



Universidad
Zaragoza

El análisis de datos en
departamentos de marketing
de entidades financieras

TRABAJO FIN DE GRADO

ANEXOS

DANIEL ALDEA GONZÁLEZ

ÍNDICE

Anexo 1: Análisis exploratorio.....	2
1.1 Análisis discriminante.....	2
1.2 Residuos y contraste Chi-cuadrado.....	5
Anexo II: Modelo de regresión logística	9

Anexo 1: Análisis exploratorio.

1.1 Análisis discriminante

Resumen de proceso del caso de análisis

Casos sin ponderar	N	Porcentaje
Válido	4521	100,0
Excluido Código de grupo perdidos o fuera de rango	0	,0
Como mínimo, falta una variable discriminatoria	0	,0
Ambos códigos, los perdidos o los que están fuera de rango y, como mínimo, una discriminación que falta	0	,0
Total	4521	100,0

Estadísticas de grupo

Suscripción		N válido (por lista)	
		No ponderados	Ponderados
No ha suscrito	Edad	4000	4000,000
	Saldo medio anual	4000	4000,000
Sí ha suscrito	Edad	521	521,000
	Saldo medio anual	521	521,000
Total	Edad	4521	4521,000
	Saldo medio anual	4521	4521,000

Autovalores

Función	Autovalor	% de varianza	% acumulado	Correlación canónica
1	,002 ^a	100,0	100,0	,047

a. Se utilizaron las primeras 1 funciones discriminantes canónicas en el análisis.

Lambda de Wilks

Prueba de funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1	,998	10,105	2	,006

Coeficientes de función

discriminante canónica

estandarizadas

	Función
	1
Edad	,929
Saldo medio anual	,301

Matriz de estructuras

	Función
	1
Edad	,954
Saldo medio anual	,378

Correlaciones dentro de grupos combinados entre las variables discriminantes y las funciones discriminantes canónicas estandarizadas
 Variables ordenadas por el tamaño absoluto de la correlación dentro de la función.

Resumen de proceso de clasificación

Procesado		4521
Excluido	Código de grupo perdidos o fuera de rango	0
	Como mínimo, falta una variable discriminatoria	0
Utilizado en resultado		4521

Probabilidades previas para grupos

Suscripción	Previa	Caso utilizados en análisis	
		No ponderados	Ponderados
No ha suscrito	,500	4000	4000,000
Sí ha suscrito	,500	521	521,000
Total	1,000	4521	4521,000

Resultados de clasificación^a

Original	Recuento	Suscripción	Pertenencia a grupos pronosticada		Total
			No ha suscrito	Sí ha suscrito	
Original	Recuento	No ha suscrito	2334	1666	4000
		Sí ha suscrito	285	236	521
	%	No ha suscrito	58,4	41,7	100,0
		Sí ha suscrito	54,7	45,3	100,0

a. 56,8% de casos agrupados originales clasificados correctamente.

1.2 Residuos y contraste Chi-cuadrado

Tabla cruzada

			Estado civil			Total
			Soltero	Casado	Divorciado o viudo	
Suscripción	No ha suscrito	Recuento	1029	2520	451	4000
		Residuo corregido	-3,1	4,3	-2,3	
	Sí ha suscrito	Recuento	167	277	77	521
		Residuo corregido	3,1	-4,3	2,3	
Total		Recuento	1196	2797	528	4521

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (2 caras)
Chi-cuadrado de Pearson	19,030 ^a	2	,000
Razón de verosimilitud	18,630	2	,000
Asociación lineal por lineal	1,023	1	,312
N de casos válidos	4521		

a. 0 casillas (,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 60,85.

Tabla cruzada

			Educación			Total
			Primaria	Secundaria	Terciaria	
Suscripción	No ha suscrito	Recuento	614	2061	1157	3832
		Residuo corregido	1,9	2,1	-3,8	
	Sí ha suscrito	Recuento	64	245	193	502
		Residuo corregido	-1,9	-2,1	3,8	
Total		Recuento	678	2306	1350	4334

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (2 caras)
Chi-cuadrado de Pearson	14,815 ^a	2	,001
Razón de verosimilitud	14,458	2	,001
Asociación lineal por lineal	13,283	1	,000
N de casos válidos	4334		

a. 0 casillas (,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 78,53.

Tabla cruzada

Suscripción	No ha suscrito	Recuento	Préstamo vivienda		Total
			No	Si	
Suscripción	No ha suscrito	Recuento	1661	2339	4000
		Residuo corregido	-7,0	7,0	
	Sí ha suscrito	Recuento	301	220	521
		Residuo corregido	7,0	-7,0	
Total		Recuento	1962	2559	4521

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (2 caras)	Significación exacta (2 caras)	Significación exacta (1 cara)
Chi-cuadrado de Pearson	49,544 ^a	1	,000		
Corrección de continuidad ^b	48,885	1	,000		
Razón de verosimilitud	49,057	1	,000		
Prueba exacta de Fisher				,000	,000
Asociación lineal por lineal	49,533	1	,000		
N de casos válidos	4521				

a. 0 casillas (,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 226,10.

b. Sólo se ha calculado para una tabla 2x2

Tabla cruzada

			Préstamo personal		Total
			No	Si	
Suscripción	No ha suscrito	Recuento	3352	648	4000
		Residuo corregido	-4,7	4,7	
	Sí ha suscrito	Recuento	478	43	521
		Residuo corregido	4,7	-4,7	
Total		Recuento	3830	691	4521

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (2 caras)	Significación exacta (2 caras)	Significación exacta (1 cara)
Chi-cuadrado de Pearson	22,481 ^a	1	,000		
Corrección de continuidad ^b	21,872	1	,000		
Razón de verosimilitud	25,777	1	,000		
Prueba exacta de Fisher				,000	,000
Asociación lineal por lineal	22,476	1	,000		
N de casos válidos	4521				

a. 0 casillas (,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 79,63.

b. Sólo se ha calculado para una tabla 2x2

Tabla cruzada

			EdadCuali4				Total
			1,00	2,00	3,00	4,00	
Suscripción	No ha suscrito	Recuento	2164	1039	718	79	4000
		Residuo corregido	1,1	1,2	1,2	-9,4	
	Sí ha suscrito	Recuento	268	123	82	48	521
		Residuo corregido	-1,1	-1,2	-1,2	9,4	
Total		Recuento	2432	1162	800	127	4521

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (2 caras)
Chi-cuadrado de Pearson	88,853 ^a	3	,000
Razón de verosimilitud	61,321	3	,000
Asociación lineal por lineal	13,981	1	,000
N de casos válidos	4521		

a. 0 casillas (,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 14,64.

Tabla cruzada

			Rdo campaña anterior		Total
			Fracaso	Éxito	
Suscripción	No ha suscrito	Recuento	427	46	473
		Residuo corregido	12,3	-12,3	
	Sí ha suscrito	Recuento	63	83	146
		Residuo corregido	-12,3	12,3	
Total		Recuento	490	129	619

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (2 caras)	Significación exacta (2 caras)	Significación exacta (1 cara)
Chi-cuadrado de Pearson	150,178 ^a	1	,000		
Corrección de continuidad ^b	147,335	1	,000		
Razón de verosimilitud	132,219	1	,000		
Prueba exacta de Fisher				,000	,000
Asociación lineal por lineal	149,935	1	,000		
N de casos válidos	619				

a. 0 casillas (,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 30,43.

b. Sólo se ha calculado para una tabla 2x2

Anexo II: Modelo de regresión logística

Resumen de procesamiento de casos

Casos sin ponderar ^a		N	Porcentaje
Casos seleccionados	Incluido en el análisis	4334	95,9
	Casos perdidos	187	4,1
	Total	4521	100,0
Casos no seleccionados		0	,0
Total		4521	100,0

a. Si la ponderación está en vigor, consulte la tabla de clasificación para el número total de casos.

Codificación de variable

dependiente

Valor original	Valor interno
No ha suscrito	0
Sí ha suscrito	1

Codificaciones de variables categóricas

		Frecuencia	Codificación de parámetro		
			(1)	(2)	(3)
EdadCuali4	1,00	2366	,000	,000	,000
	2,00	1108	1,000	,000	,000
	3,00	743	,000	1,000	,000
	4,00	117	,000	,000	1,000
Educación	Primaria	678	,000	,000	
	Secundaria	2306	1,000	,000	
	Terciaria	1350	,000	1,000	
Estado civil	Soltero	1150	,000	,000	
	Casado	2680	1,000	,000	
	Divorciado o viudo	504	,000	1,000	
Préstamo personal	No	3650	,000		
	Si	684	1,000		
Préstamo vivienda	No	1858	,000		
	Si	2476	1,000		

Bloque 0: Bloque de inicio

Tabla de clasificación^{a,b}

Observado		Pronosticado			Corrección de porcentaje	
		Suscripción		No ha suscrito		
		Sí ha suscrito				
Paso 0	Suscripción	No ha suscrito		0	3832 ,0	
		Sí ha suscrito		0	502 100,0	
	Porcentaje global				11,6	

a. La constante se incluye en el modelo.

b. El valor de corte es ,100

Variables en la ecuación

	B	Error estándar	Wald	gl	Sig.	Exp(B)
Paso 0 Constante	-2,033	,047	1833,662	1	,000	,131

Las variables no están en la ecuación

		Puntuación	gl	Sig.
Paso 0 Variables	Estadocivil	21,506	2	,000
	Estadocivil(1)	21,468	1	,000
	Estadocivil(2)	5,351	1	,021
	Educación	14,815	2	,001
	Educación(1)	4,420	1	,036
	Educación(2)	14,097	1	,000
	Préstamo_vivienda(1)	47,411	1	,000
	Préstamo_personal(1)	22,246	1	,000
	EdadCuali4	80,130	3	,000
	EdadCuali4(1)	1,802	1	,179
	EdadCuali4(2)	1,302	1	,254
	EdadCuali4(3)	79,518	1	,000
	Estadísticos globales	158,637	9	,000

Bloque 1: Método = Avanzar por paso (Wald)

Pruebas ómnibus de coeficientes de modelo

		Chi-cuadrado	gl	Sig.
Paso 1	Escalón	55,573	3	,000
	Bloque	55,573	3	,000
	Modelo	55,573	3	,000
Paso 2	Escalón	33,287	1	,000
	Bloque	88,860	4	,000
	Modelo	88,860	4	,000
Paso 3	Escalón	22,288	2	,000
	Bloque	111,148	6	,000
	Modelo	111,148	6	,000
Paso 4	Escalón	18,586	1	,000
	Bloque	129,734	7	,000
	Modelo	129,734	7	,000
Paso 5	Escalón	9,148	2	,010
	Bloque	138,883	9	,000
	Modelo	138,883	9	,000

Resumen del modelo

Escalón	Logaritmo de la verosimilitud -2	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	3052,165 ^a	,013	,025
2	3018,878 ^a	,020	,040
3	2996,590 ^a	,025	,049
4	2978,004 ^a	,029	,058
5	2968,855 ^a	,032	,062

a. La estimación ha terminado en el número de iteración 5 porque las estimaciones de parámetro han cambiado en menos de ,001.

Prueba de Hosmer y Lemeshow

Escalón	Chi-cuadrado	gl	Sig.
1	,000	1	1,000
2	2,886	4	,577
3	5,565	7	,591
4	8,788	8	,361
5	5,088	8	,748

Tabla de contingencia para la prueba de Hosmer y Lemeshow

		Suscripción = No ha suscrito		Suscripción = Sí ha suscrito		Total
		Observado	Esperado	Observado	Esperado	
Paso 1	1	666	666,000	77	77,000	743
	2	992	992,000	116	116,000	1108
	3	2174	2174,000	309	309,000	2483
Paso 2	1	292	289,710	21	23,290	313
	2	594	601,156	61	53,844	655
	3	1368	1362,026	129	134,974	1497
	4	374	376,290	56	53,710	430
	5	398	390,844	55	62,156	453
	6	806	811,974	180	174,026	986
Paso 3	1	216	212,412	12	15,588	228
	2	796	800,814	67	62,186	863
	3	439	441,450	37	34,550	476
	4	265	265,052	32	31,948	297
	5	302	303,992	40	38,008	342
Paso 4	6	472	461,425	48	58,575	520
	7	389	395,048	58	51,952	447
	8	404	397,534	55	61,466	459
	9	549	554,274	153	147,726	702
	1	379	384,320	25	19,680	404
	2	344	336,844	19	26,156	363
	3	643	645,350	57	54,650	700
	4	395	394,389	34	34,611	429
	5	288	287,605	37	37,395	325
	6	408	397,928	44	54,072	452
	7	289	297,699	50	41,301	339
	8	275	276,167	41	39,833	316
	9	380	374,964	61	66,036	441
	10	431	436,734	134	128,266	565

Paso 5	1	392	394,492	22	19,508	414
	2	330	328,427	21	22,573	351
	3	404	397,629	26	32,371	430
	4	382	382,910	34	33,090	416
	5	366	375,042	48	38,958	414
	6	367	364,304	41	43,696	408
	7	402	403,951	56	54,049	458
	8	351	349,571	52	53,429	403
	9	303	298,025	46	50,975	349
	10	535	537,649	156	153,351	691

Tabla de clasificación^a

		Observado	Pronosticado		
			Suscripción		Corrección de porcentaje
			No ha suscrito	Sí ha suscrito	
Paso 1	Suscripción	No ha suscrito	0	3832	,0
		Sí ha suscrito	0	502	100,0
		Porcentaje global			11,6
Paso 2	Suscripción	No ha suscrito	2254	1578	58,8
		Sí ha suscrito	211	291	58,0
		Porcentaje global			58,7
Paso 3	Suscripción	No ha suscrito	1451	2381	37,9
		Sí ha suscrito	116	386	76,9
		Porcentaje global			42,4
Paso 4	Suscripción	No ha suscrito	1727	2105	45,1
		Sí ha suscrito	133	369	73,5
		Porcentaje global			48,4
Paso 5	Suscripción	No ha suscrito	1994	1838	52,0
		Sí ha suscrito	164	338	67,3
		Porcentaje global			53,8

a. El valor de corte es ,100

Tabla de clasificación^a

		Observado	Pronosticado		
			Suscripción		Corrección de porcentaje
			No ha suscrito	Sí ha suscrito	
Paso 1	Suscripción	No ha suscrito	3759	73	98,1
		Sí ha suscrito	458	44	8,8
		Porcentaje global			87,7
Paso 2	Suscripción	No ha suscrito	3759	73	98,1
		Sí ha suscrito	458	44	8,8
		Porcentaje global			87,7
Paso 3	Suscripción	No ha suscrito	3759	73	98,1
		Sí ha suscrito	458	44	8,8
		Porcentaje global			87,7
Paso 4	Suscripción	No ha suscrito	3759	73	98,1
		Sí ha suscrito	458	44	8,8
		Porcentaje global			87,7
Paso 5	Suscripción	No ha suscrito	3538	294	92,3
		Sí ha suscrito	400	102	20,3
		Porcentaje global			84,0

a. El valor de corte es ,200

Variables en la ecuación

	B	Error estándar	Wald	gl	Sig.	Exp(B)
Paso 1 ^a	EdadCuali4		66,345	3	,000	
	EdadCuali4(1)	-,076	,118	,413	,521	,927
	EdadCuali4(2)	-,087	,137	,404	,525	,917
	EdadCuali4(3)	1,564	,202	60,150	,000	4,779
	Constante	-2,070	,065	1008,746	,000	,126
Paso 2 ^b	Préstamo_vivienda(1)	-,574	,100	33,174	,000	,563
	EdadCuali4		47,394	3	,000	
	EdadCuali4(1)	-,101	,118	,730	,393	,904
	EdadCuali4(2)	-,209	,139	2,266	,132	,811
	EdadCuali4(3)	1,281	,207	38,238	,000	3,600
	Constante	-1,738	,084	429,070	,000	,176

Paso 3 ^c	Estadocivil		22,378	2	,000	
	Estadocivil(1)	-,491	,117	17,547	1	,000
	Estadocivil(2)	-,041	,164	,062	1	,803
	Préstamo_vivienda(1)	-,533	,100	28,232	1	,000
	EdadCuali4		52,312	3	,000	
	EdadCuali4(1)	,008	,125	,004	1	,950
	EdadCuali4(2)	-,057	,148	,145	1	,703
	EdadCuali4(3)	1,470	,216	46,324	1	,000
	Constante	-1,531	,098	243,274	1	,000
						,216
Paso 4 ^d	Estadocivil		21,509	2	,000	
	Estadocivil(1)	-,473	,118	16,152	1	,000
	Estadocivil(2)	-,011	,164	,004	1	,947
	Préstamo_vivienda(1)	-,533	,100	28,157	1	,000
	Préstamo_personal(1)	-,668	,167	16,018	1	,000
	EdadCuali4		45,756	3	,000	
	EdadCuali4(1)	,007	,125	,004	1	,952
	EdadCuali4(2)	-,035	,149	,057	1	,812
	EdadCuali4(3)	1,387	,217	40,967	1	,000
	Constante	-1,463	,099	217,416	1	,000
Paso 5 ^e	Estadocivil		18,732	2	,000	
	Estadocivil(1)	-,437	,118	13,644	1	,000
	Estadocivil(2)	,008	,165	,002	1	,962
	Educación		9,040	2	,011	
	Educación(1)	,237	,156	2,297	1	,130
	Educación(2)	,457	,163	7,875	1	,005
	Préstamo_vivienda(1)	-,502	,101	24,713	1	,000
	Préstamo_personal(1)	-,659	,167	15,485	1	,000
	EdadCuali4		49,106	3	,000	
	EdadCuali4(1)	,042	,126	,109	1	,742
	EdadCuali4(2)	,032	,151	,044	1	,834
	EdadCuali4(3)	1,483	,221	45,242	1	,000
	Constante	-1,806	,178	103,397	1	,000
						,164

- a. Variables especificadas en el paso 1: EdadCuali4.
- b. Variables especificadas en el paso 2: Préstamo_vivienda.
- c. Variables especificadas en el paso 3: Estadocivil.
- d. Variables especificadas en el paso 4: Préstamo_personal.
- e. Variables especificadas en el paso 5: Educación.

Las variables no están en la ecuación

			Puntuación	gl	Sig.
Paso 1	Variables	Estadocivil	27,736	2	,000
		Estadocivil(1)	27,371	1	,000
		Estadocivil(2)	4,896	1	,027
		Educación	17,976	2	,000
		Educación(1)	3,187	1	,074
		Educación(2)	15,541	1	,000
		Préstamo_vivienda(1)	33,782	1	,000
		Préstamo_personal(1)	17,745	1	,000
		Estadísticos globales	81,435	6	,000
Paso 2	Variables	Estadocivil	22,637	2	,000
		Estadocivil(1)	22,602	1	,000
		Estadocivil(2)	5,308	1	,021
		Educación	12,790	2	,002
		Educación(1)	1,629	1	,202
		Educación(2)	10,327	1	,001
		Préstamo_personal(1)	17,278	1	,000
		Estadísticos globales	47,996	5	,000
Paso 3	Variables	Educación	9,857	2	,007
		Educación(1)	1,516	1	,218
		Educación(2)	8,126	1	,004
		Préstamo_personal(1)	16,548	1	,000
Paso 4	Variables	Estadísticos globales	25,676	3	,000
		Educación	9,096	2	,011
		Educación(1)	,897	1	,344
		Educación(2)	6,907	1	,009
		Estadísticos globales	9,096	2	,011