

Examen Final

[¿A quienes hemos preguntado?](#)

[¿Cómo están las finanzas de los encuestados?](#)

[¿Qué variables predicen el éxito de la campaña?](#)

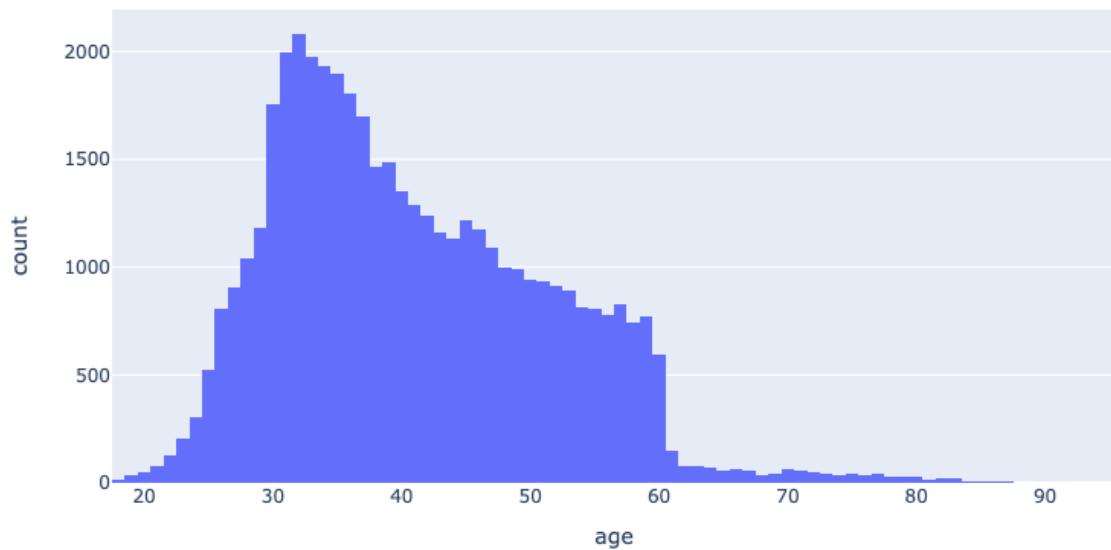
[¿Qué valores de estas variables predicen el éxito?](#)

[Conclusión](#)

[Capturas del Dash](#)

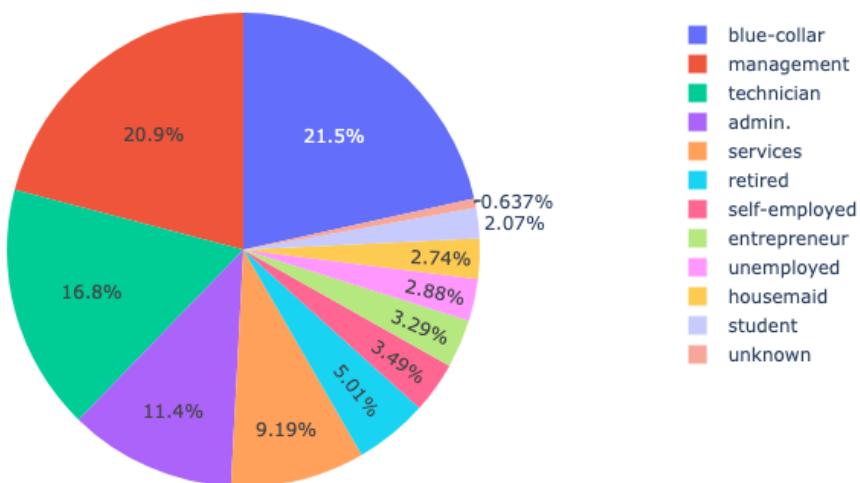
¿A quienes hemos preguntado?

En los siguientes gráficos podemos ver la distribución de la edad de los encuestados, su trabajo, estado marital y su educación



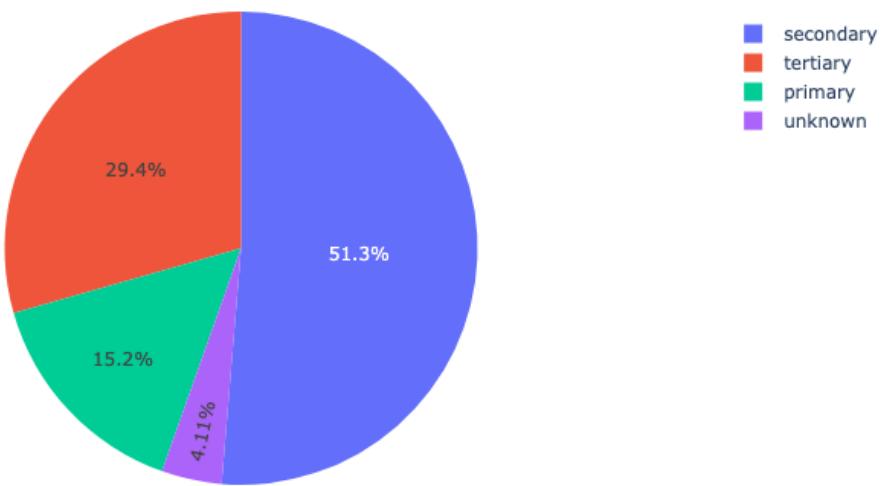
Distribución de la edad de los encuestados

La mayoría de las personas encuestadas se sitúan entre los 30 y 50 años, seguido de las personas entre 50 y 60 años.



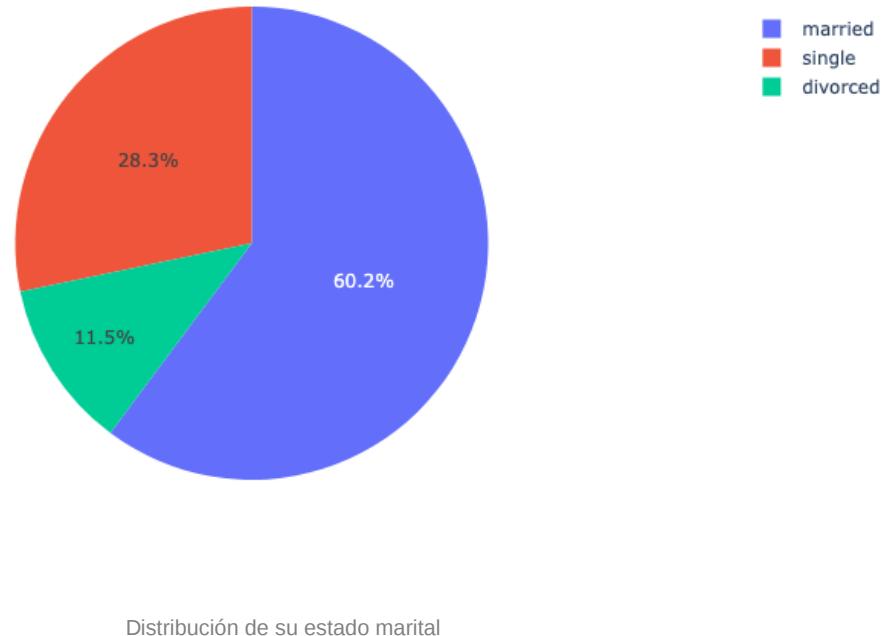
Distribución de trabajos

Las personas encuestadas son mayoritariamente, *blue-collar*, dirección y técnicos, cubriendo entre esto tres grupos más del 50% de las encuestas.



Distribución de su educación

A su vez, la mayoría de encuestados tiene la educación secundaria, seguidos un 30% con educación terciaria y un 15% con educación primaria.

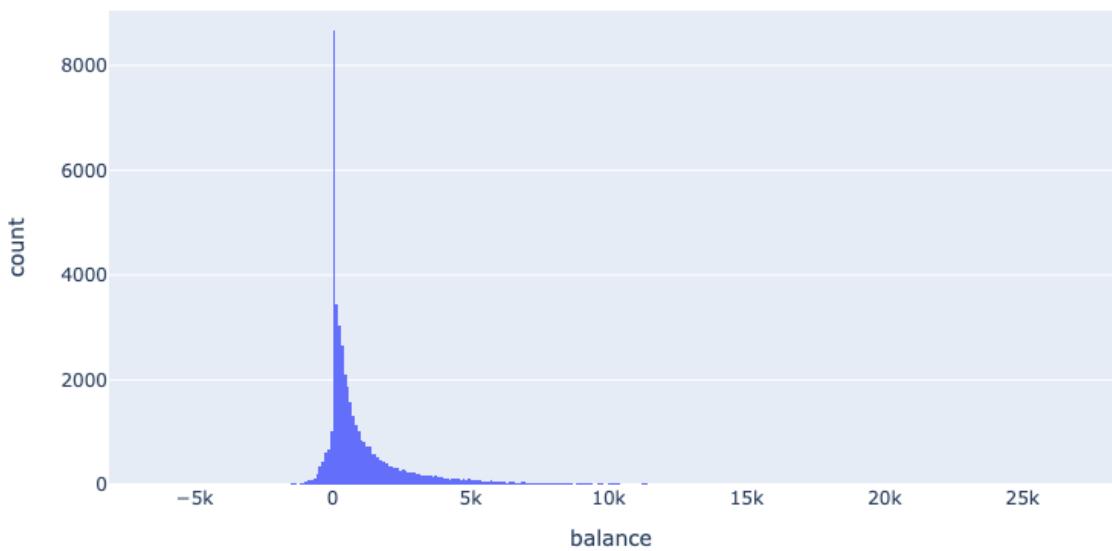


Finalmente, la mayoría de encuestados está casado (lo que guarda cierta relación con la edad media del encuestado).

Por tanto, podríamos decir que el perfil medio encuestado es una persona de 35 años, *blue-collar*, con educación secundaria y casado.

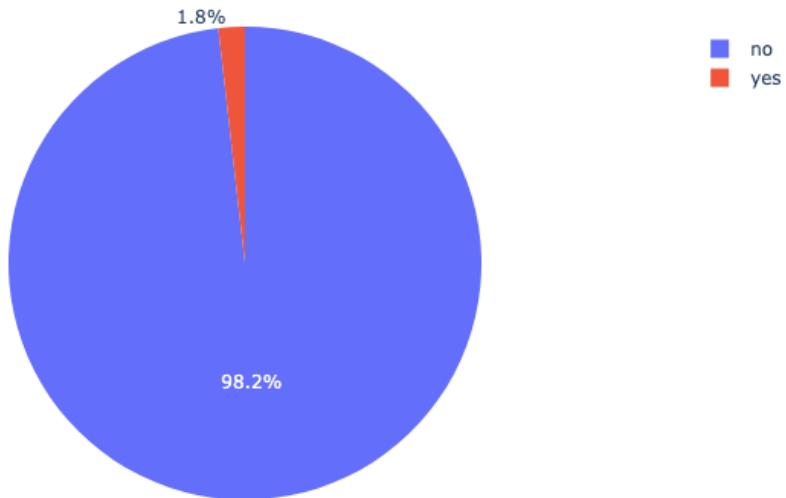
¿Cómo están las finanzas de los encuestados?

A continuación evaluaremos la situación económica de las personas encuestadas.



Distribución del balance

Podemos ver por el histograma del balance, que la mayoría de encuestados tienen un balance medio anual de entre 0 y 2500€.



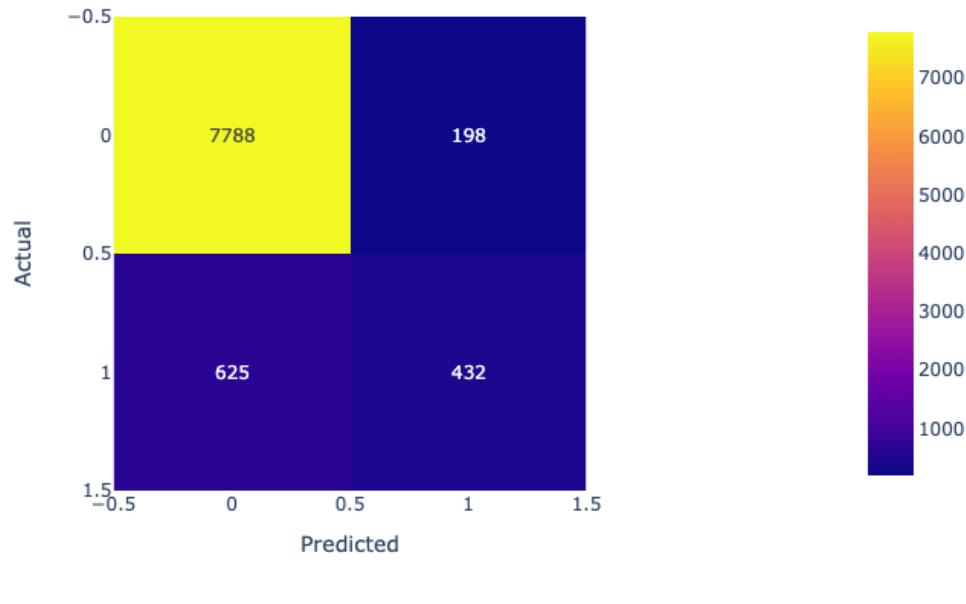
% de encuestados que no han pagado sus deudas

Y en este gráfico podemos ver que la mayoría de encuestados han pagado sus deudas (no han *defaulted*).

¿Qué variables predicen el éxito de la campaña?

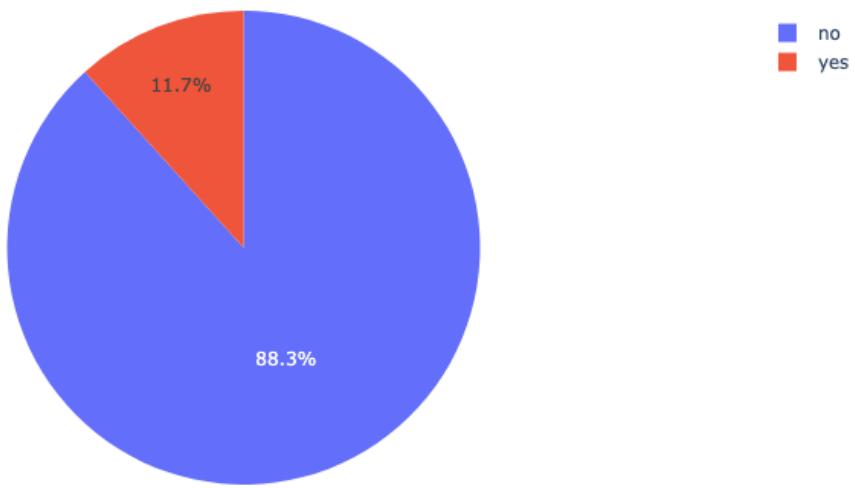
Más importante que conocer el perfil de los encuestados es saber qué variables predicen el éxito de la campaña. Para extraer esta información se ha entrenado un *random forest* y se ha extraído la importancia de cada variable.

Antes de ver qué variables son más importantes, vamos a evaluar la predicción del clasificador.



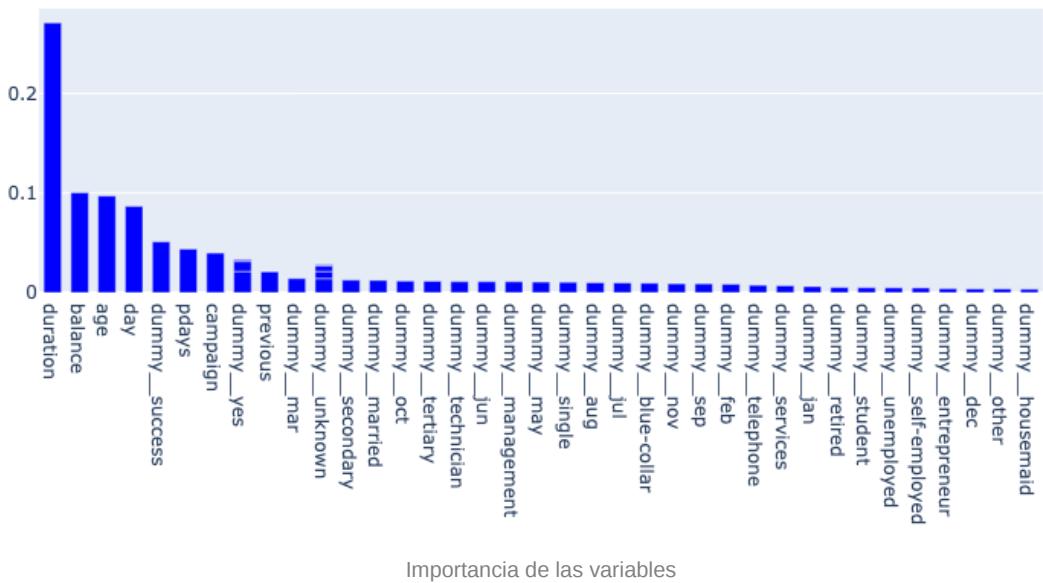
Matriz de confusión sobre el set de test

Podemos ver que el clasificador tiene una precisión del 91% y un *weighted recall* del 91% también. Además cabe destacar que hay más verdaderos positivos que ningún otro valor. Esto se podría deber a que hay más noes que síes y por tanto los datos pueden estar sesgados hacia ese valor.



La precisión del modelo nos indica que este (aparentemente) es un buen modelo para predecir la respuesta de la encuesta.

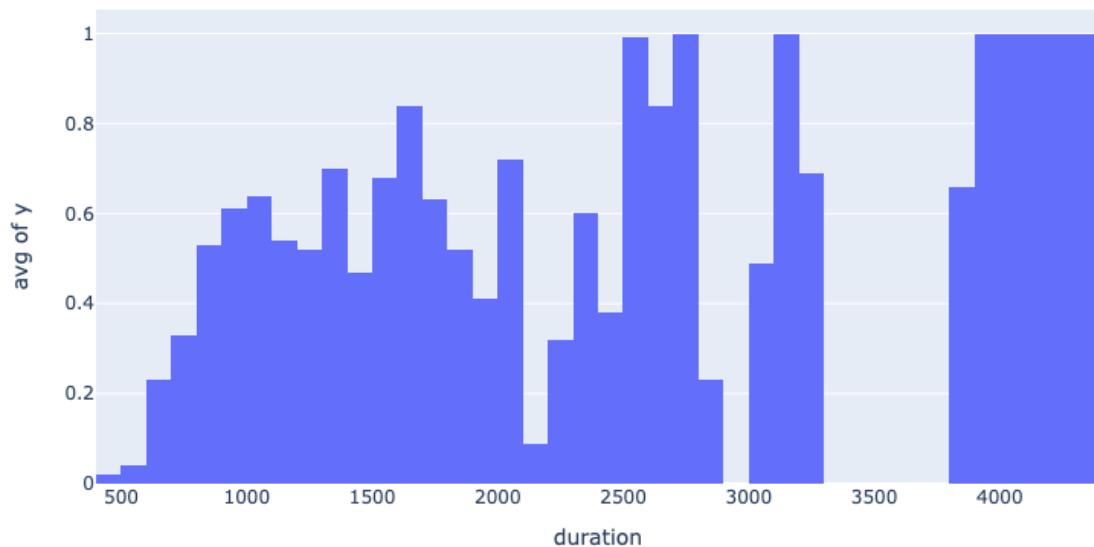
Sabiendo que el modelo funciona correctamente, podemos evaluar qué variables son más significativas. La siguiente gráfica muestra la importancia de las variables ordenadas de mayor a menor.



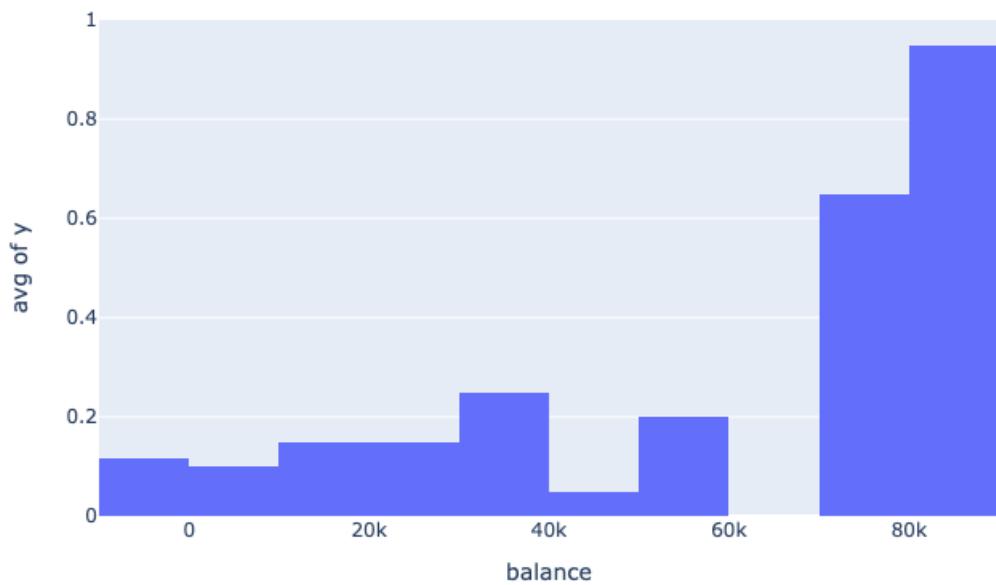
Podemos ver en el gráfico que la variable más importante a la hora de predecir es la duración, seguido del balance anual, la edad y el día. Estas son las variables que más predicen el éxito de la campaña: **duración, balance y edad**.

¿Qué valores de estas variables predicen el éxito?

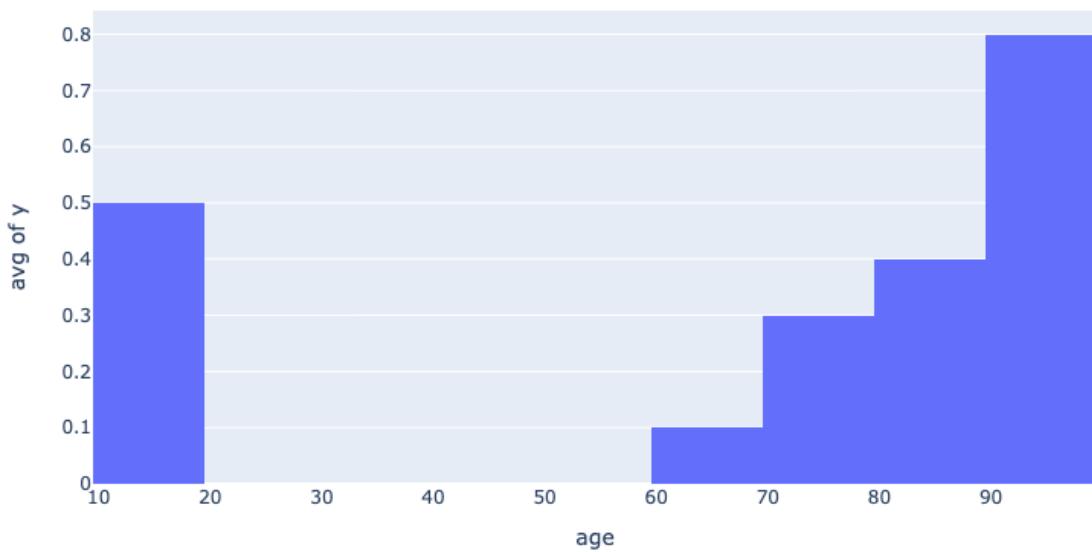
La última pregunta a responder es, ¿qué valores de estas variables predicen el éxito? Con esta información podemos optimizar las ventas para maximizar el número de conversiones.



Este gráfico muestra cómo responde la persona para diferentes duraciones. Cuando la y vale 0, el cliente ha comprado el depósito, por tanto, **cuanto menor, mejor**. Podemos ver entonces que, para duraciones entre 0 y 500 segundos, y entre 3200 y 3800, la conversión es mayor. A partir de los 500 segundos, es más improbable que el cliente compre el depósito a no ser que el cliente se mantenga en línea hasta los 3200. A partir de los 3800, el cliente se puede dar por perdido.



En cuanto al balance, personas con menos de 70.000€ medios anuales tienen altas probabilidades de contratar el depósito, sobre todo las personas en el rango entre 60-70k€.

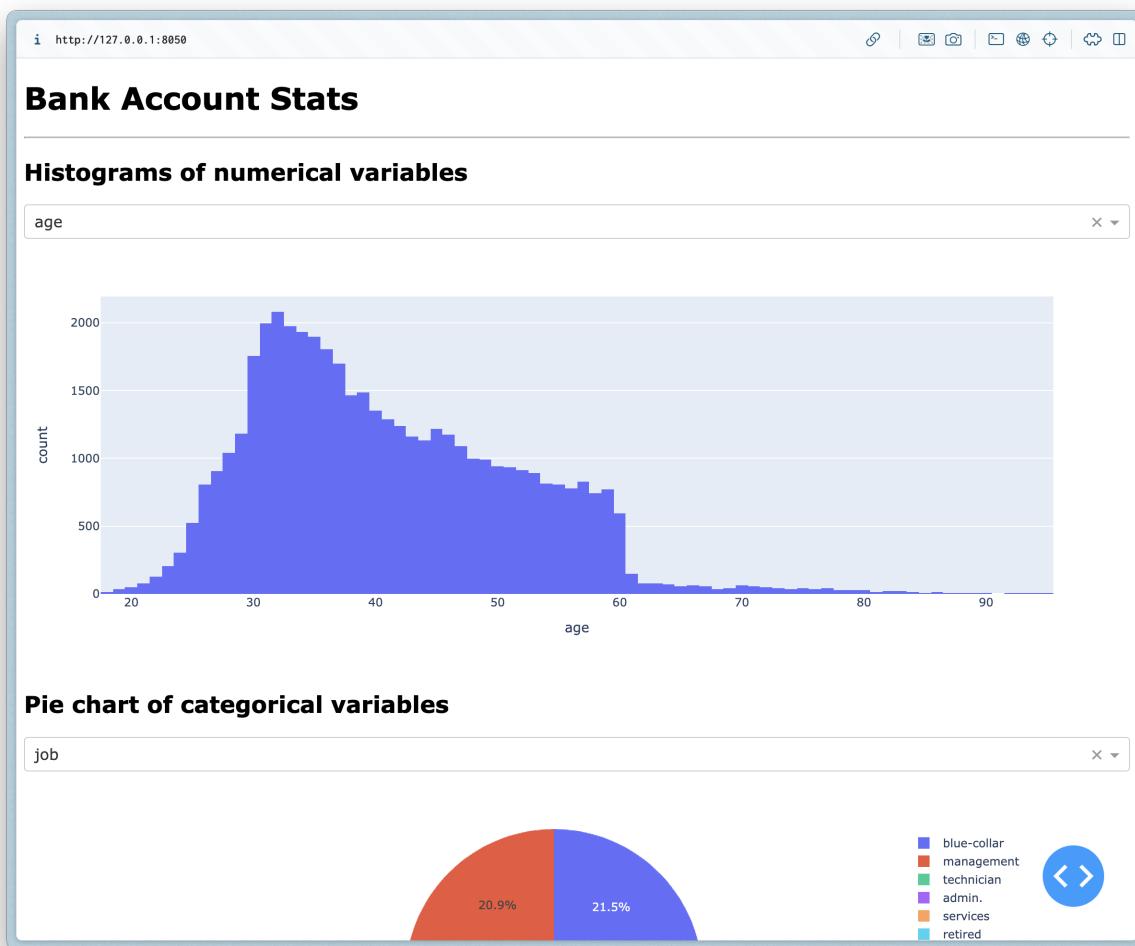


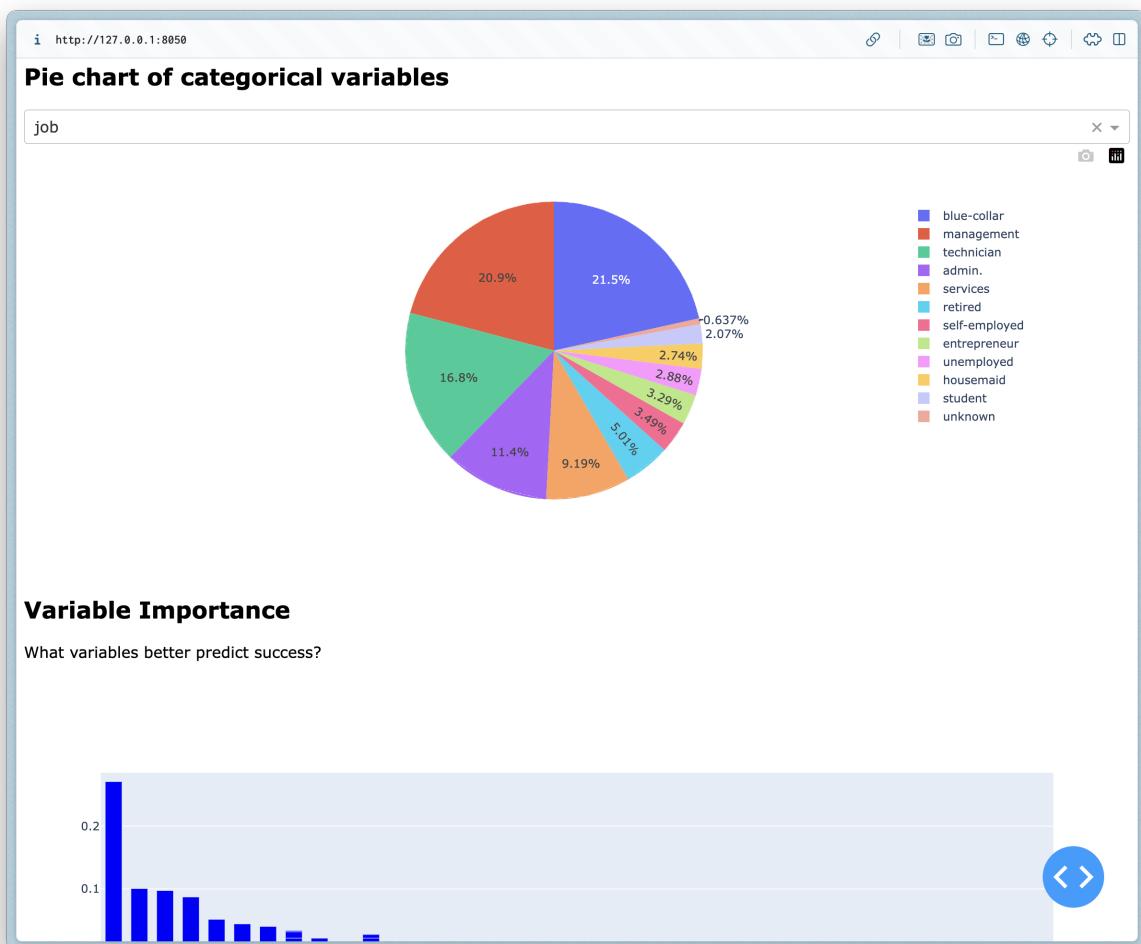
Finalmente, los grupos de edad que más contratan el depósito son las personas entre 20 y 60 años. A partir de los 60 años cada vez es menos probable que contraten el depósito.

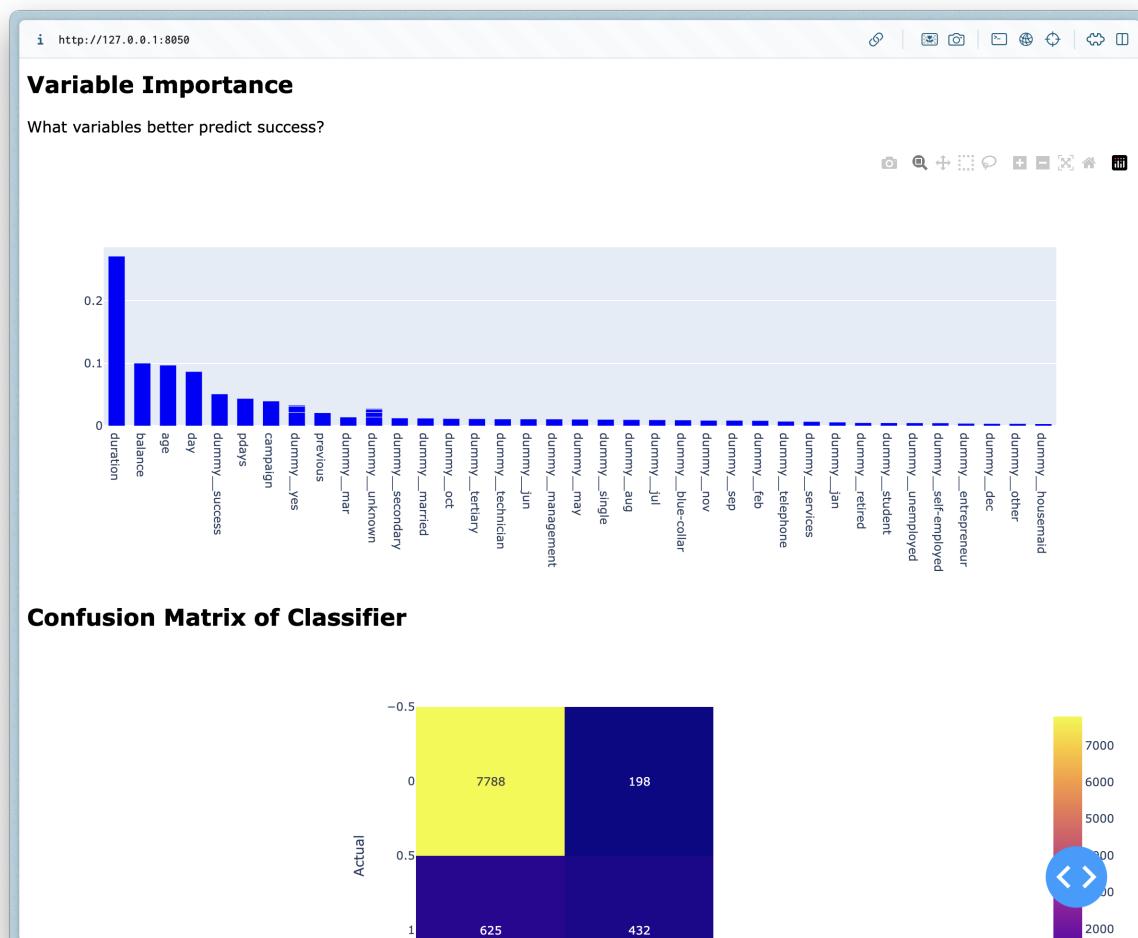
Conclusión

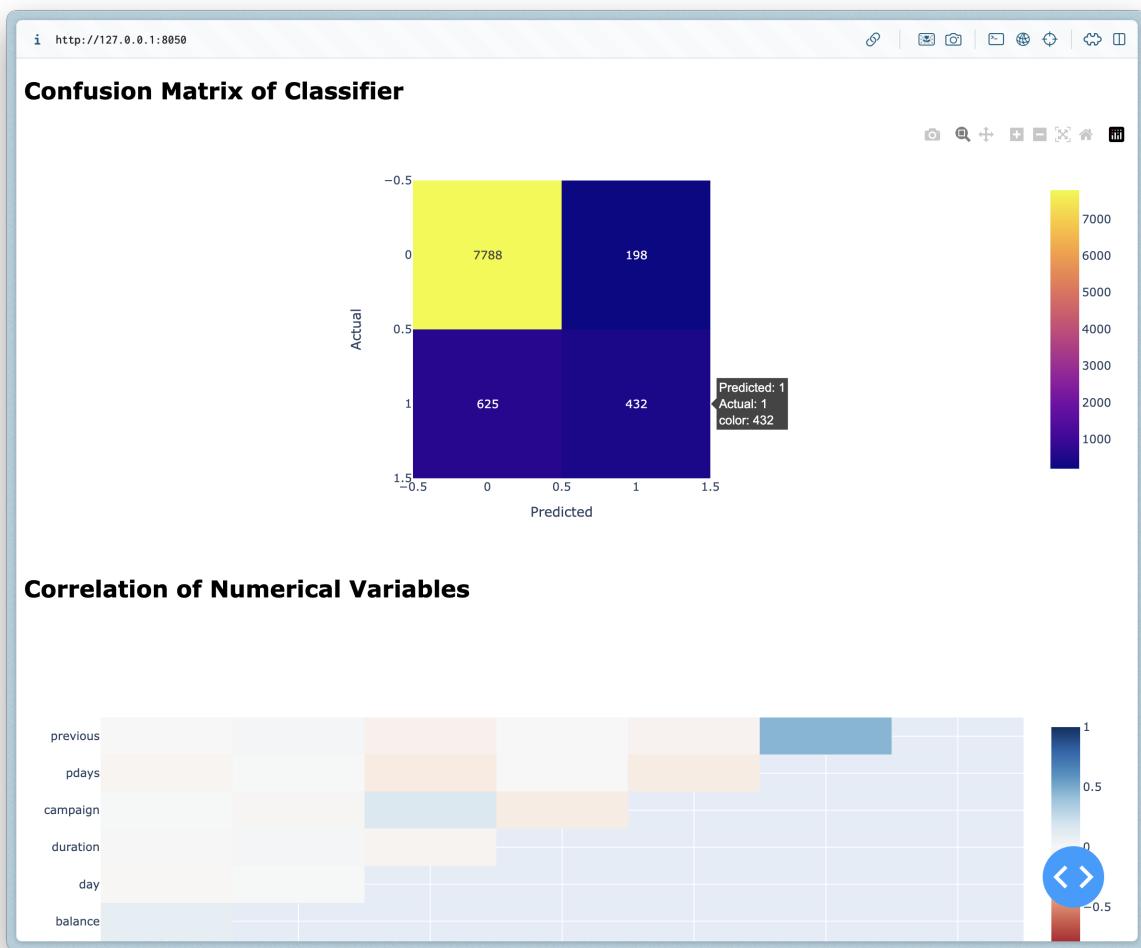
Para maximizar las ventas hay que centrarse en conversaciones cortas (menos de 500 segundos), con personas de menos de 70.000€ anuales y que estén entre los 20 y 60 años.

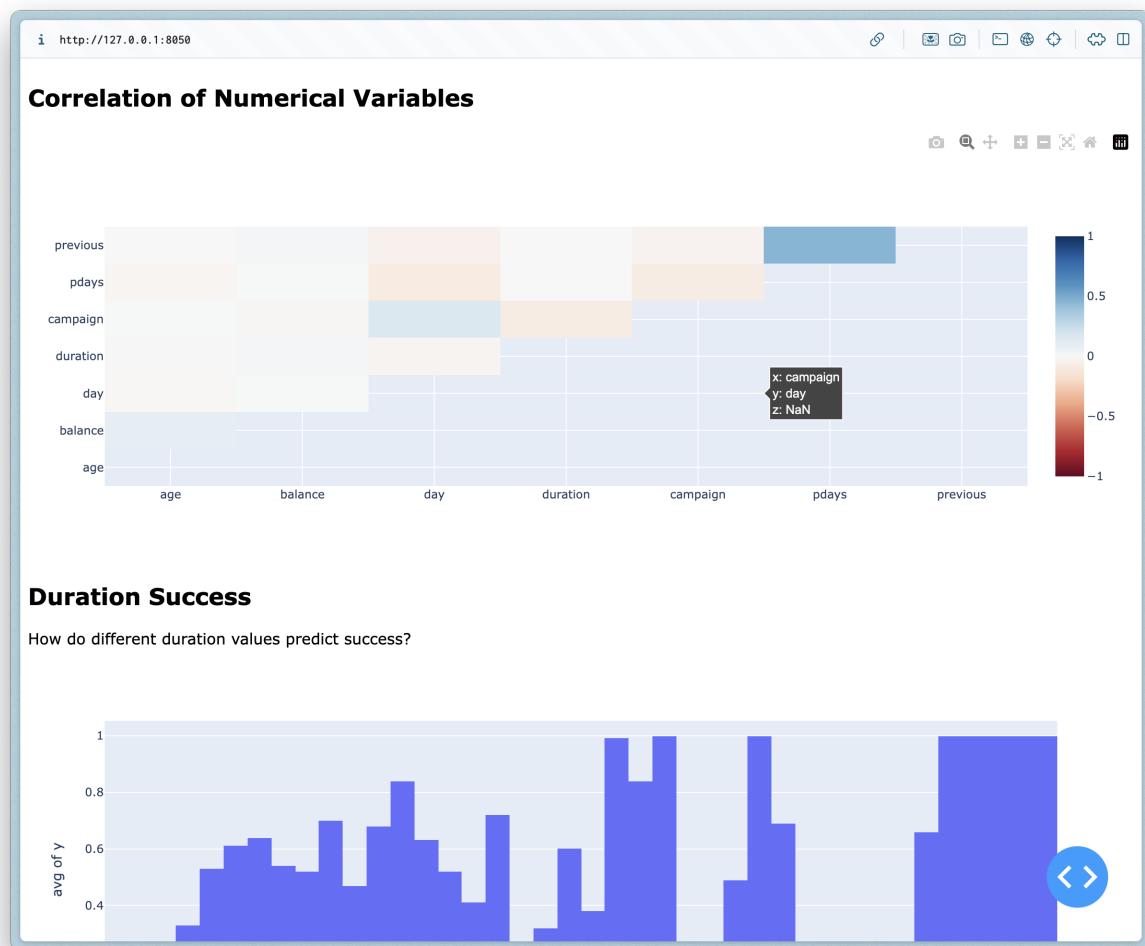
Capturas del Dash

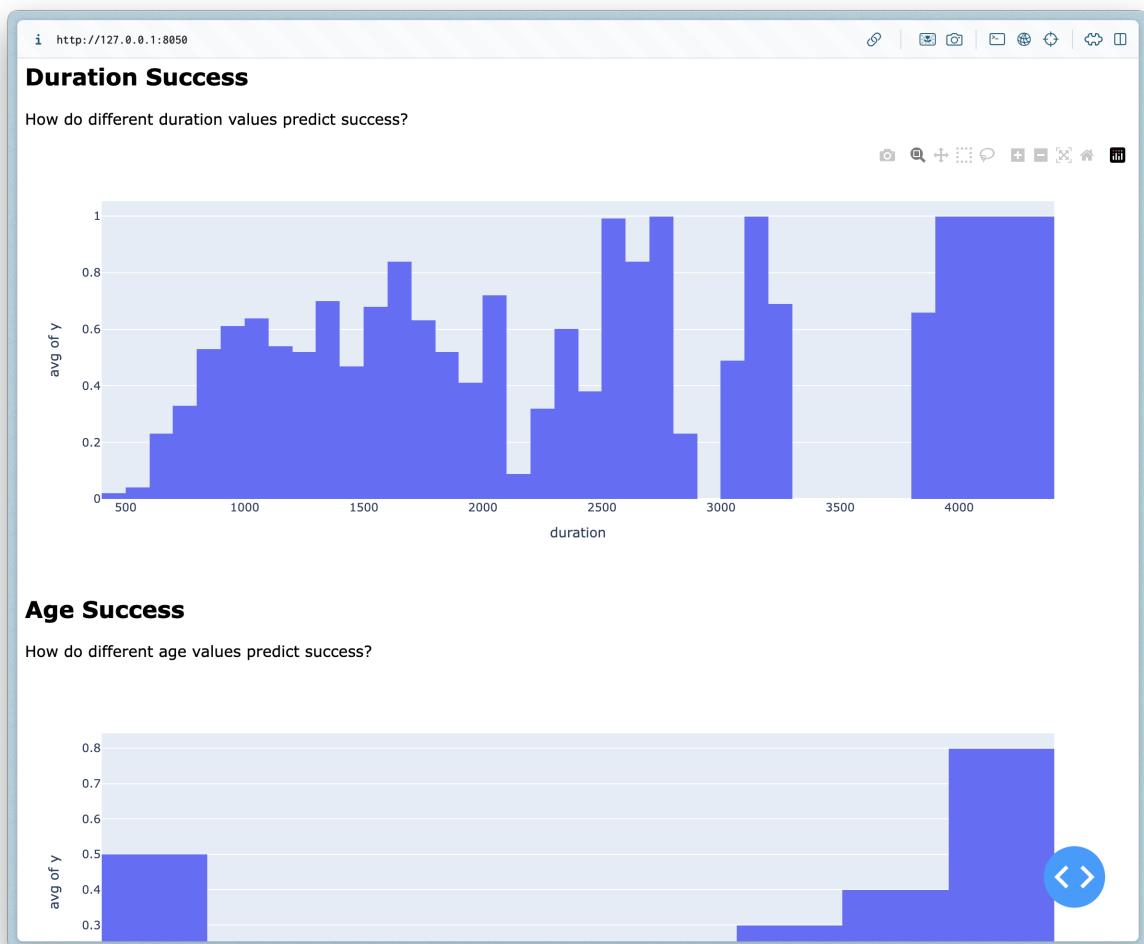


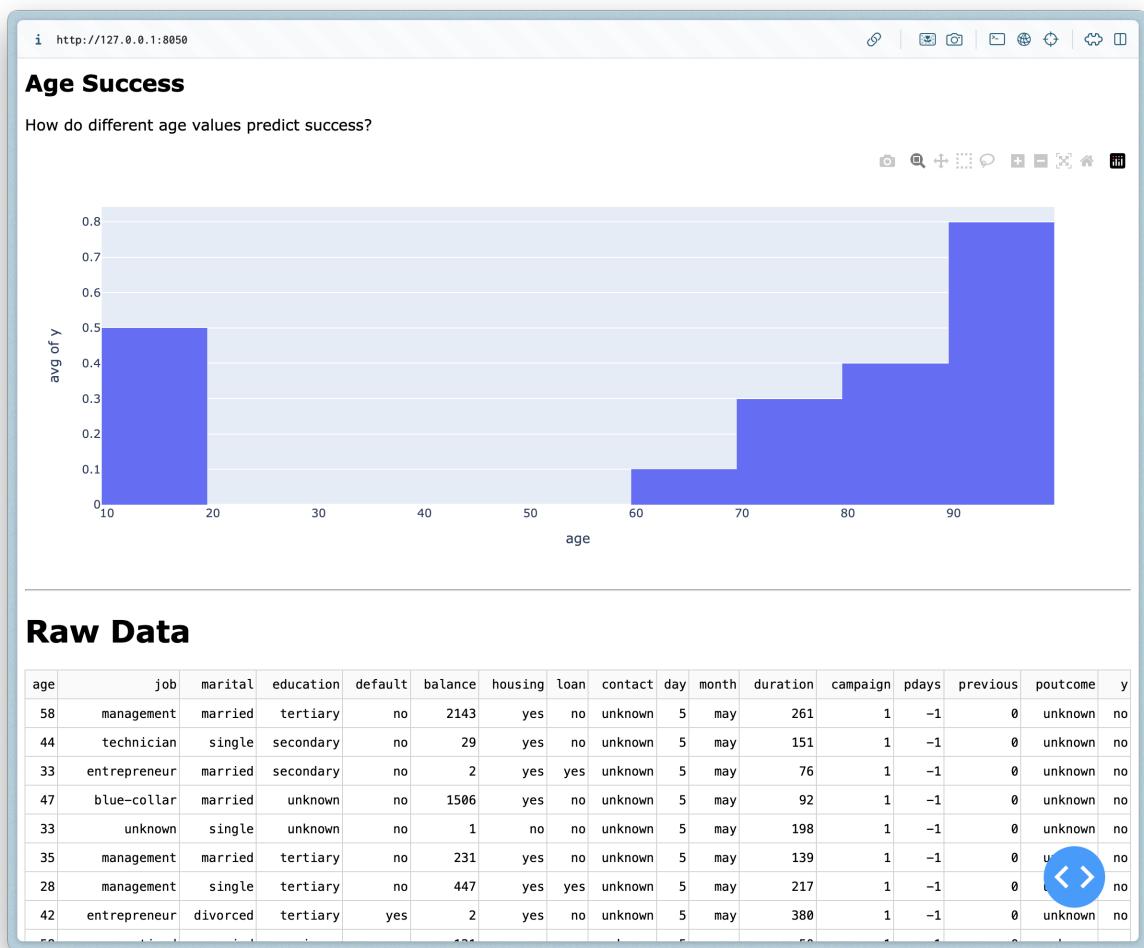












i http://127.0.0.1:8050

Raw Data

age	job	marital	education	default	balance	housing	loan	contact	day	month	duration	campaign	pdays	previous	poutcome	y
58	management	married	tertiary	no	2143	yes	no	unknown	5	may	261	1	-1	0	unknown	no
44	technician	single	secondary	no	29	yes	no	unknown	5	may	151	1	-1	0	unknown	no
33	entrepreneur	married	secondary	no	2	yes	yes	unknown	5	may	76	1	-1	0	unknown	no
47	blue-collar	married	unknown	no	1506	yes	no	unknown	5	may	92	1	-1	0	unknown	no
33	unknown	single	unknown	no	1	no	no	unknown	5	may	198	1	-1	0	unknown	no
35	management	married	tertiary	no	231	yes	no	unknown	5	may	139	1	-1	0	unknown	no
28	management	single	tertiary	no	447	yes	yes	unknown	5	may	217	1	-1	0	unknown	no
42	entrepreneur	divorced	tertiary	yes	2	yes	no	unknown	5	may	380	1	-1	0	unknown	no
58	retired	married	primary	no	121	yes	no	unknown	5	may	50	1	-1	0	unknown	no
43	technician	single	secondary	no	593	yes	no	unknown	5	may	55	1	-1	0	unknown	no
41	admin.	divorced	secondary	no	270	yes	no	unknown	5	may	222	1	-1	0	unknown	no
29	admin.	single	secondary	no	390	yes	no	unknown	5	may	137	1	-1	0	unknown	no
53	technician	married	secondary	no	6	yes	no	unknown	5	may	517	1	-1	0	unknown	no
58	technician	married	unknown	no	71	yes	no	unknown	5	may	71	1	-1	0	unknown	no
57	services	married	secondary	no	162	yes	no	unknown	5	may	174	1	-1	0	unknown	no
51	retired	married	primary	no	229	yes	no	unknown	5	may	353	1	-1	0	unknown	no
45	admin.	single	unknown	no	13	yes	no	unknown	5	may	98	1	-1	0	unknown	no
57	blue-collar	married	primary	no	52	yes	no	unknown	5	may	38	1	-1	0	unknown	no
60	retired	married	primary	no	60	yes	no	unknown	5	may	219	1	-1	0	unknown	no
33	services	married	secondary	no	0	yes	no	unknown	5	may	54	1	-1	0	unknown	no
28	blue-collar	married	secondary	no	723	yes	yes	unknown	5	may	262	1	-1	0	unknown	no
56	management	married	tertiary	no	779	yes	no	unknown	5	may	164	1	-1	0	unknown	no
32	blue-collar	single	primary	no	23	yes	yes	unknown	5	may	160	1	-1	0	unknown	no
25	services	married	secondary	no	50	yes	no	unknown	5	may	342	1	-1	0	unknown	no
40	retired	married	primary	no	0	yes	yes	unknown	5	may	181	1	-1	0	unknown	no
44	admin.	married	secondary	no	-372	yes	no	unknown	5	may	172	1	-1	0	unknown	no
39	management	single	tertiary	no	255	yes	no	unknown	5	may	296	1	-1	0	unknown	no