

Default of credit card clients

Agustina Gonzalez Crespo L.61743



Contiene información sobre clientes de tarjetas de crédito de un banco en Taiwan desde Abril del 2005 a Septiembre del 2005

25 variables

**30,000
registros**

Archive UCI

Objetivo: Predecir si el cliente va a entrar en default o no, es decir si va a saldar su cuenta con el banco o no

Variables

- ID: ID de cada cliente
- LIMIT_BAL: cantidad de crédito otorgado en NT\$ (incluye crédito individual y familiar)
- SEX: Genero (1=hombre, 2=mujer)
- EDUCATION: (1=escuela de posgrado, 2=universidad, 3=secundaria, 4=otros)
- MARRIAGE: estado civil (1=casado, 2=soltero, 3=otros)
- AGE: edad en años
- PAY_o..6: estatus del pago anterior (-1=pago a tiempo, 1=tardo 1 mes en pagar,...,9=tardó 9 meses o más en pagar)
- BILL_AMT1..6: monto del estado de cuentas para Septiembre,2005 a Abril,2005 (NT\$)
- PAY_AMT1..6: monto del pago anterior desde Septiembre,2005 a Abril,2005 (NT\$)
- default.payment.next.month: Default payment (1=sí, 0=no)

Variable	Cantidad de ceros	Porcentaje de ceros	Cantidad NAs	Porcentaje NAs	Tipo	Unique
ID	0	0.00	0	0	integer	30,000
LIMIT_BAL	0	0.00	0	0	integer	81
SEX	0	0.00	0	0	integer	2
EDUCATION	14	0.05	0	0	integer	7
MARRIAGE	54	0.18	0	0	integer	4
AGE	0	0.00	0	0	integer	56
PAY_0	14,737	49.12	0	0	integer	11
PAY_2	15,730	52.43	0	0	integer	11
PAY_3	15,764	52.55	0	0	integer	11
PAY_4	16,455	54.85	0	0	integer	11
PAY_5	16,947	56.49	0	0	integer	10
PAY_6	16,286	54.29	0	0	integer	10
BILL_AMT1	2,008	6.69	0	0	integer	22,723
BILL_AMT2	2,506	8.35	0	0	integer	22,346
BILL_AMT3	2,870	9.57	0	0	integer	22,026
BILL_AMT4	3,195	10.65	0	0	integer	21,548
BILL_AMT5	3,506	11.69	0	0	integer	21,010
BILL_AMT6	4,020	13.40	0	0	integer	20,604

PAY_AMT1	5,249	17.50	0	0	integer	7,943
PAY_AMT2	5,396	17.99	0	0	integer	7,899
PAY_AMT3	5,968	19.89	0	0	integer	7,518
PAY_AMT4	6,408	21.36	0	0	integer	6,937
PAY_AMT5	6,703	22.34	0	0	integer	6,897
PAY_AMT6	7,173	23.91	0	0	integer	6,939
default.payment.next.month	23,364	77.88	0	0	integer	2

Luego se eliminó la variable ID y PAY_0 se lo cambio a PAY_1

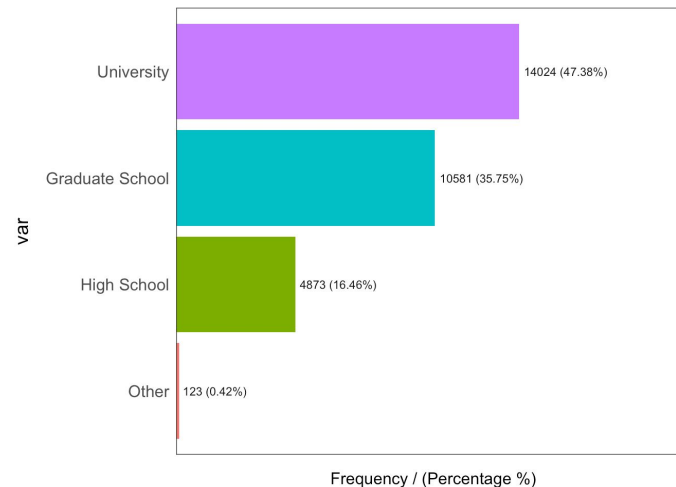
Variables categoricas

SEX	MARRIAGE	EDUCATION
1:11888	0: 54	0: 14
2:18112	1:13659	1:10585
NA	2:15964	2:14030
NA	3: 323	3: 4917
NA	NA	4: 123
NA	NA	5: 280
NA	NA	6: 51

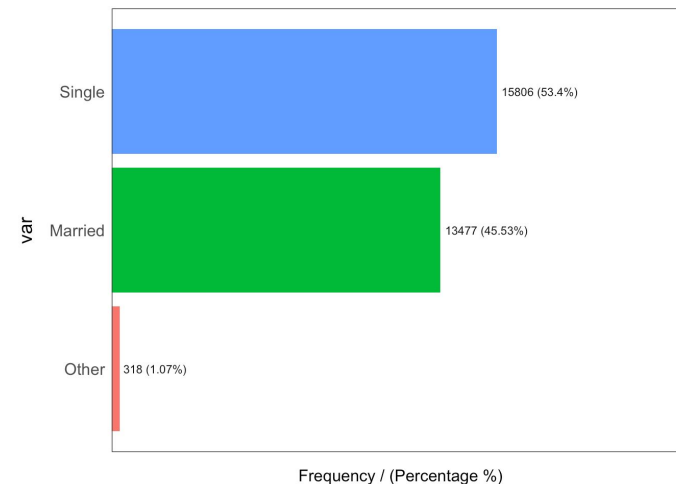
Education: 0,5,6 representan un **1,15%**

Marriage: 0 representa **0,18%**

• Frecuencia de Education



• Frecuencia de Marriage



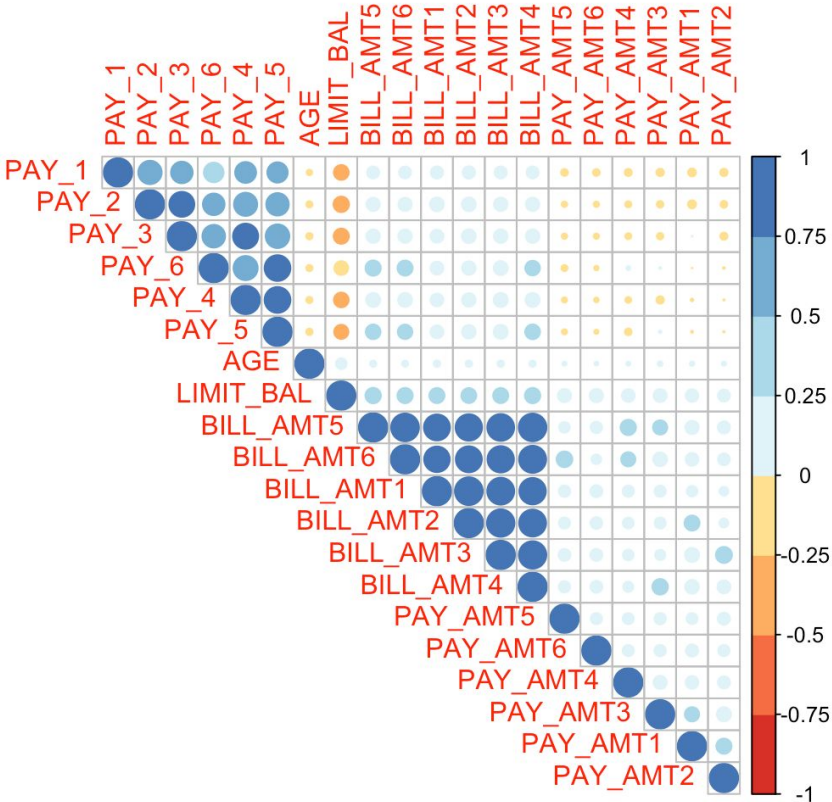
Variable PAY_1..6

PAY_1	PAY_2	PAY_3	PAY_4	PAY_5	PAY_6
Min. :-2.00000	Min. :-2.0000	Min. :-2.0000	Min. :-2.0000	Min. :-2.000	Min. :-2.0000
1st Qu.: -1.00000	1st Qu.: -1.0000	1st Qu.: -1.0000	1st Qu.: -1.0000	1st Qu.: -1.000	1st Qu.: -1.0000
Median : 0.00000	Median : 0.0000	Median : 0.0000	Median : 0.0000	Median : 0.000	Median : 0.0000
Mean :-0.01493	Mean :-0.1313	Mean :-0.1634	Mean :-0.2183	Mean :-0.264	Mean :-0.2876
3rd Qu.: 0.00000	3rd Qu.: 0.0000	3rd Qu.: 0.0000	3rd Qu.: 0.0000	3rd Qu.: 0.000	3rd Qu.: 0.0000
Max. : 8.00000	Max. : 8.0000	Max. : 8.0000	Max. : 8.0000	Max. : 8.000	Max. : 8.0000

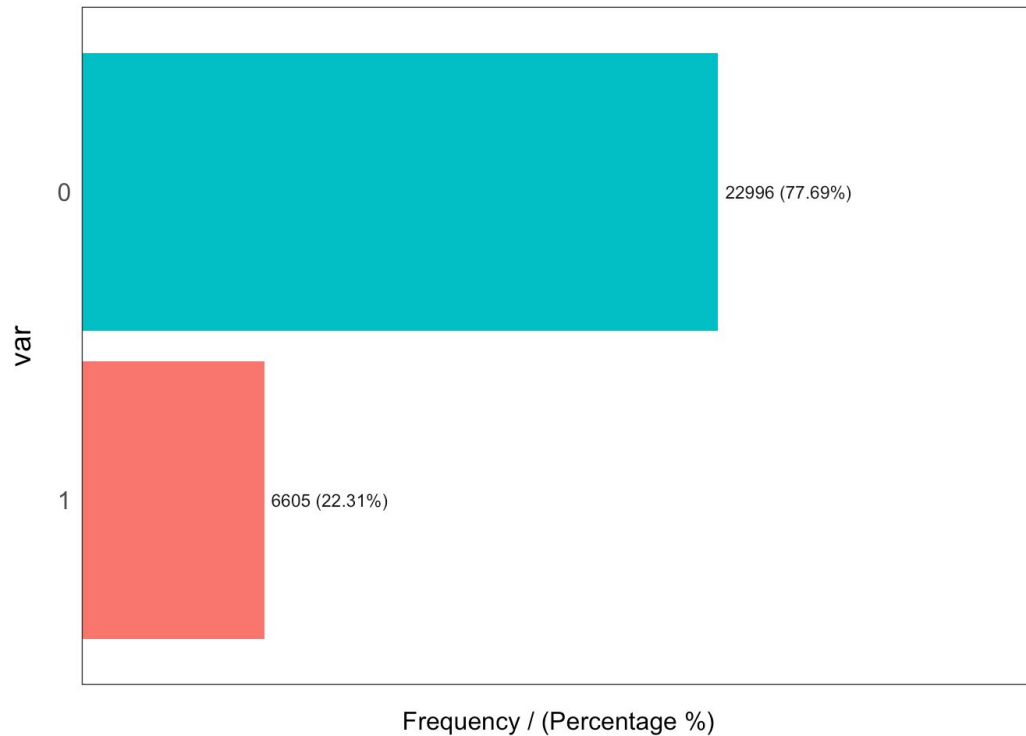
Variable BILL_AMT1..6

BILL_AMT1	BILL_AMT2	BILL_AMT3	BILL_AMT4	BILL_AMT5	BILL_AMT6
Min. :-165580	Min. :-69777	Min. :-157264	Min. :-170000	Min. :-81334	Min. :-339603
1st Qu.: 3528	1st Qu.: 2970	1st Qu.: 2652	1st Qu.: 2329	1st Qu.: 1780	1st Qu.: 1278
Median : 22259	Median : 21050	Median : 20035	Median : 19005	Median : 18091	Median : 17118
Mean : 50957	Mean : 48942	Mean : 46803	Mean : 43123	Mean : 40236	Mean : 38858
3rd Qu.: 66623	3rd Qu.: 63497	3rd Qu.: 59830	3rd Qu.: 54271	3rd Qu.: 50072	3rd Qu.: 49121
Max. : 964511	Max. :983931	Max. :1664089	Max. : 891586	Max. :927171	Max. : 961664

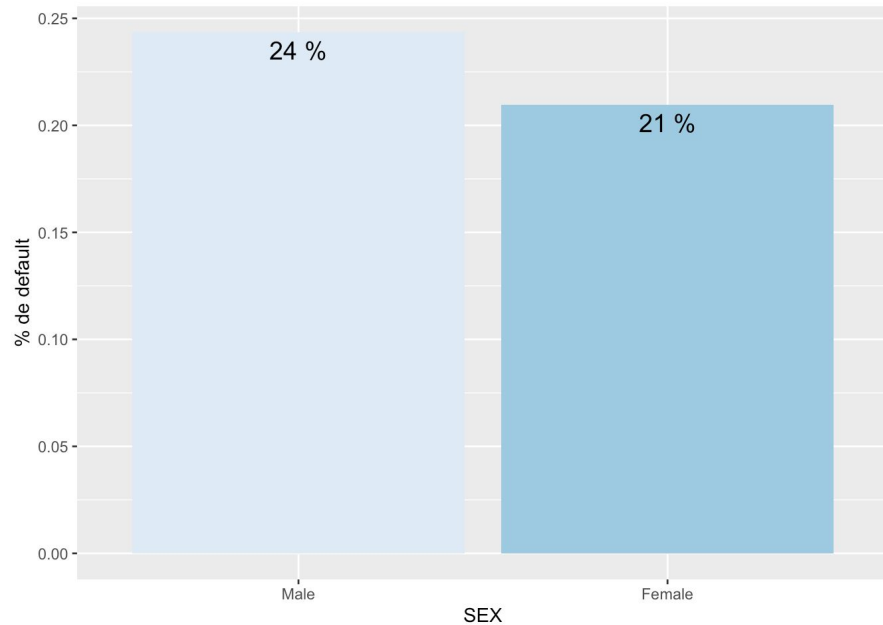
Correlacion



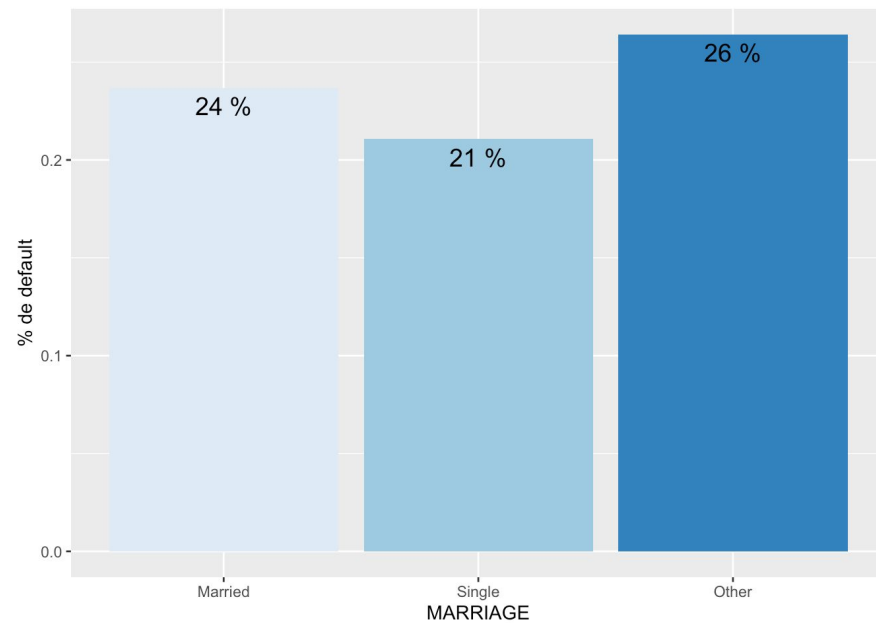
Variable Default



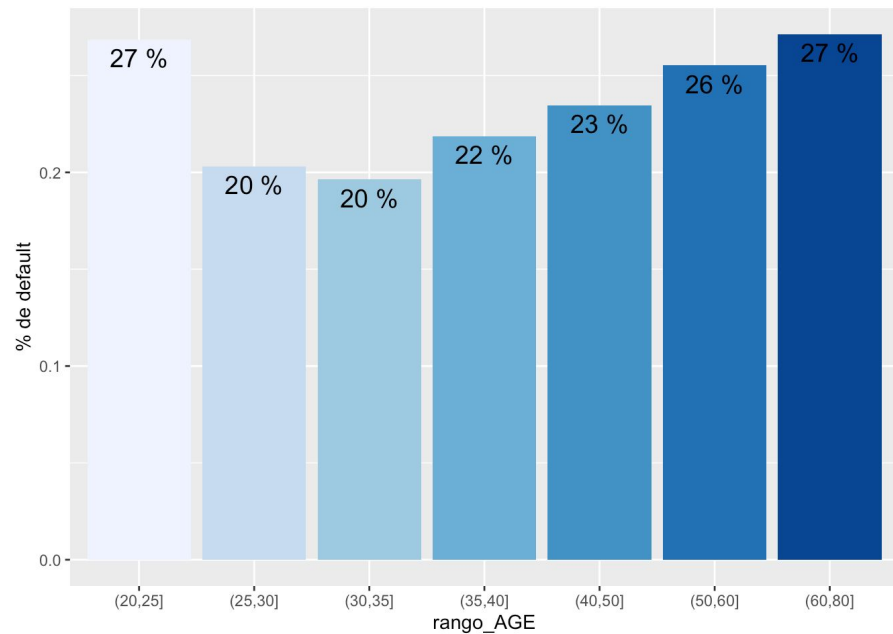
Genero



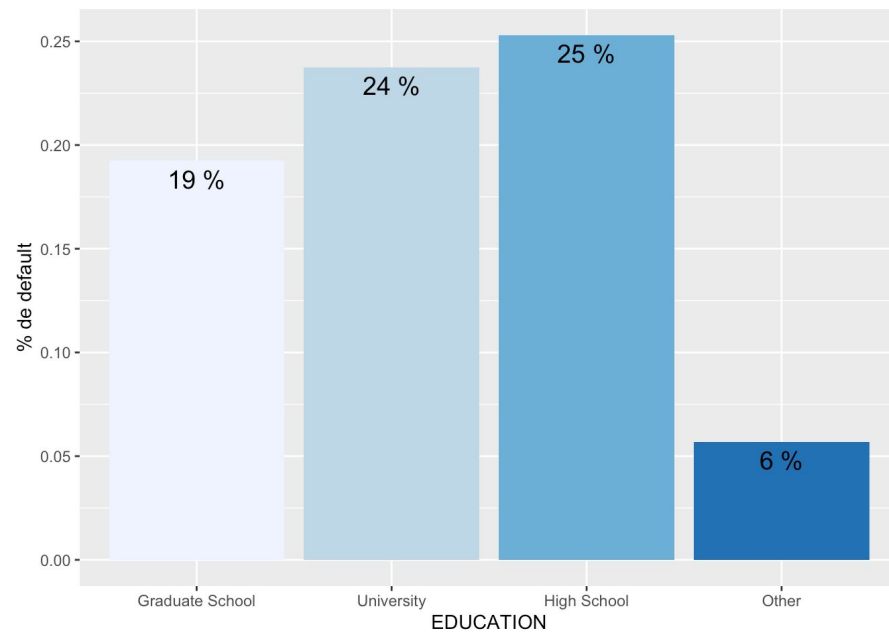
Estado civil



Edad

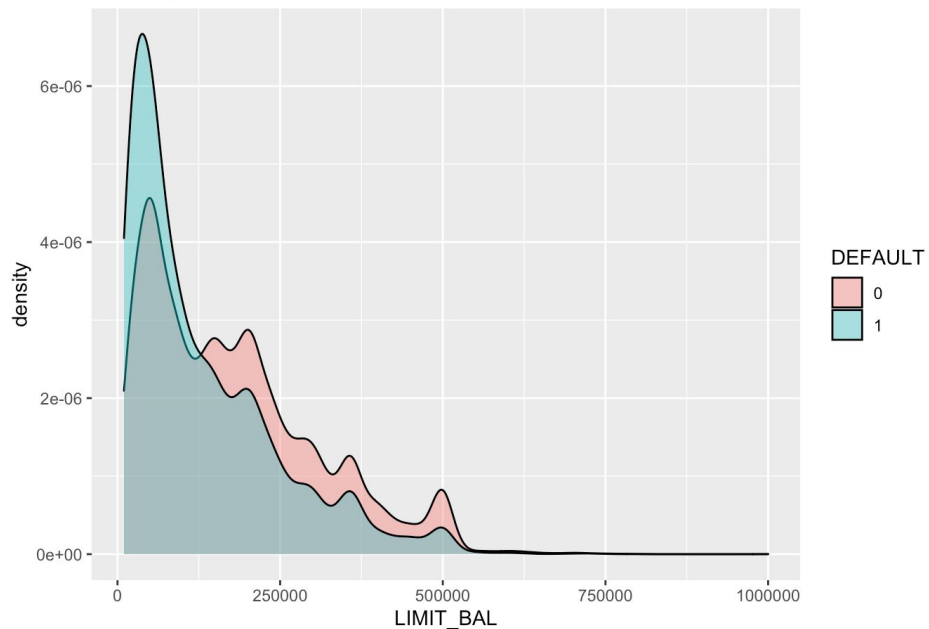


Nivel educativo

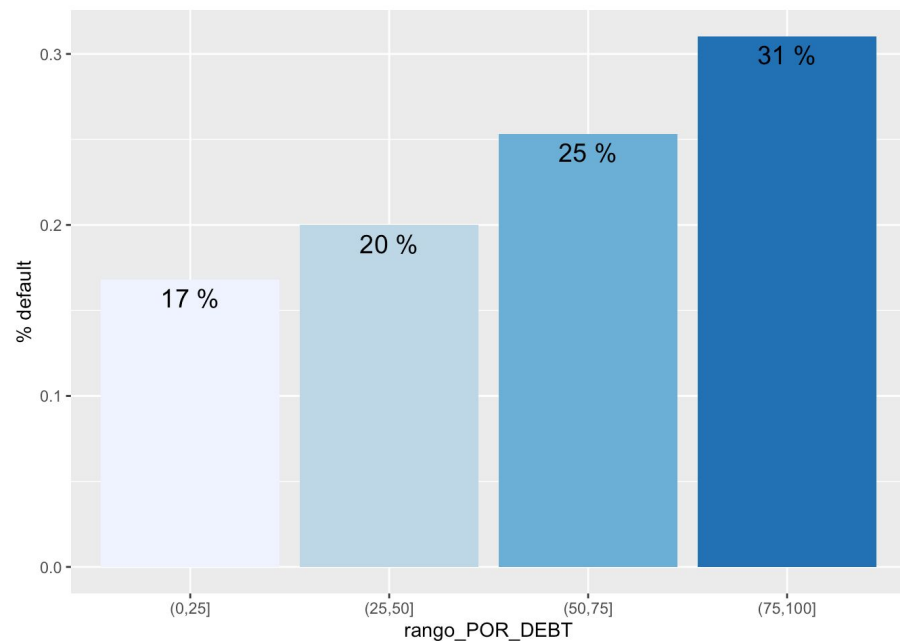


Límite de crédito

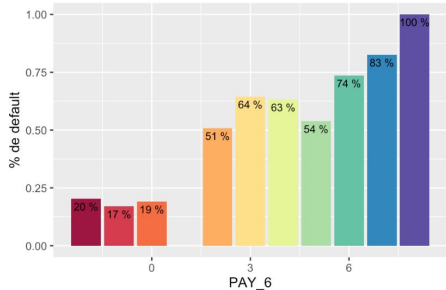
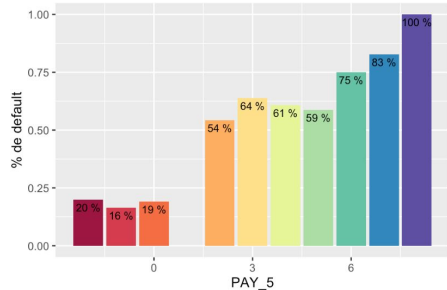
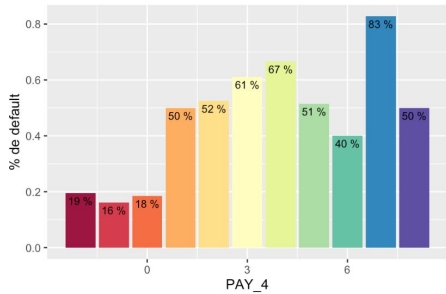
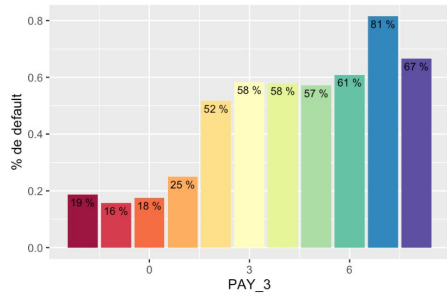
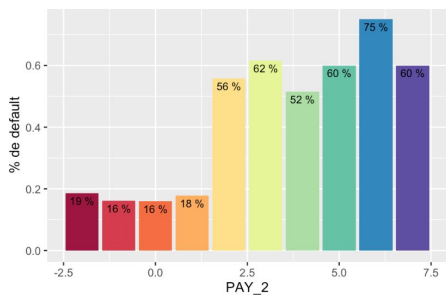
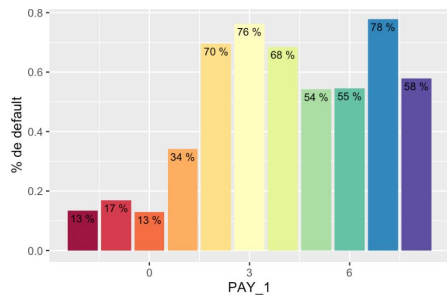
LIMIT_BAL Distribution



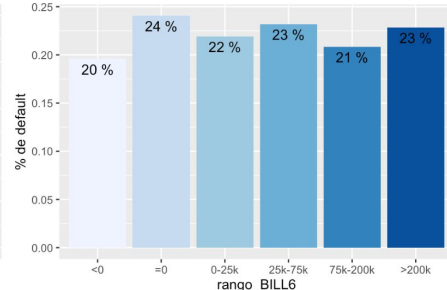
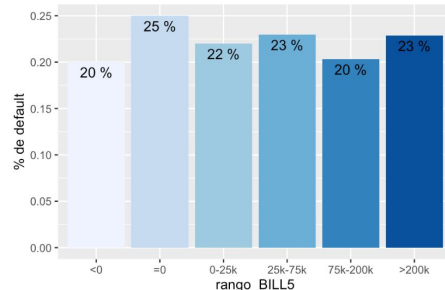
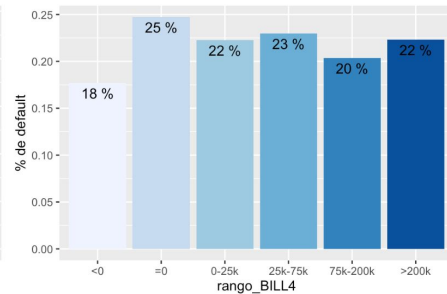
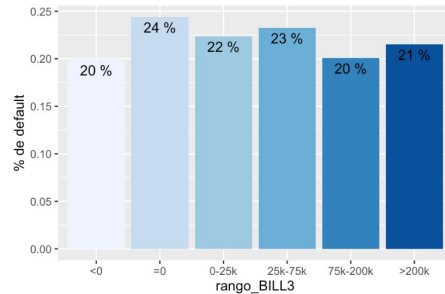
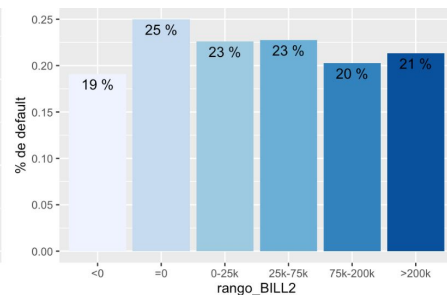
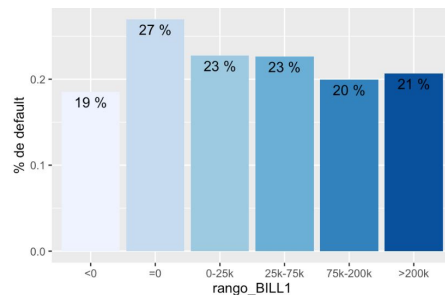
Porcentaje de consumo del límite



Retraso en pagar



Monto del estado de cuentas



Problemas e Información que se puede agregar

Problemas

Valores ingresados no son consistentes

Base de datos debe ser actualizada todos los meses

Debería cambiar la estructura para agregar mas datos

Información

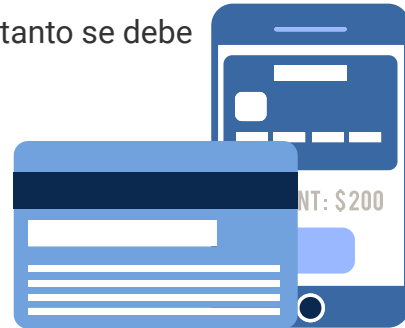
Utilizar datos del censo para ver la representatividad

Agregar más información sobre el resto del año

Situación económica de ese año en Taiwan

Conclusiones

- Los hombres tienen más probabilidades de default
- Los clientes que están divorciados o viudos tienen más probabilidades de default
- Cuanto mayor sea el límite de crédito, menor es la probabilidad de default
- Cuanto mayor sea el porcentaje de consumo sobre el límite de crédito, mayor es la probabilidad de default.
- Aquellos clientes con un porcentaje de consumo de su límite alto, tienen probabilidades más altas de default
- Puede ser que las variables **BILL_AMT1..6** no tienen significancia en el modelo
- Las variables **PAY_1..6** demuestran que cuanto más tarde el cliente en pagar, más grandes son las probabilidades de default. Estas variables también están correlacionadas entre sí, por lo tanto se debe analizar su significancia



Uso de la información

Información para decidir quién debería recibir una tarjeta de crédito y/o el límite de crédito



Conocimiento más profundo sobre los clientes actuales o futuros



Permite planificar servicios o productos específicos para los clientes



**¡Muchas
gracias!**