ignore	48	36	married	no	partime	45	130	750	0	1100	1511
20	good	1	parents	60	31	single	no	fixed	35	137	0	0	1250	1253
21	bad	2	rent	60	25	single	no	fixed	46	107	0	0	1500	2189



Introducción

- **Status:** Estado del crédito
- **Seniority:** Antigüedad en años del solicitante en su empleo actual.
- **Home:** Tipo de vivienda del solicitante (alquiler, propietario, ...)
- **Time:** Duración del préstamo en meses
- **Age:** Edad del solicitante
- **Marital:** Estado civil
- **Records:** Historial crediticio ("yes" si tiene antecedentes negativos, "no" si no los tiene)
- **Job:** Tipo de empleo ("fijo", "freelance",...)
- **Expenses:** Gastos mensuales del solicitante.
- **Income:** Ingresos mensuales del solicitante
- **Assets:** Valor de los activos del solicitante (propiedades, ahorros, etc.).
- **Debt:** Deuda actual del solicitante.
- **Amount:** Monto solicitado en el préstamo.
- **Price:** Precio del bien que se quiere adquirir con el préstamo.



Preprocesado de los datos



Paso 1:

Estudio de valores nulos



Paso 2:

Estudio de valores atípicos



Paso 1:

Balanceo de clases



Paso 1:

Transformaciones de variables



UNIVERSIDAD
COMPLUTENSE
MADRID

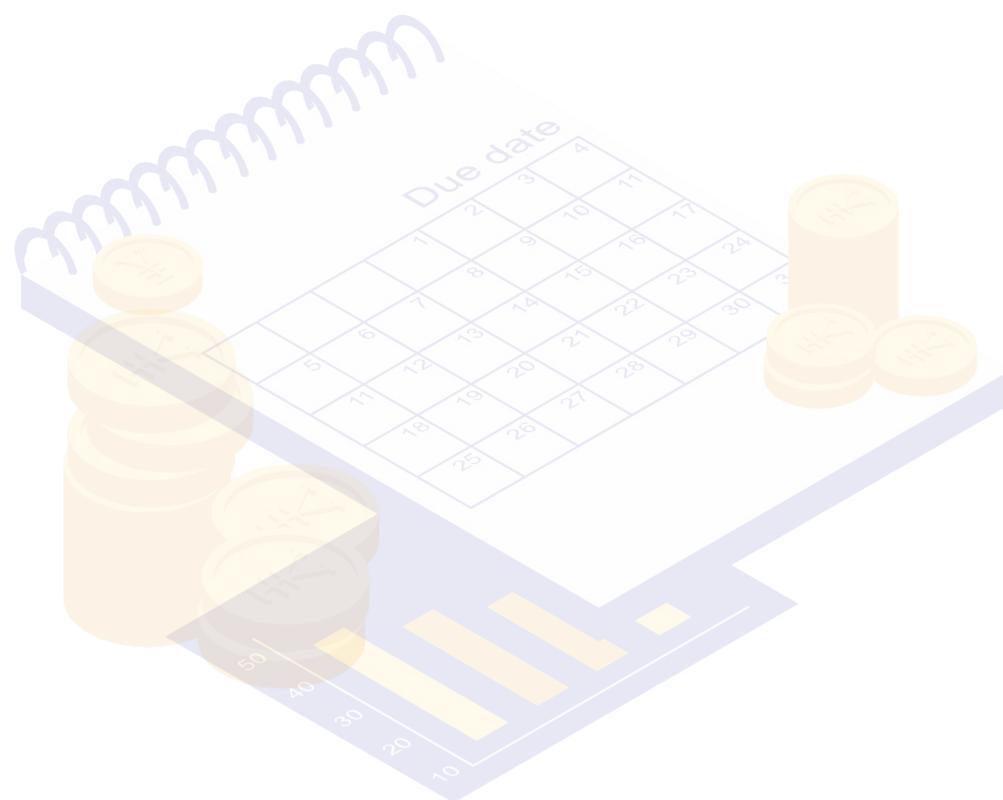


POLITÉCNICA

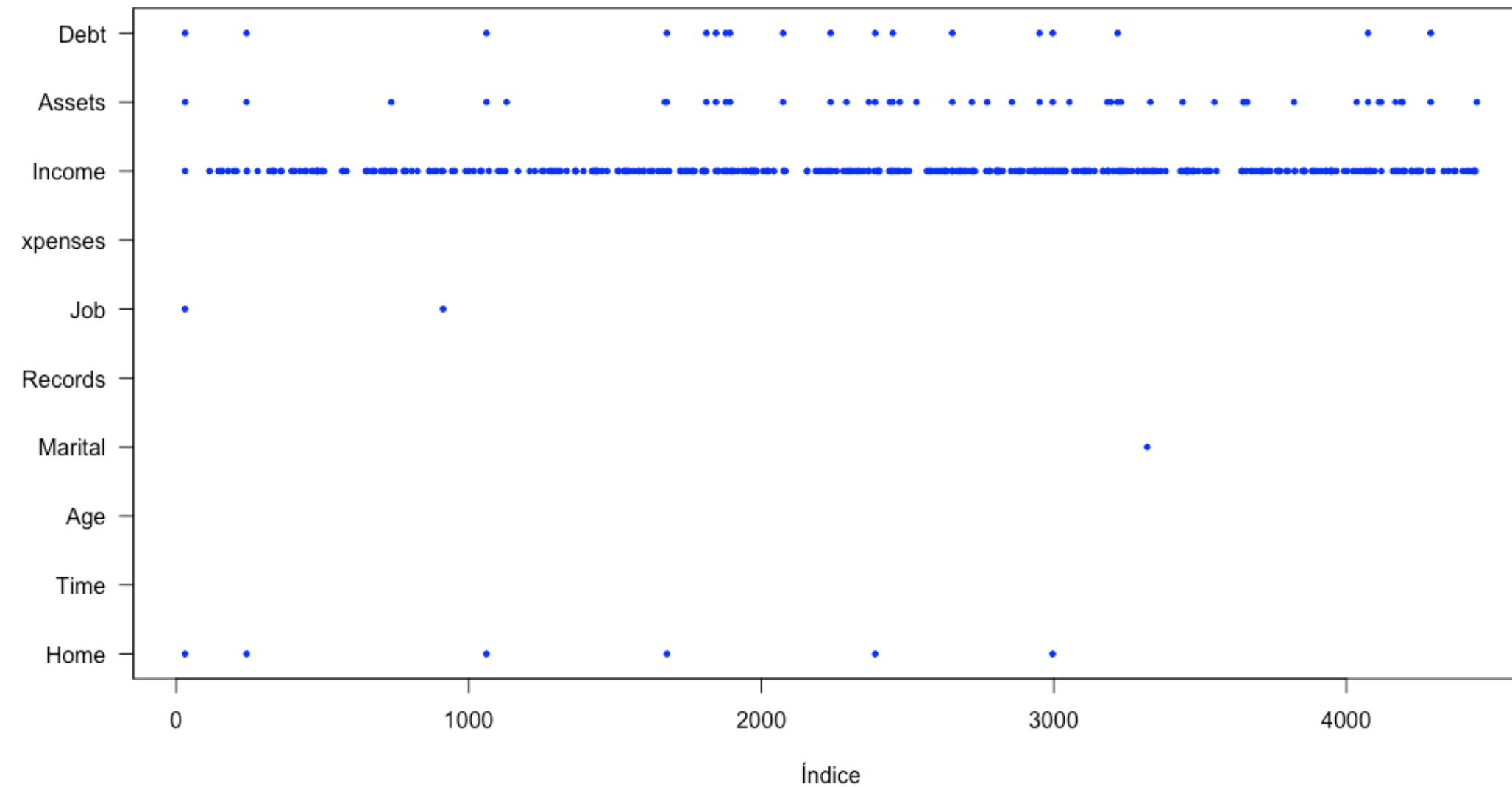
UNIVERSIDAD
POLÍTÉCNICA
DE MADRID

Paso 1: Estudio de valores nulos

El número de valores nulos es:
455



Valores nulos en la base de datos



Paso 1: Estudio de valores nulos

El número de valores nulos es:

455

Veamos cuantos valores nulos tiene cada instancia de la base de datos

Se decide eliminar las entradas que tengan más de un valor nulo

```
# A tibble: 415 x 2
  Fila Cantidad_NA
  <int>      <int>
1     1          5
2     2          1
3     3          1
4     4          1
5     5          1
6     6          1
7     7          1
8     8          1
9     9          3
10    10         1
# i 405 more rows
# i Use `print(n = ...)` to see more rows
```



UNIVERSIDAD
COMPLUTENSE
MADRID

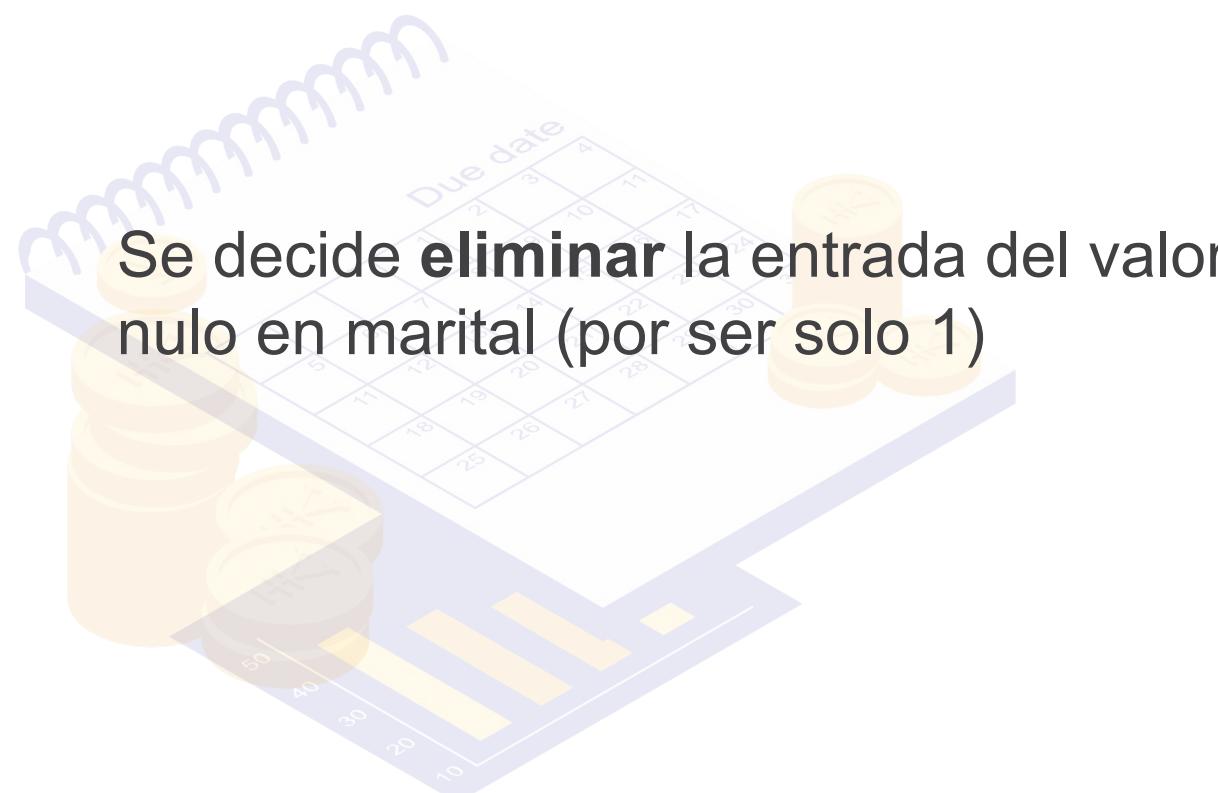


POLITÉCNICA

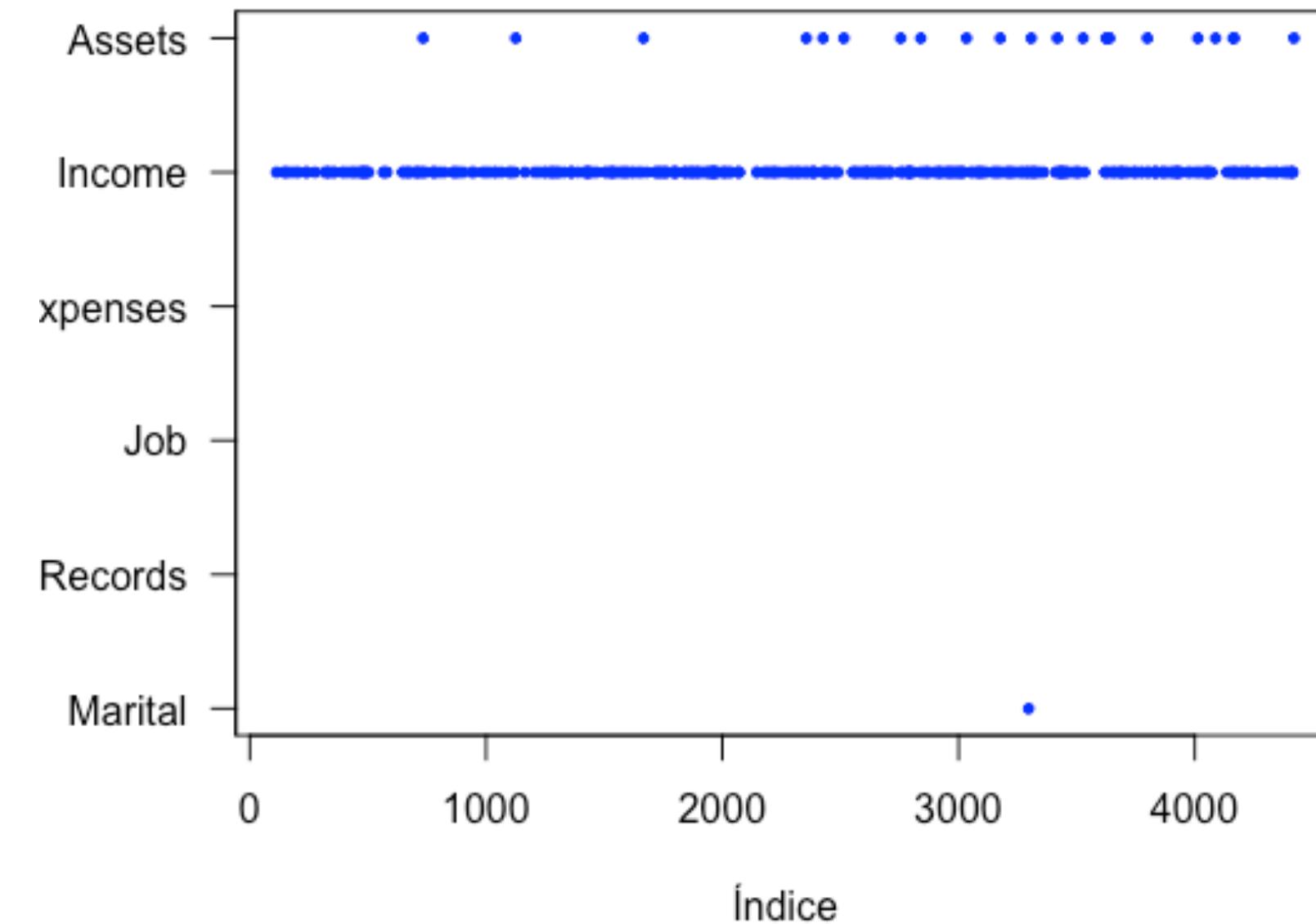
UNIVERSIDAD
POLITÉCNICA
DE MADRID

Paso 1: Estudio de valores nulos

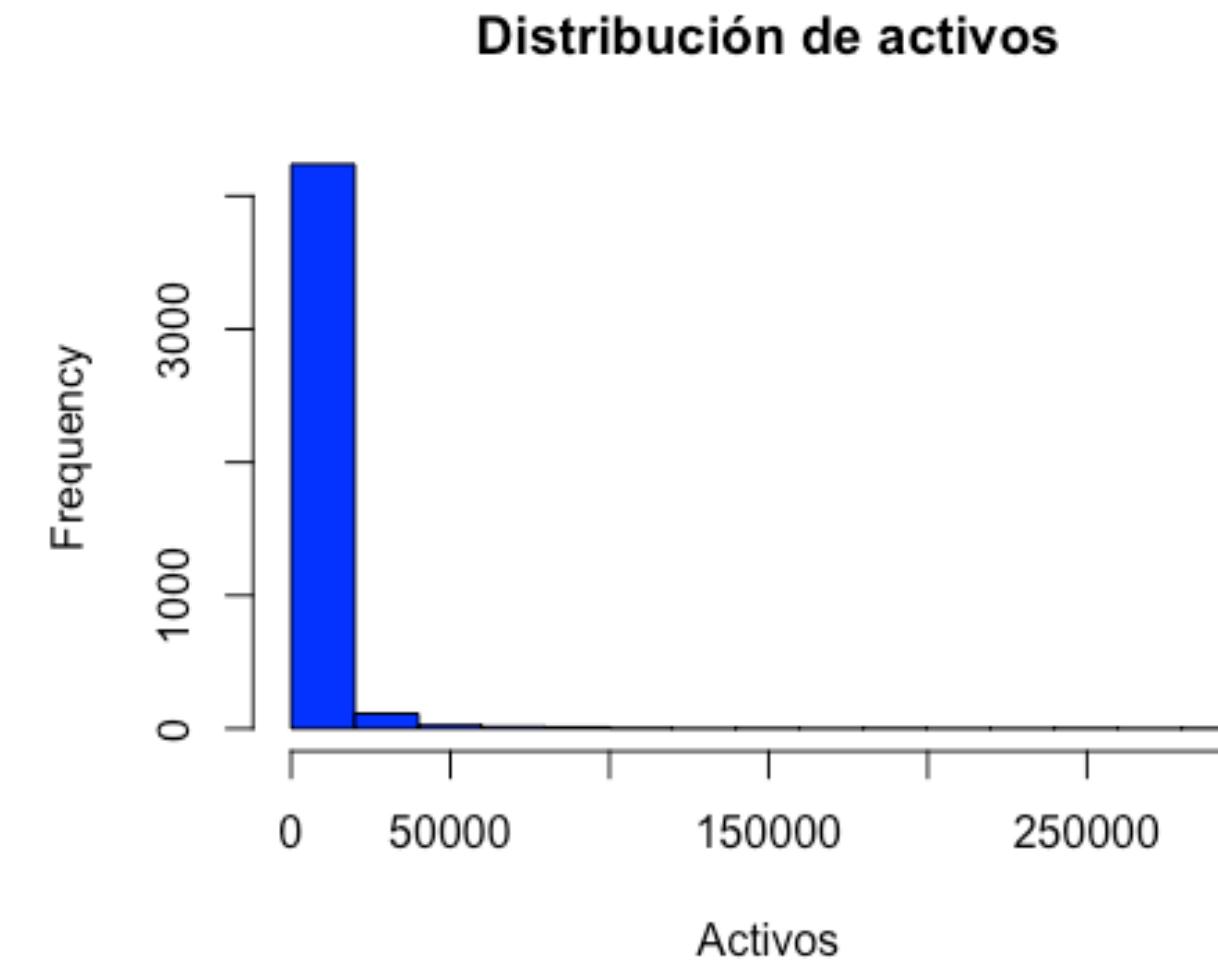
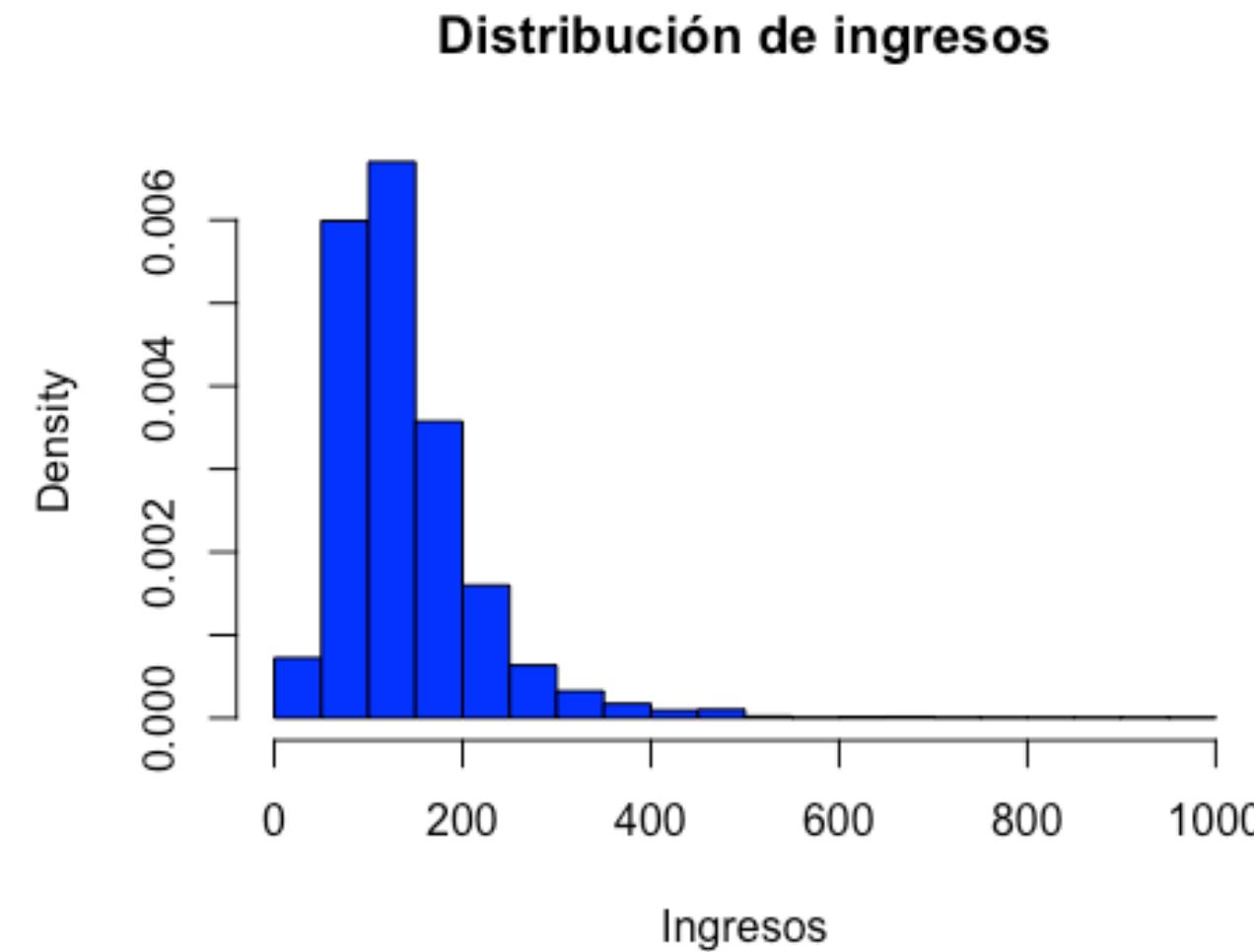
Tras eliminarlos, nos quedamos con estos nulos



Valores nulos en la base de datos



Paso 1: Estudio de valores nulos

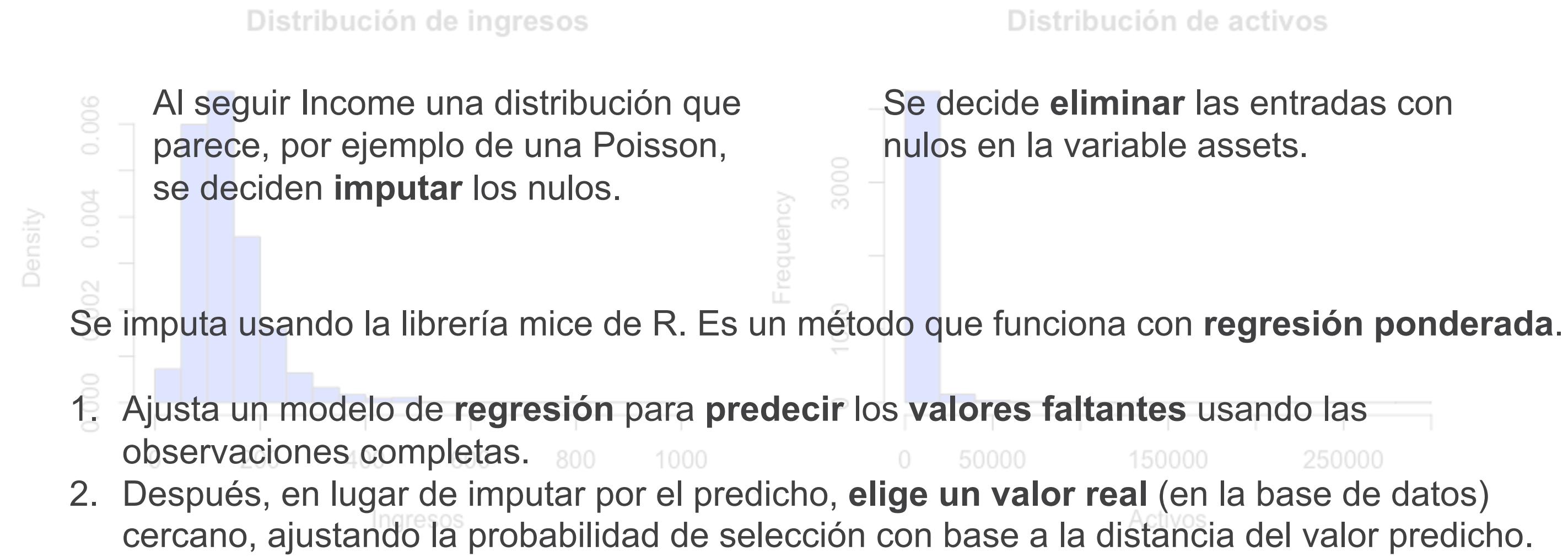


Al seguir Income una distribución que parece, por ejemplo de una Poisson, se deciden **imputar** los nulos.

Se decide **eliminar** las entradas con nulos en la variable assets.



Paso 1: Estudio de valores nulos

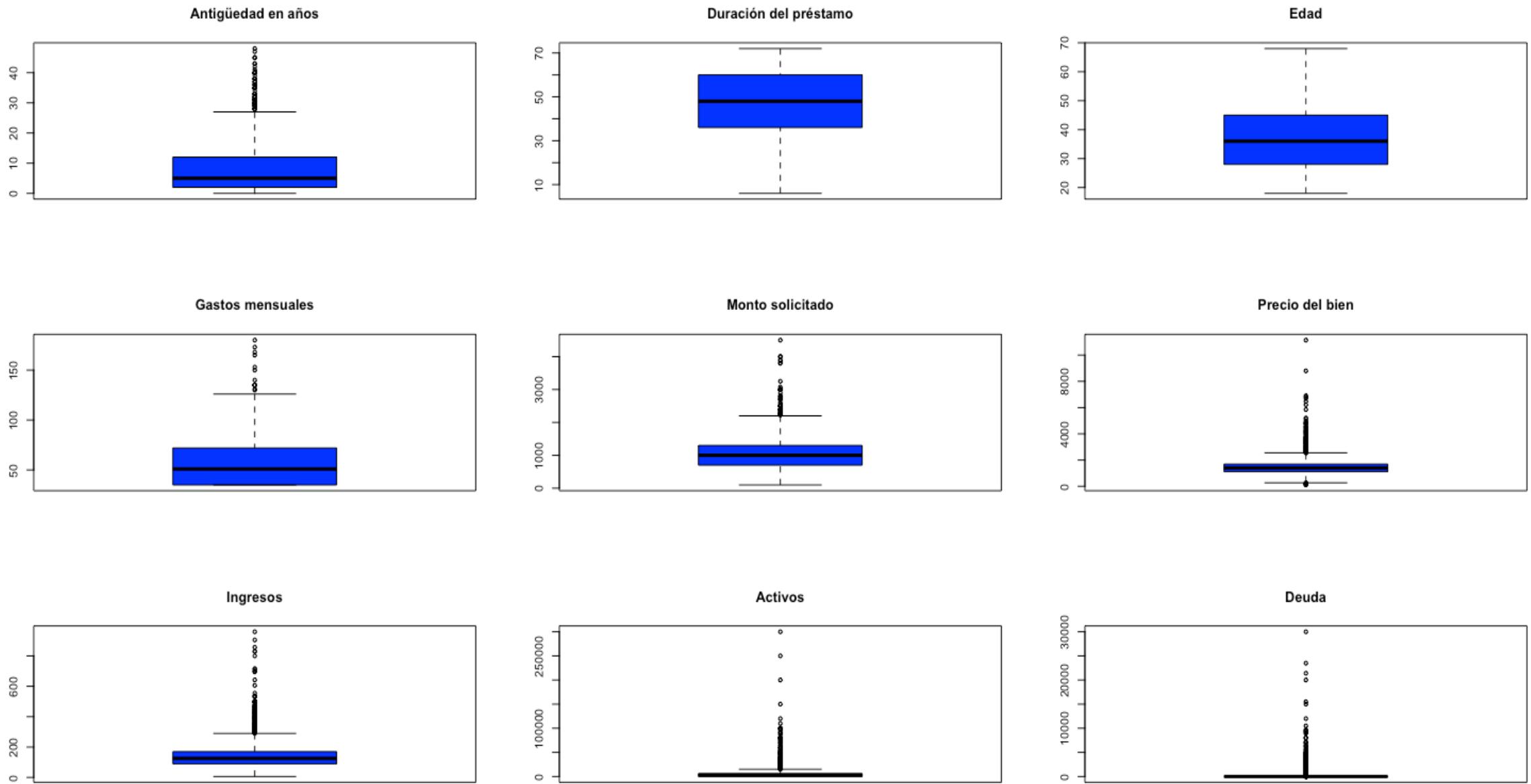


Paso 2: Estudio de valores atípicos

Para ver los valores atípicos, se usarán **diagramas de caja y bigote**.

A excepción de en **duración del préstamo** y en **edad**, tenemos bastantes valores atípicos en el resto de variables.

De momento los dejamos así. Eliminarlos podría **quitar información relevante** para el análisis, como por ejemplo, clientes con mucha deuda.



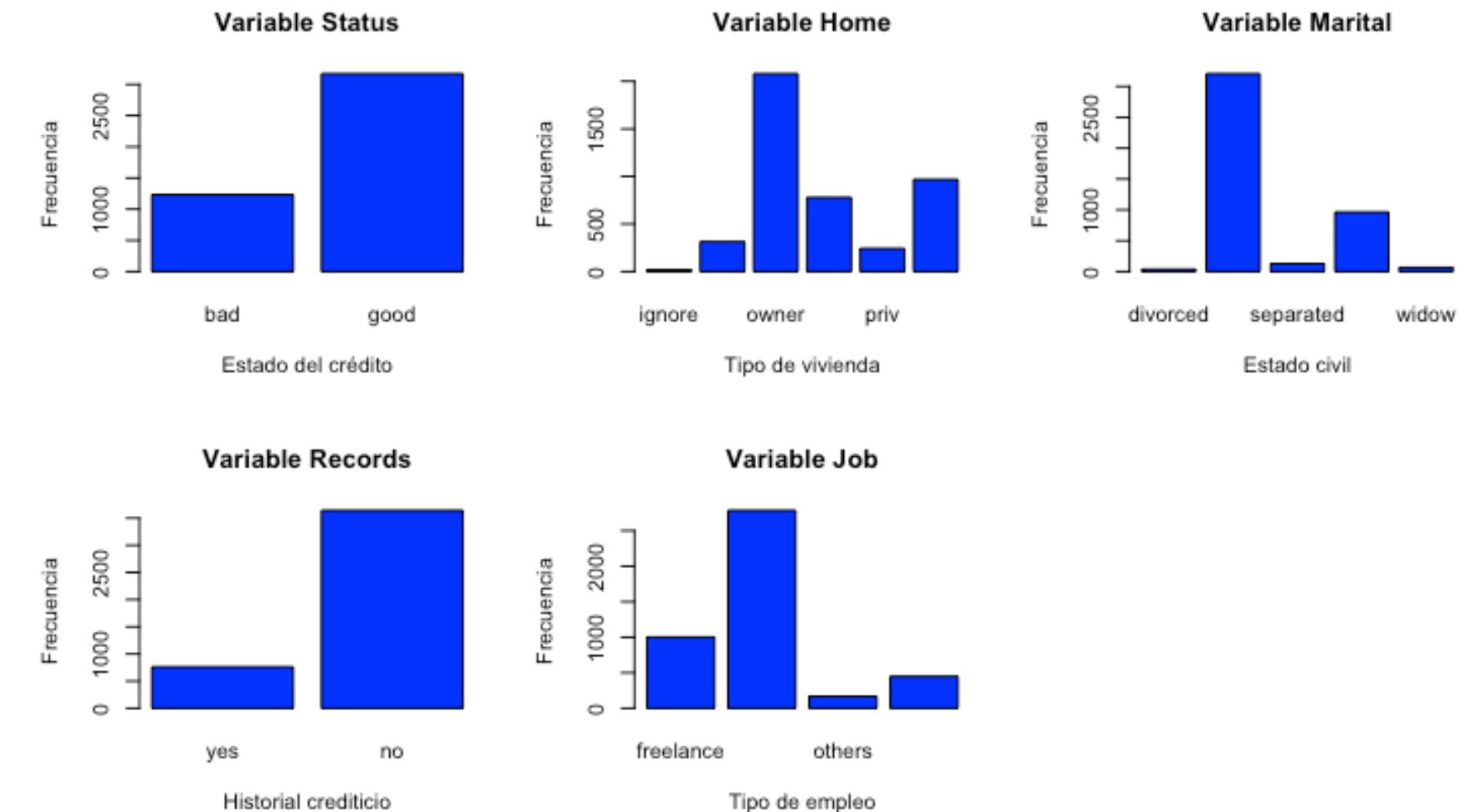
Paso 3: Balanceo de clases categóricas

Veamos los **diagramas de barras** de las clases categóricas

Podemos ver que todas las clases están desbalanceadas.

En concreto, la variable Status cuenta con:

bad	good
0.2799092	0.7200908

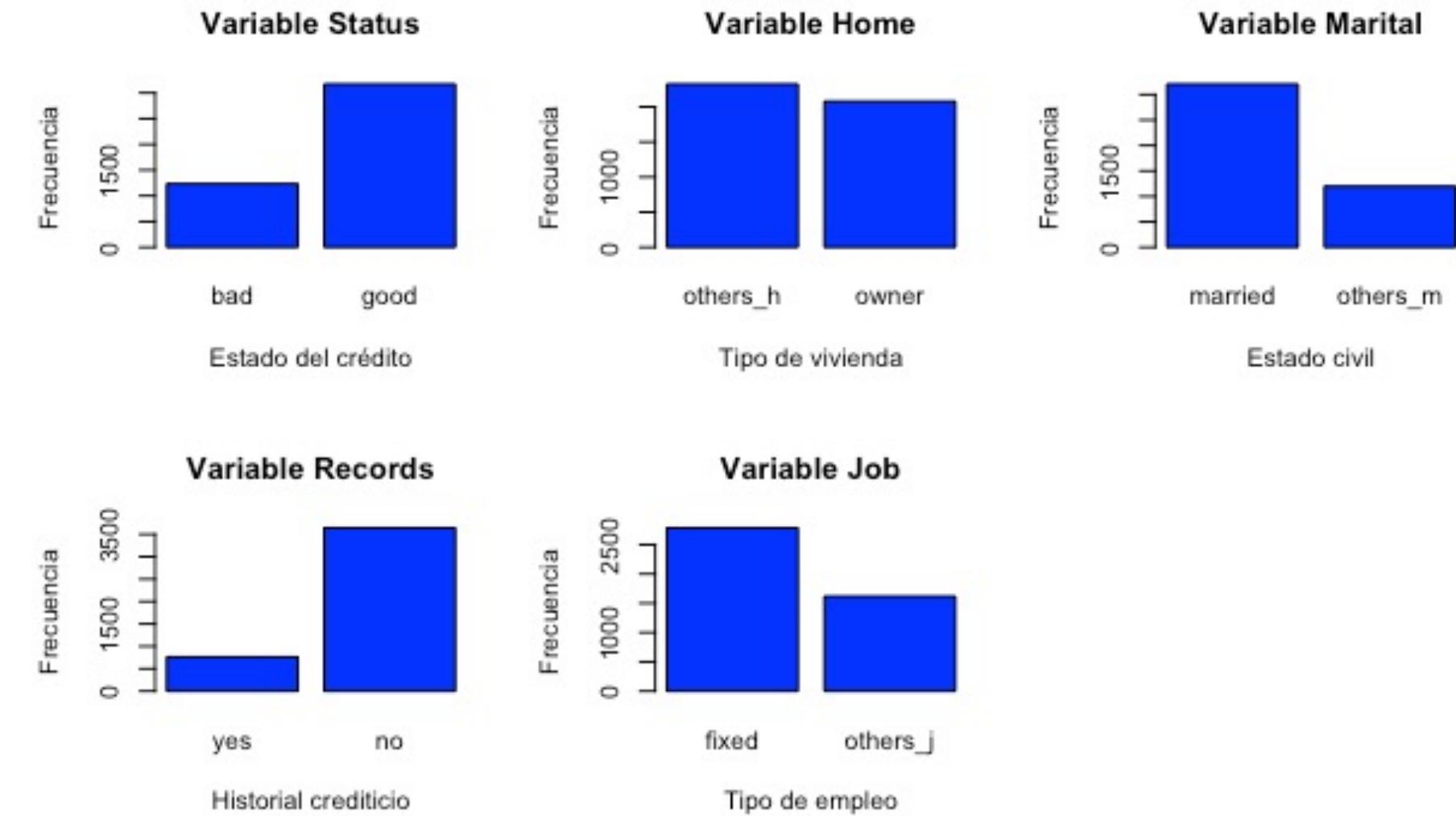


Paso 3: Balanceo de clases categóricas

El desbalanceo de la variable **status** lo trataremos más adelante en la base de datos

Para intentar mitigar el desbalanceo en el resto de variables de la base de datos, se dicotomizan, creando:

- **Home:** propietario/a y otros
- **Marital:** casado/a y otros
- **Job:** Empleo fijo y otros



Paso 3: Transformaciones de las variables

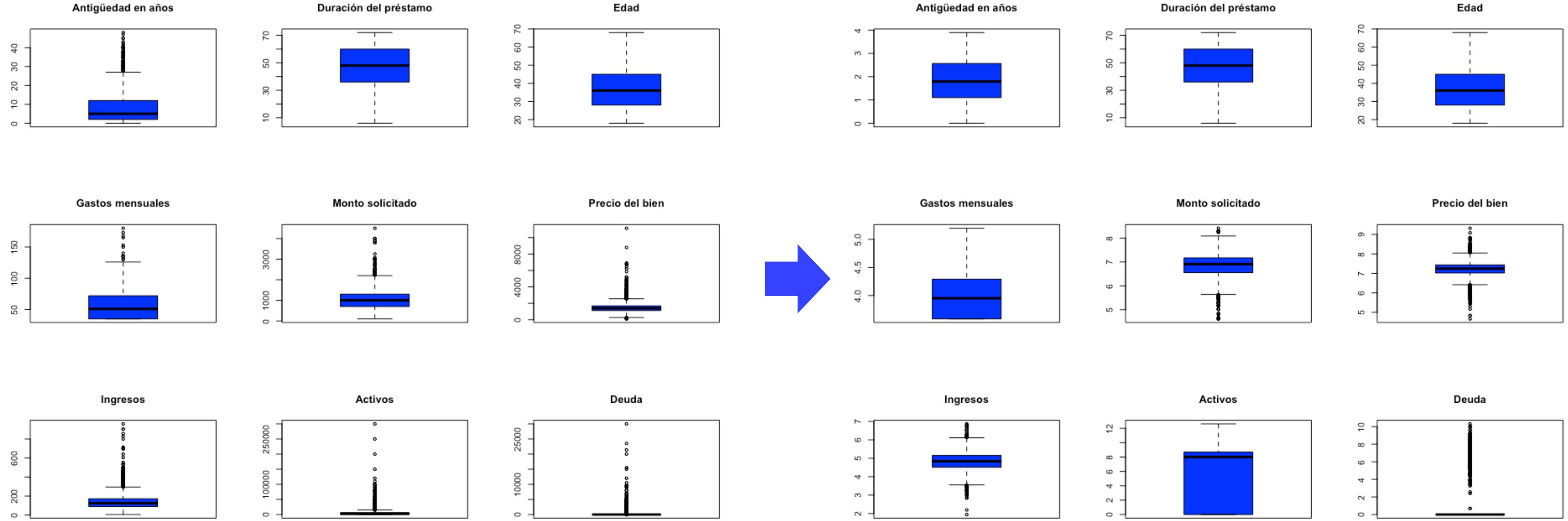
Para intentar mitigar los valores atípicos que veíamos en los diagramas de caja y bigote (y corregir el sesgo de dichas variables). Se aplica la siguiente transformación:

$$y = \log(1 + x)$$

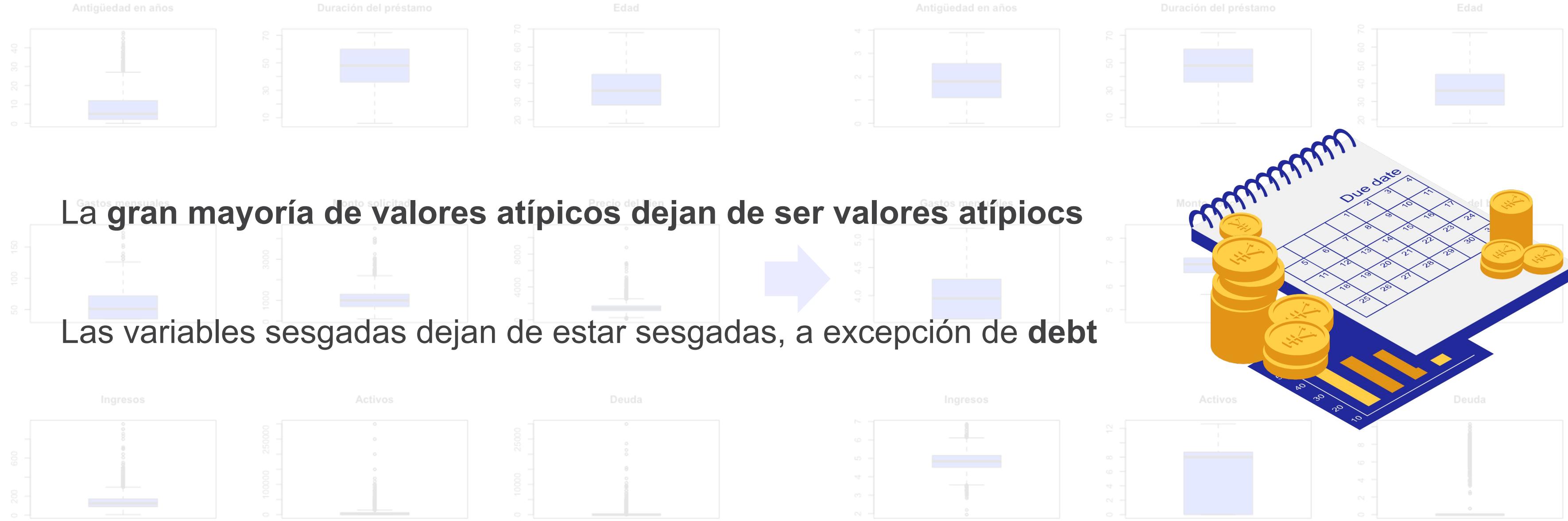
Se aplica a las variables: Seniority, Expenses, Amount, Price, Income, Assets y Debt



Paso 3: Transformaciones de las variables



Paso 3: Transformaciones de las variables



Paso 3: Transformaciones de las variables

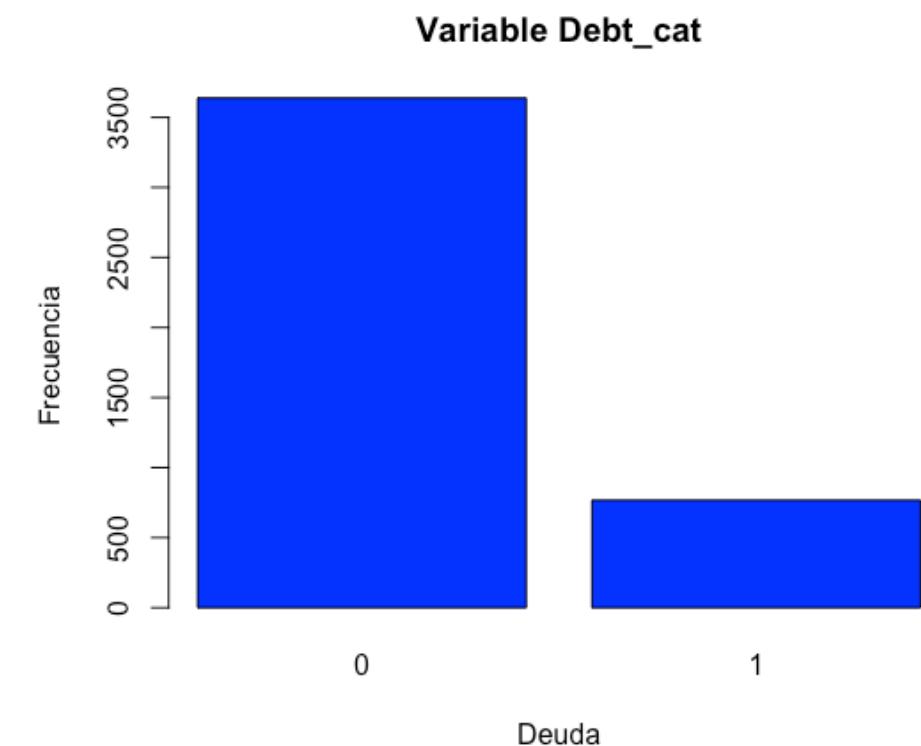
La gran mayoría de valores atípicos dejan de ser valores atípicos

Las variables sesgadas dejan de estar sesgadas, a excepción de **debt**

Vamos a categorizar la variable **debt**

Si la deuda es 0 le damos el valor 0

Si la deuda > 0 le damos el valor 1



Tarea 1:

Tablas de Contingencia



Tarea 1: Tablas de Contingencia

- 1.1 Plantear el objetivo del estudio
Construir las tablas de contingencia

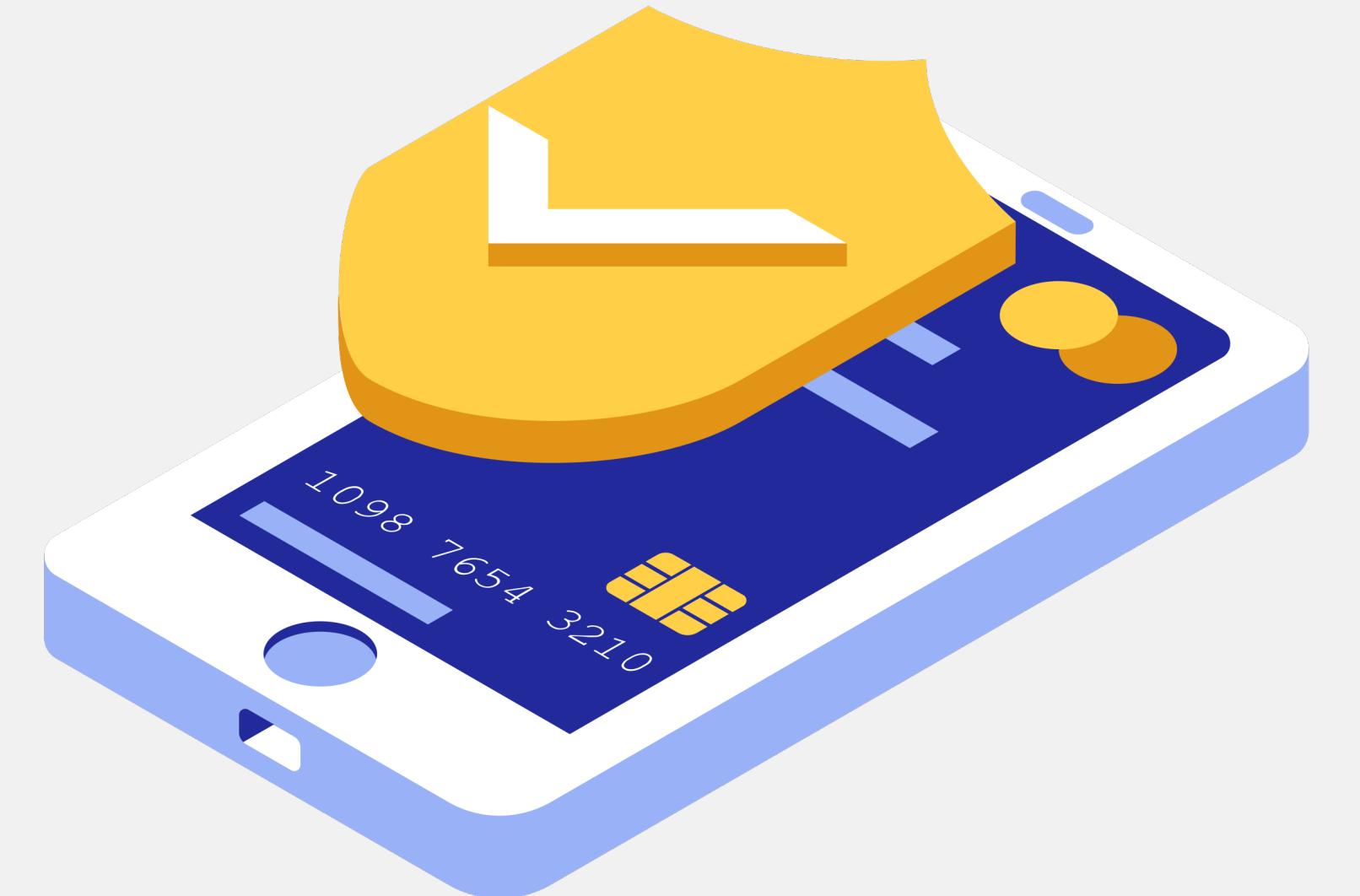
- 1.2 Estudiar la posible influencia de variables de interacción o confusión

- 1.3 Analizar la hipótesis de independencia entre variables



Objetivo del estudio

El objetivo del estudio es **analizar si la variable Status** (estado del crédito) **está relacionada** con las **variables Home, Marital, Records, Job y Debt_cat**. Entender estas relaciones puede ser fundamental para saber si conceder un crédito a una persona o no.



Tablas de contingencia

	Status	
Home	bad	good
others_h	855	1470
owner	378	1702

	Status	
Marital	bad	good
married	816	2391
others_m	417	781

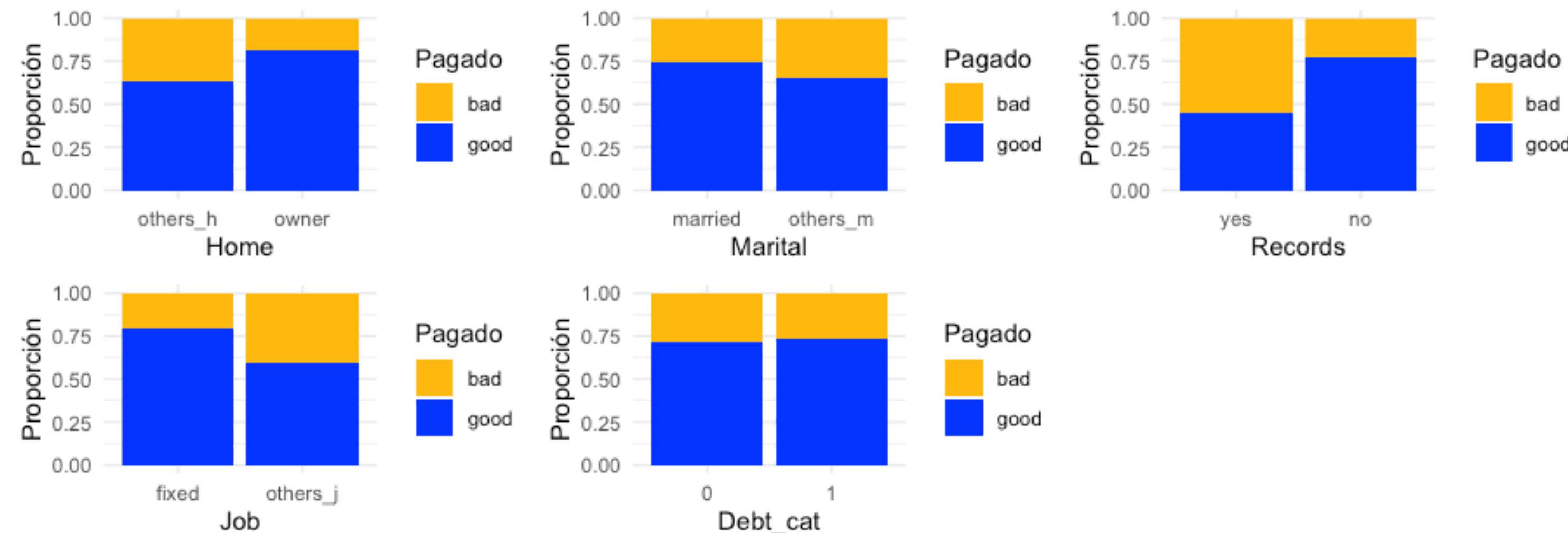
	Status	
Records	bad	good
yes	422	340
no	811	2832

	Status	
Job	bad	good
fixed	578	2208
others_j	655	964

	Status	
Debt_cat	bad	good
0	1032	2606
1	201	566



Tablas de contingencia



Tablas de contingencia

Vamos a analizar ahora las medidas de asociación entre la variable **Status** y las variables anteriores

- Diferencia de proporciones
 - Riesgo relativo
- Odds ratio (diferencia de plausibilidad)



Tablas de contingencia

Vamos a analizar ahora las medidas de asociación entre la variable **Status** y las variables anteriores

Variable	DR (IC 95%)	RR (IC 95%)	OR (IC 95%)
Home	0.186 (0.160, 0.212)	2.024 (1.821, 2.249)	2.619 (2.277, 3.012)
Marital	-0.094 (-0.125, -0.063)	0.731 (0.663, 0.806)	0.639 (0.554, 0.737)
Records	0.331 (0.293, 0.369)	2.488 (2.278, 2.717)	4.334 (3.683, 5.100)
Job	-0.197 (-0.225, -0.169)	0.513 (0.467, 0.563)	0.385 (0.337, 0.441)
Debt_cat	0.022 (-0.013, 0.056)	1.082 (0.951, 1.232)	1.115 (0.935, 1.330)



Tablas de contingencia

Vamos a analizar ahora las medidas de asociación entre la variable **Status** y las variables anteriores

Medida	¿Qué indica?
DR (Diferencia de Proporciones)	Diferencia entre la proporción de <i>impagos</i> (bad) en un grupo frente al otro.
RR (Relative Risk)	Cuántas veces es más probable (o menos) que ocurra el impago en un grupo frente al otro.
OR (Odds Ratio)	Relación entre las odds de impago en un grupo respecto al otro.



Tablas de contingencia

Vamos a analizar ahora las medidas de asociación entre la variable **Status** y las variables anteriores

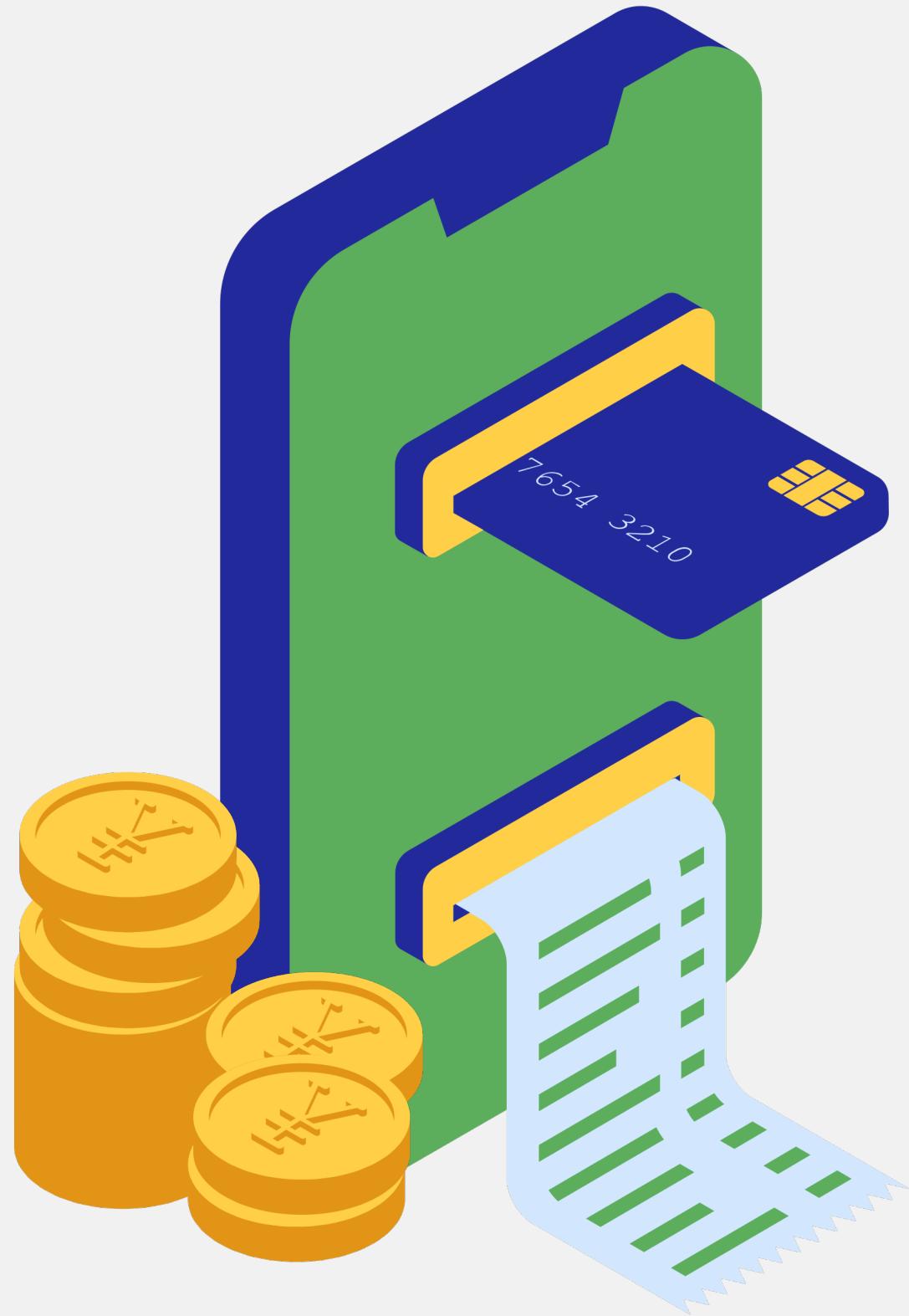
1. Home (tipo de vivienda)

- DR: 0.186 → 18.6% más impagos en el grupo "others_h" (no propietarios) que en "owner".
- RR: 2.02 → los que **no son propietarios** tienen **el doble de probabilidad de impago** que los propietarios.
- OR: 2.62 → las *odds* de impago son 2.6 veces mayores si **no eres propietario**.

3. Records (antecedentes crediticios negativos)

- DR: 0.331 → hay un **33.1% más de impagos entre quienes tienen antecedentes**.
- RR: 2.49 → riesgo de impago **es 2.5 veces mayor con antecedentes**.
- OR: 4.33 → odds de impago **se cuadriplican con historial negativo**.





Qué conclusiones sacamos?

- No ser propietario está fuertemente asociado al **impago**.
- Estar **casado** se asocia con **menor riesgo de impago**.
- Tener **antecedentes crediticios negativos** es el factor más asociado al **impago**.
- Tener **trabajo fijo** reduce notablemente el **riesgo de impago**.
- Tener **deuda previa** no muestra asociación estadísticamente significativa con el **impago**.



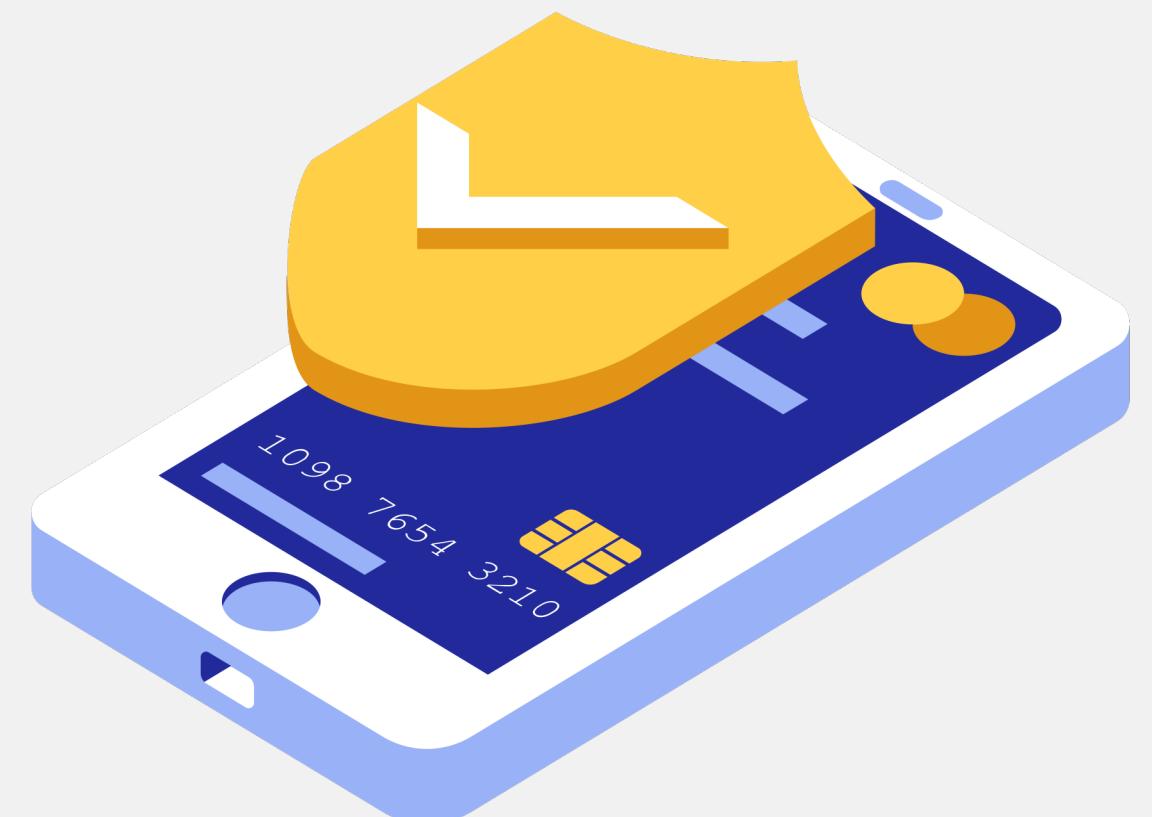
Influencia (confusión o interacción)

Variable de confusión

Por ejemplo, si parece que los que tienen Records tienen más impagos (Status = bad), pero ese efecto desaparece cuando controlamos por Job, entonces Job es una variable de confusión.

Variable de interacción

Por ejemplo, si el efecto de Records sobre Status es mucho más fuerte en personas sin trabajo fijo, entonces Job es un modificador de efecto.

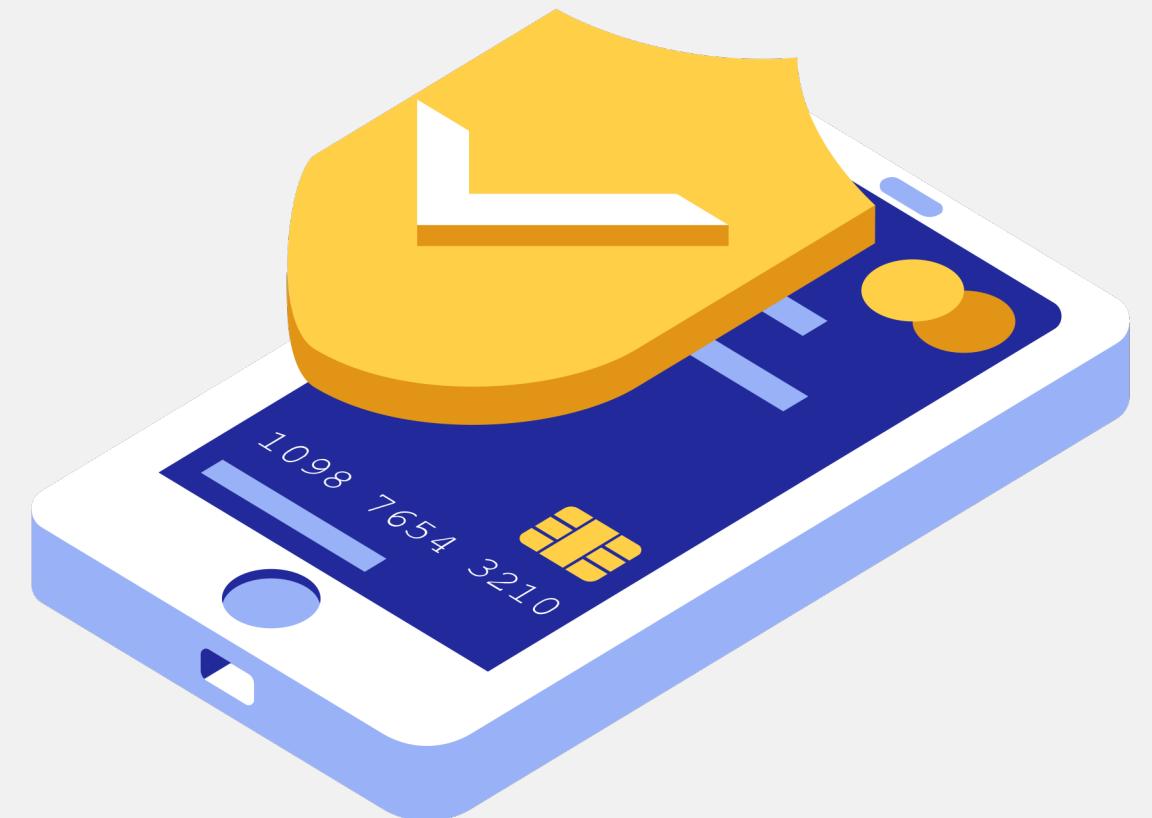


Influencia (confusión o interacción)

Variable de confusión

Variable de interacción

Hagamos tablas de contingencia
estratificadas por la variable que
creemos que puede ser de
confusión o de interacción.



UNIVERSIDAD
COMPLUTENSE
MADRID



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID

Efecto del historial crediticio según tipo de empleo

- En personas **con trabajo fijo**, tener antecedentes negativos **multiplica por 5,38** las odds de impago (IC 95%: 4,34 – 6,67).
- En personas **sin trabajo fijo**, las odds también aumentan significativamente, pero en menor medida: **OR = 3,48** (IC 95%: 2,68 – 4,52).

También se observa un mayor **riesgo relativo (RR)** y **diferencia de proporciones (DR)** en el grupo con empleo fijo.



El **tipo de empleo modifica** el efecto del historial crediticio sobre el impago.

La variable **Job actúa como modificador de efecto**, lo que **sugiere una interacción** entre ambas variables.

, , Job = fixed

Status

Records	bad	good
yes	225	234
no	353	1974

, , Job = others_j

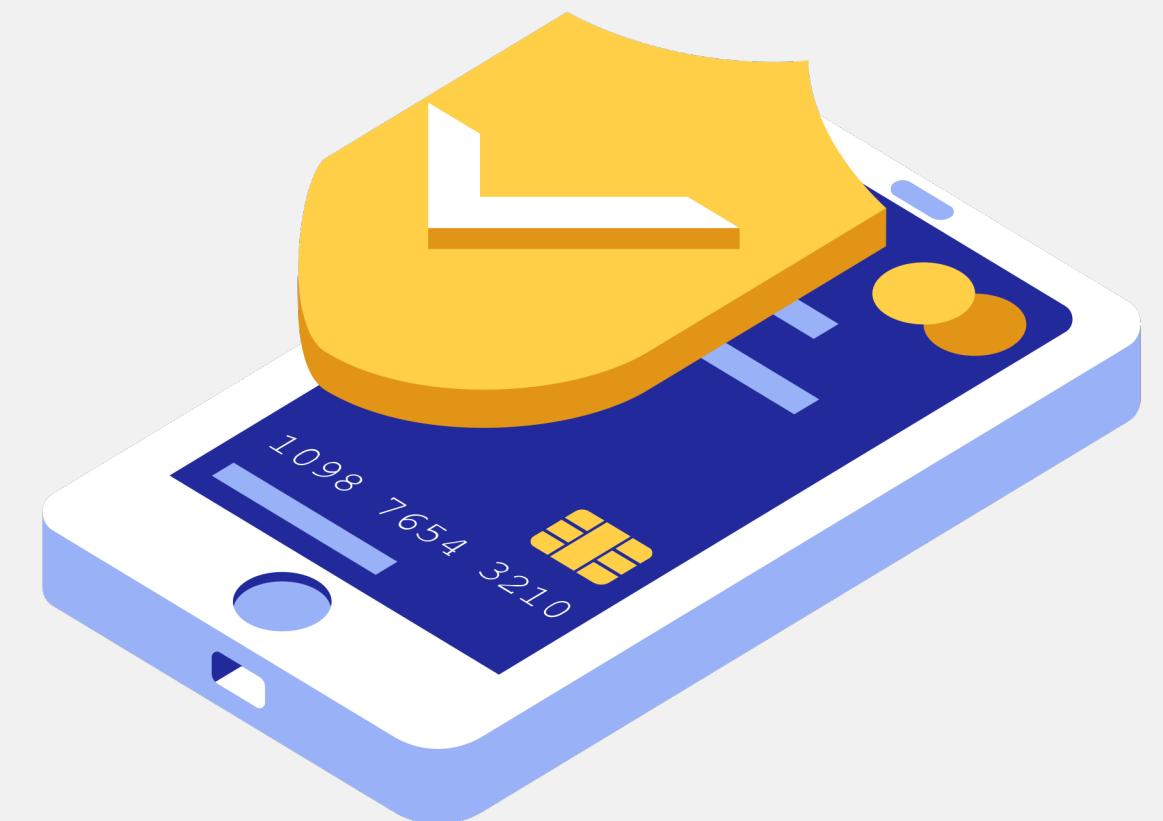
Status

Records	bad	good
yes	197	106
no	458	858



Efecto del historial crediticio según el estado civil (marital)

Al hacer el mismo procedimiento con la variable marital, vemos que las diferencias de las medidas de asociación entre los estratos son pequeñas. Esto indica que la variable marital no modifica sustancialmente el efecto de Records sobre el estado del crédito.



Hipótesis de independencia entre variables

Comprobemos si las variables son o no independientes de la variable **Status** mediante el test de la Chi-Cuadrado

El contraste es el siguiente:

$$\begin{cases} H_0: \text{Las variables son independientes.} \\ H_1: \text{Hay asociación entre las variables (no son independientes).} \end{cases}$$



Hipótesis de independencia entre variables

Comprobemos si las variables son o no independientes de la variable **Status** mediante el test de la Chi-Cuadrado

Variable	X ²	gl	p-valor	Decisión sobre H ₀
Home	188.46	1	< 0.001	Se rechaza H ₀ → Hay asociación
Marital	37.94	1	< 0.001	Se rechaza H ₀ → Hay asociación
Records	342.93	1	< 0.001	Se rechaza H ₀ → Hay asociación
Job	197.37	1	< 0.001	Se rechaza H ₀ → Hay asociación
Debt_cat	1.47	1	0.226	No se rechaza H ₀ → No hay asociación



Tarea 2: Regresión logística



Tarea 2: Regresión logística

- 2.1 Establecer una pregunta de investigación
- 2.2 Explicar si el modelo que se aplicará será explicativo o predictivo
- 2.3 Aplicar el modelo de regresión logística, indicando el procedimiento para la elección de variables. Interpretar los resultados.



Pregunta de investigación

El objetivo de este análisis es:

predecir si un crédito será pagado o no,

es decir, si el cliente caerá en impago,
utilizando como variable respuesta Status
(binaria: "good" o "bad").



¿Modelo explicativo o predictivo?

Se aplica un **modelo de regresión logística predictivo**, que permite:

- Estimar la **probabilidad de impago** (Status = "bad")
- **Clasificar** a los clientes según su nivel de riesgo
- Evaluar el rendimiento con métricas como la **matriz de confusión** y el **AUC-ROC**

Aunque el modelo también permite interpretar el efecto de cada variable, **el foco principal está en su capacidad de predicción.**



Modelo de regression logística

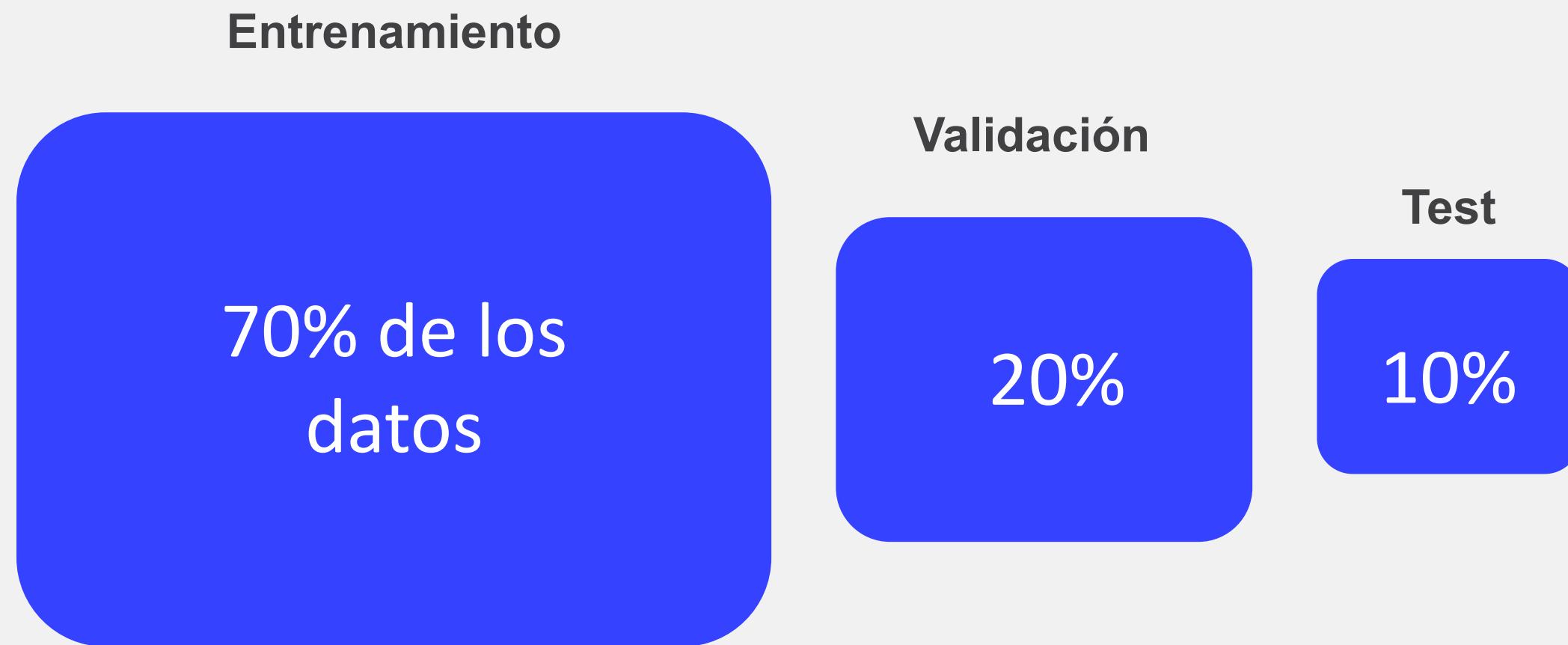
1 - Tratamiento de los datos:

100% de los
datos



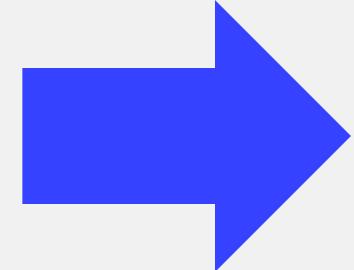
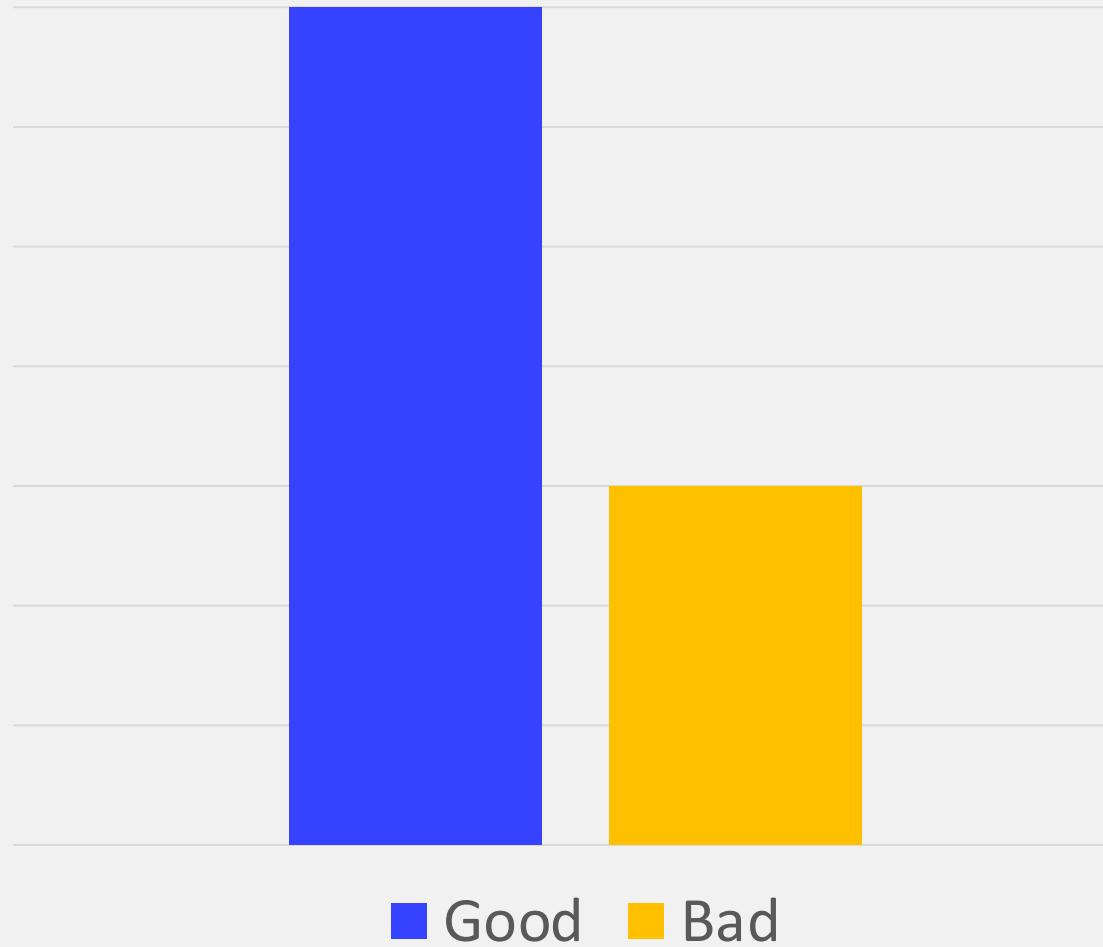
Modelo de regression logística

1 - Tratamiento de los datos:

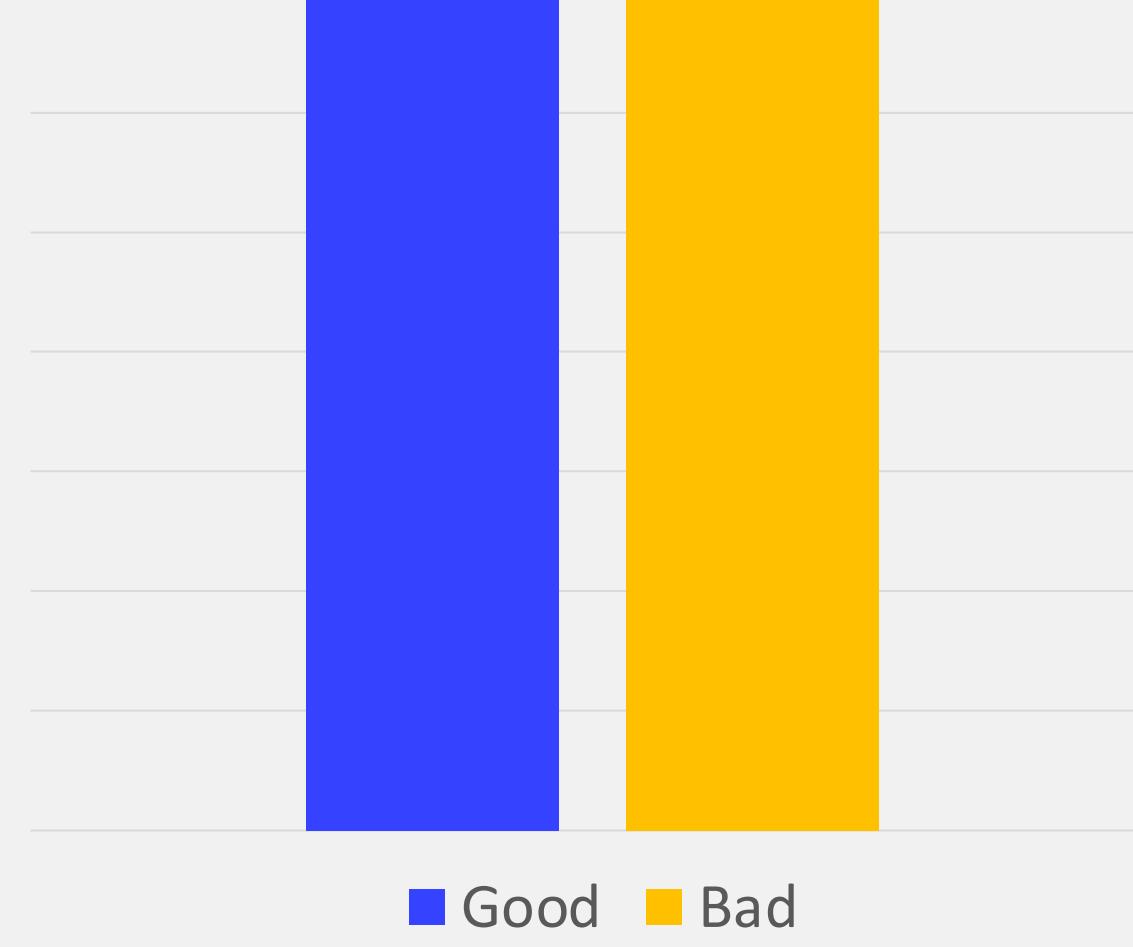


Modelo de regression logística

2 – Balanceo de la clase a predecir



sobremuestreo



Modelo de regression logística

3 – Ajuste del modelo:

En primer lugar probamos con los modelos univariantes

Variable	p_value
Home	0.0000
Marital	0.0000
Records	0.0000
Job	0.0000
log_Seniority	0.0000
Time	0.0000
Age	0.0000
log_Income	0.0000
log_Assets	0.0000
log_Amount	0.0000
log_Price	0.0001
log_Expenses	0.1332



Modelo de regression logística

3 – Ajuste del modelo:

Ajustamos ahora el modelo máximo de regression logística

Tiene un AIC de 4525.7

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.9299606	0.8696698	2.219	0.02647 *
Homeowner	-0.3349693	0.1065139	-3.145	0.00166 **
Maritalothers_m	0.2916809	0.1057018	2.759	0.00579 **
Recordsno	-1.8080471	0.0989303	-18.276	< 2e-16 ***
Jobothers_j	0.8887583	0.0796896	11.153	< 2e-16 ***
log_Seniority	-0.6218989	0.0426286	-14.589	< 2e-16 ***
Time	0.0106054	0.0032852	3.228	0.00125 **
Age	0.0003043	0.0039898	0.076	0.93920
log_Expenses	1.0338707	0.1416147	7.301	2.87e-13 ***
log_Income	-1.0830893	0.0826626	-13.103	< 2e-16 ***
log_Assets	-0.0681881	0.0122592	-5.562	2.66e-08 ***
log_Amount	1.4568601	0.1384410	10.523	< 2e-16 ***
log_Price	-1.2430768	0.1422027	-8.742	< 2e-16 ***

Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’
	0.1 ‘ ’	1		



Modelo de regression logística

3 – Ajuste del modelo:

Mediante stepAIC llegamos al siguiente modelo

Tiene un AIC de 4523.7

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.935784	0.866303	2.235	0.02545 *
Homeowner	-0.334157	0.105982	-3.153	0.00162 **
Maritalothers_m	0.290454	0.104466	2.780	0.00543 **
Recordsno	-1.808843	0.098385	-18.385	< 2e-16 ***
Jobothers_j	0.889661	0.078809	11.289	< 2e-16 ***
log_Seniority	-0.621016	0.041028	-15.137	< 2e-16 ***
Time	0.010596	0.003283	3.228	0.00125 **
log_Expenses	1.034814	0.141071	7.335	2.21e-13 ***
log_Income	-1.083080	0.082660	-13.103	< 2e-16 ***
log_Assets	-0.068127	0.012233	-5.569	2.56e-08 ***
log_Amount	1.457050	0.138412	10.527	< 2e-16 ***
log_Price	-1.243202	0.142190	-8.743	< 2e-16 ***

Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’
	0.1 ‘ ’	1		



Modelo de regression logística

3 – Ajuste del modelo:

Probamos ahora modelos con interacciones.

El modelo completo es

Tiene un AIC de 4506.3

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-11.215811	4.139695	-2.709	0.006742 **
Homeowner	-0.484300	0.113184	-4.279	1.88e-05 ***
Maritalothers_m	0.141203	0.118833	1.188	0.234739
Recordsno	-1.919350	0.122236	-15.702	< 2e-16 ***
Jobothers_j	0.689521	0.180351	3.823	0.000132 ***
log_Seniority	-0.620322	0.042937	-14.447	< 2e-16 ***
Time	0.009880	0.003304	2.990	0.002788 **
Age	-0.001242	0.004020	-0.309	0.757387
log_Expenses	4.414475	1.036686	4.258	2.06e-05 ***
log_Income	1.676740	0.843109	1.989	0.046728 *
log_Assets	-0.069912	0.012304	-5.682	1.33e-08 ***
log_Amount	1.439161	0.138271	10.408	< 2e-16 ***
log_Price	-1.227258	0.142108	-8.636	< 2e-16 ***
Recordsno:Jobothers_j	0.257392	0.197545	1.303	0.192590
Homeowner:Maritalothers_m	0.638183	0.188317	3.389	0.000702 ***
log_Expenses:log_Income	-0.694847	0.210770	-3.297	0.000978 ***



Modelo de regression logística

3 – Ajuste del modelo:

Si aplicamos StepAIC llegamos al siguiente modelo

Tiene un AIC de 4506.3

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-11.394176	4.138913	-2.753	0.005906	**
Homeowner	-0.484030	0.112937	-4.286	1.82e-05	***
Maritalothers_m	0.148946	0.116984	1.273	0.202942	
Recordsno	-1.825936	0.098936	-18.456	< 2e-16	***
Jobothers_j	0.897956	0.079201	11.338	< 2e-16	***
log_Seniority	-0.627378	0.041246	-15.211	< 2e-16	***
Time	0.010177	0.003298	3.086	0.002029	**
log_Exenses	4.439189	1.036398	4.283	1.84e-05	***
log_Income	1.697088	0.843351	2.012	0.044187	*
log_Assets	-0.070141	0.012275	-5.714	1.10e-08	***
log_Amount	1.431408	0.138352	10.346	< 2e-16	***
log_Price	-1.224455	0.142345	-8.602	< 2e-16	***
Homeowner:Maritalothers_m	0.635440	0.187570	3.388	0.000705	***
log_Exenses:log_Income	-0.700397	0.210782	-3.323	0.000891	***



Modelos *sin* interacciones

=====

Modelo Maximal
AUC: 0.8264

Reference	
Prediction	bad good
bad	198 174
good	53 456

Accuracy: 0.7423
Sensibilidad (Recall bad): 0.7888
Especificidad: 0.7238
Precisión (PPV): 0.5323

=====

Modelo StepAIC
AUC: 0.8264

Reference	
Prediction	bad good
bad	198 174
good	53 456

Accuracy: 0.7423
Sensibilidad (Recall bad): 0.7888
Especificidad: 0.7238
Precisión (PPV): 0.5323

=====

Modelos *con* interacciones

=====

Modelo Interacciones Max
AUC: 0.8286

Reference	
Prediction	bad good
bad	198 175
good	53 455

Accuracy: 0.7412
Sensibilidad (Recall bad): 0.7888
Especificidad: 0.7222
Precisión (PPV): 0.5308

=====

Modelo Interacciones Step
AUC: 0.8289

Reference	
Prediction	bad good
bad	197 172
good	54 458

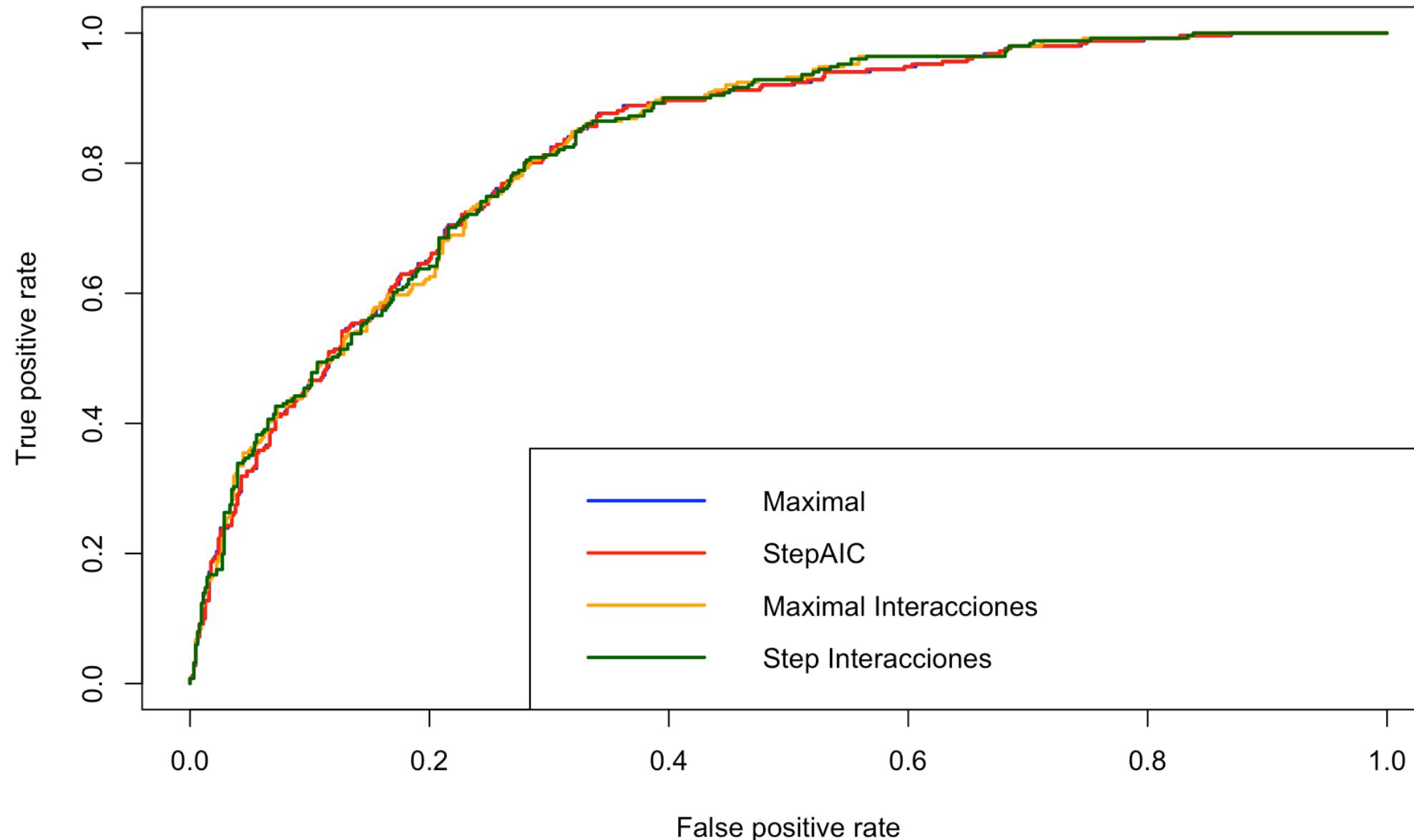
Accuracy: 0.7435
Sensibilidad (Recall bad): 0.7849
Especificidad: 0.727
Precisión (PPV): 0.5339

=====



Curvas ROC

Curvas ROC - Validación



UNIVERSIDAD
COMPLUTENSE
MADRID

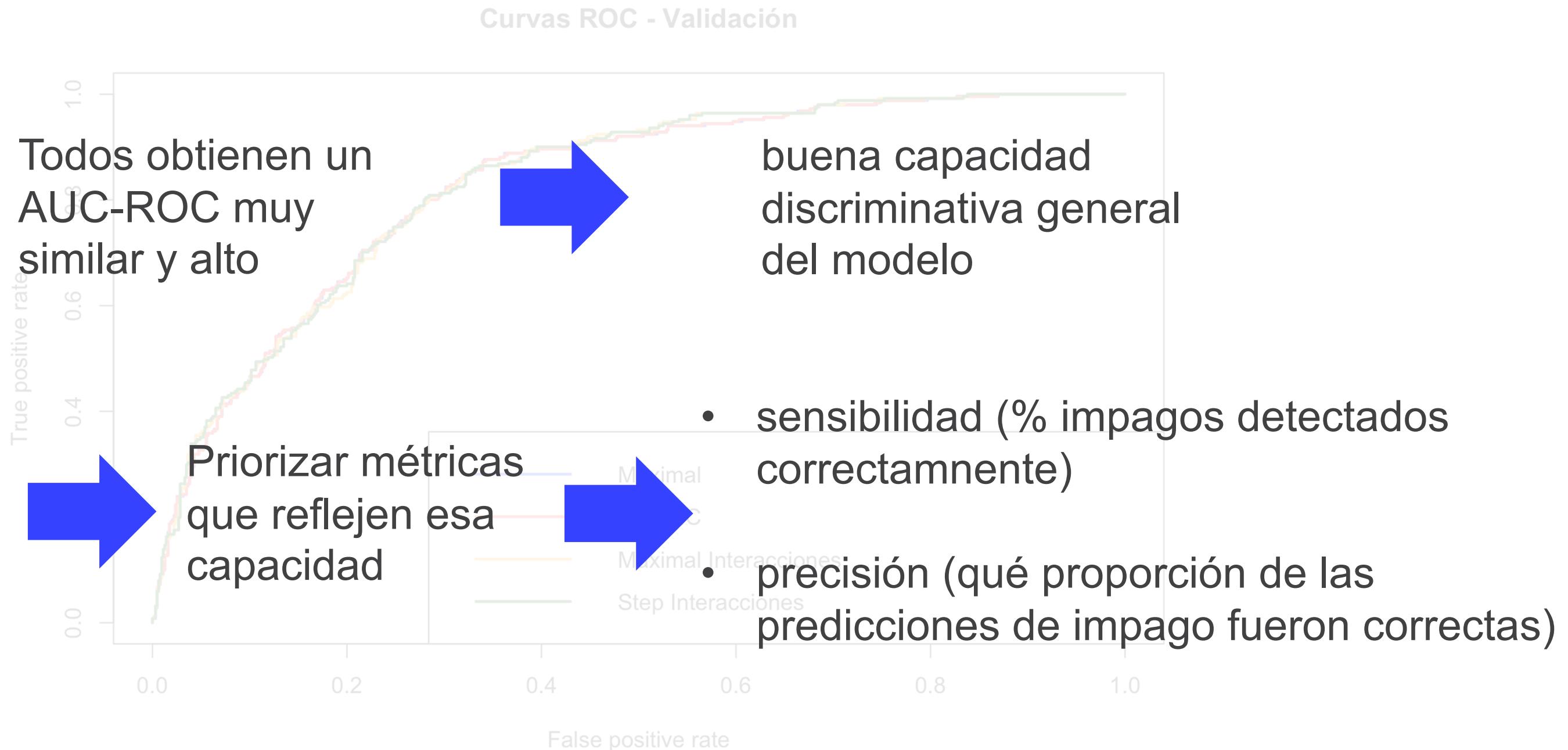


UNIVERSIDAD
POLÍTÉCNICA
DE MADRID

POLITÉCNICA

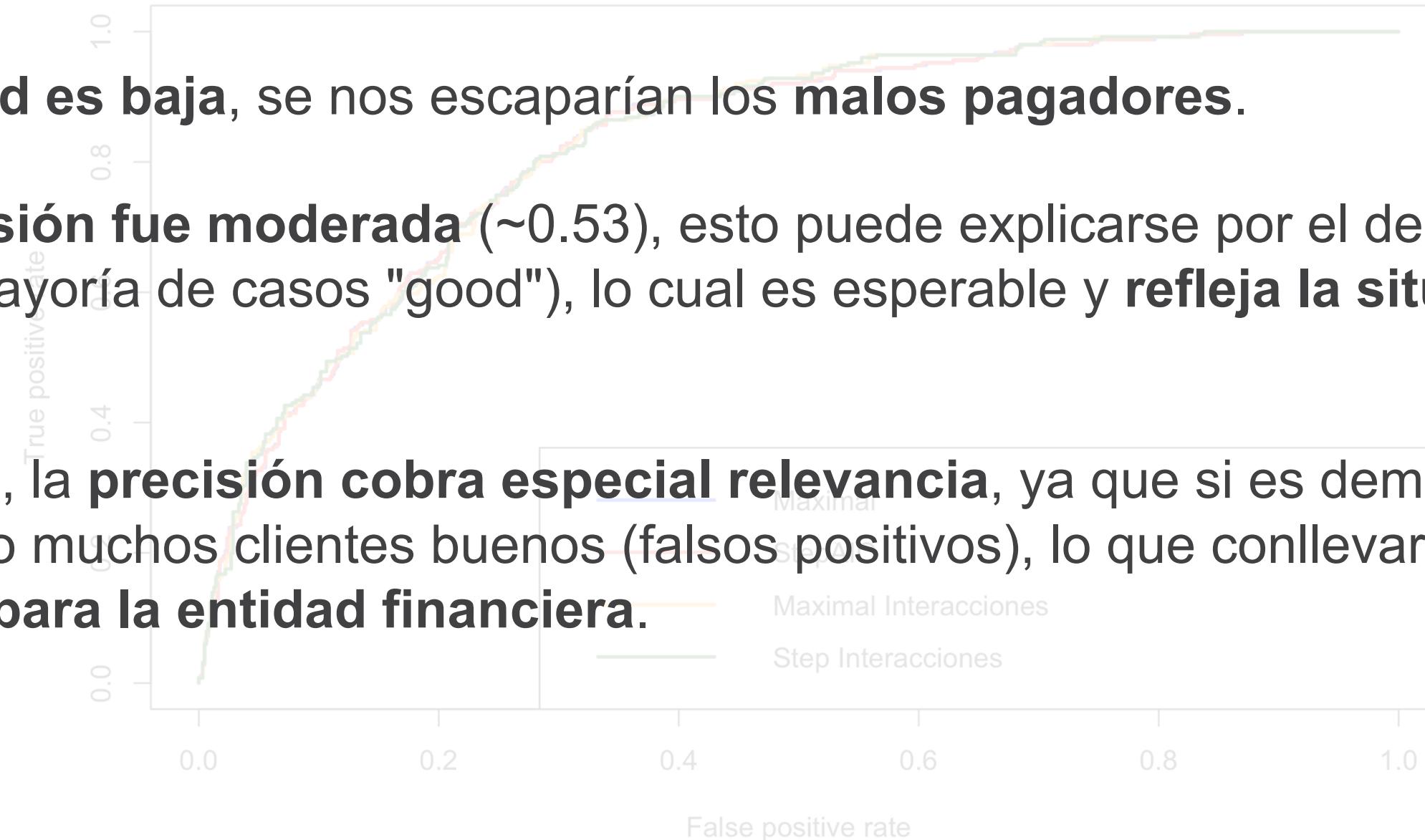
Curvas ROC

Objetivo:
prevenir impagos



Curvas ROC

Curvas ROC - Validación

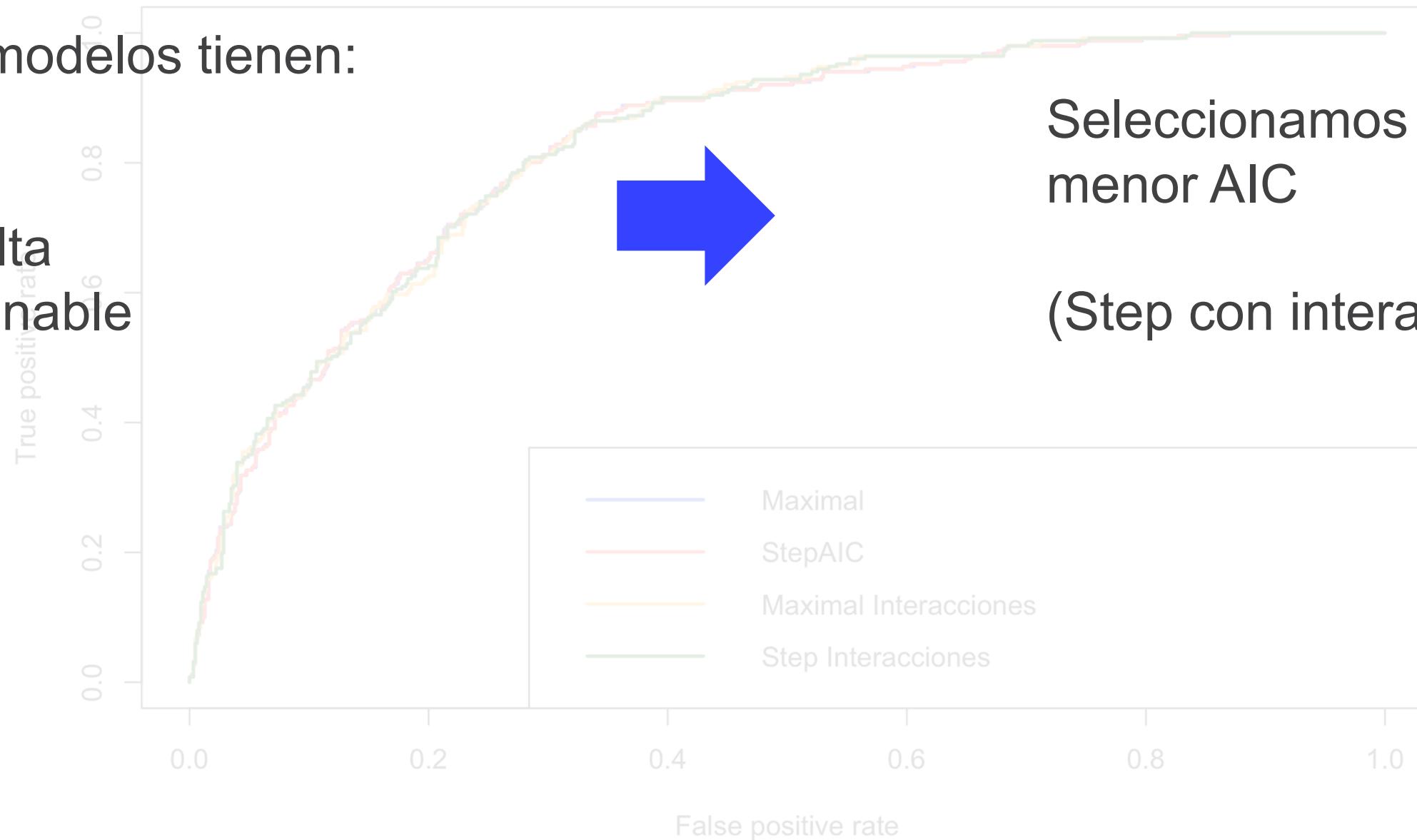


Selección del modelo

Como todos los modelos tienen:

- AUC elevado
- Sensibilidad alta
- Precision razonable

Curvas ROC - Validación



Seleccionamos el modelo de menor AIC
(Step con interacciones)

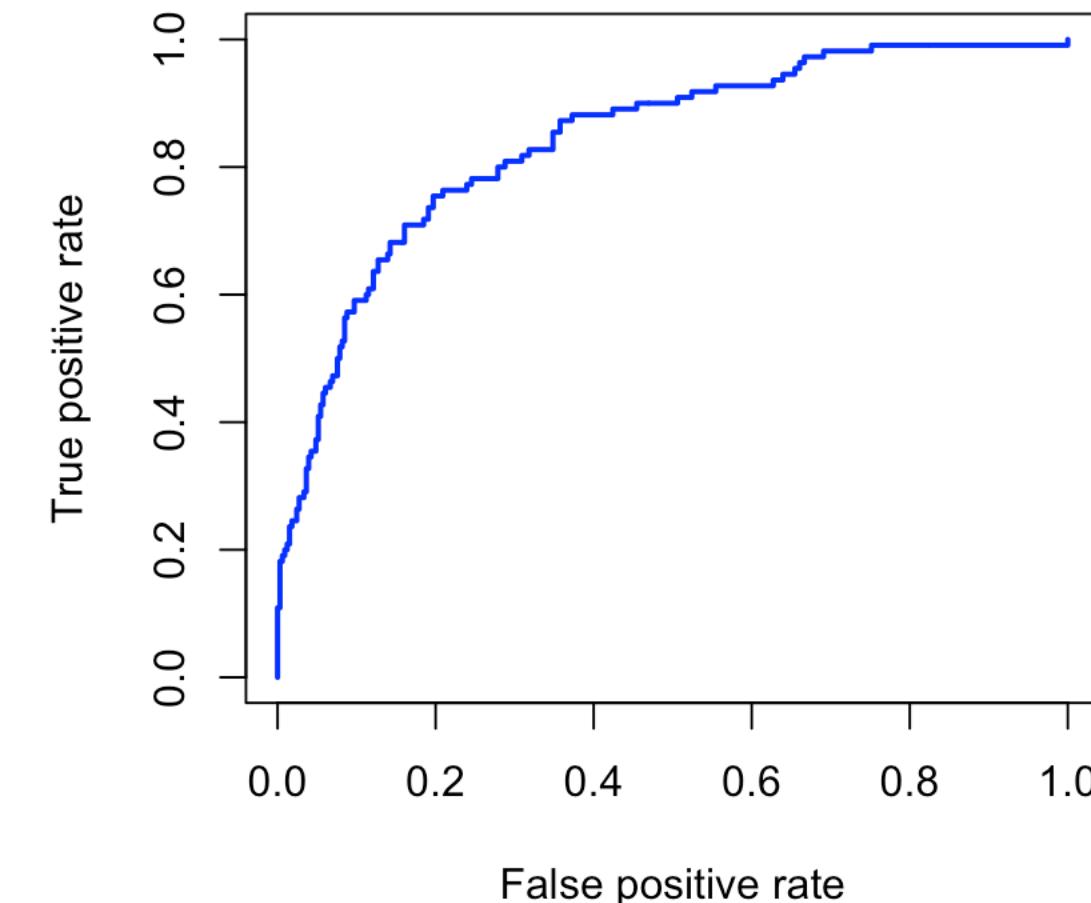


Evaluación del modelo

Ahora recurrimos al conjunto de test

```
=====
Modelo Final
AUC: 0.8409
Reference
Prediction bad good
    bad   81   64
    good  29  266
Accuracy: 0.7886
Sensibilidad (Recall bad): 0.7364
Especificidad: 0.8061
Precisión (PPV): 0.5586
=====
```

Curvas ROC - Validación



Tarea 3: Regresión logística ordinal



Tarea 3: Regresión logística ordinal

- 3.1 Establecer una pregunta de investigación
- 3.2 Aplicar un modelo de regresión logística ordinal
- 3.3 Interpretar los resultados del mismo



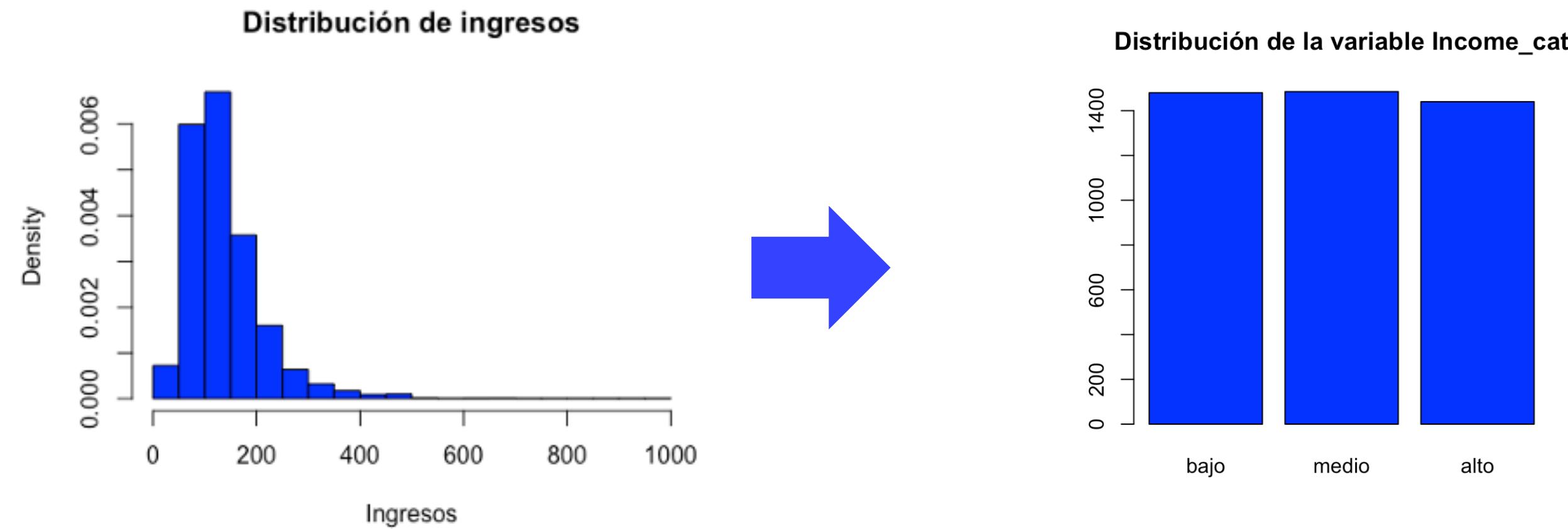
Pregunta de investigación

La pregunta que se pretende responder es:

¿Las **variables explicativas** (edad, tipo de empleo, estado civil, etc.) **están asociadas** a que un solicitante de crédito pertenezca a un **nivel bajo, medio o alto de ingresos**?



Categorizamos la variable ingresos



UNIVERSIDAD
COMPLUTENSE
MADRID



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID

Aplicamos un modelo de reg. log. ordinal

```
polr(formula = Income_cat ~ Job + Marital + Age + Records + Home,  
      data = credit_data_imputed, method = "logistic")
```

Coefficients:

	Value	Std. Error	t value
Jobothers_j	-0.32423	0.058982	-5.497
Maritalothers_m	-0.53535	0.070081	-7.639
Age	0.01301	0.002825	4.605
Recordsno	-0.22159	0.074371	-2.979
Homeowner	0.29210	0.060172	4.854

Intercepts:

	Value	Std. Error	t value
bajolmedio	-0.5333	0.1330	-4.0107
mediolalto	0.9321	0.1335	6.9832

Interpretando los Odds Ratio estimados con el modelo:

- Los solicitantes sin trabajo fijo tienen **26% menos odds** de estar en un nivel superior de ingresos respecto a quienes tienen empleo fijo.
- Las personas **no casadas** tienen **menor probabilidad de pertenecer a niveles altos de ingresos** frente a las casadas.
- La **probabilidad de mayores ingresos aumenta ligeramente con la edad**.
- **Ser propietario** se asocia con **mayor probabilidad** de estar en los **niveles altos de ingresos**.



Comparación con un modelo nulo

Se ha usado la prueba de razón de verosimilitudes ($\text{Chi}^2 = 228.84$, $gl = 5$, $p < 0.001$)



Lo que indica que el modelo con predictores explica significativamente mejor la variabilidad en los niveles de ingresos.

Podemos concluir entonces que las variables sociodemográficas como el tipo de empleo, estado civil, edad, situación habitacional y antecedentes crediticios están asociadas con el nivel de ingresos de los solicitantes de crédito.

Estos resultados pueden ser útiles para entender el perfil económico de los clientes y orientar estrategias de segmentación o riesgo.



UNIVERSIDAD
COMPLUTENSE
MADRID



POLITÉCNICA

UNIVERSIDAD
POLÍTÉCNICA
DE MADRID

MUCHAS GRACIAS



UNIVERSIDAD
COMPLUTENSE
MADRID



POLITÉCNICA

UNIVERSIDAD
POLITÉCNICA
DE MADRID