

EDA CON PYTHON

PROYECTO EDUCATIVO
THEPOWER

Julia Pirogova

Tabla de contenido

<u>INTRODUCCIÓN</u>	<u>2</u>
<u>ANÁLISIS DESCRIPTIVO</u>	<u>3</u>
<u>CLIENTES:</u>	<u>3</u>
<u>CAMPAÑA:</u>	<u>8</u>
<u>FACTOR TIEMPO</u>	<u>12</u>
<u>MARCADORES MACROECONÓMICOS</u>	<u>14</u>
<u>CONCLUSIONES FINALES</u>	<u>15</u>

INTRODUCCIÓN

1. Objetivo de trabajo:

Hemos recibido varias bases de datos sobre los clientes del banco (43 mil personas) y la campaña telefónica realizada durante varios años. El final de este análisis era recopilar los datos, hacer la limpieza para poder sacar la mayor información y llegar a las posteriores conclusiones sobre el éxito del producto ofrecido y, además, ver que ha salido bien y lo que no tanto para poder enfocar mejor las próximas campañas de marketing.

2. Descripción de los datos:

Se utilizaron los archivos bank-additional.csv y customer-details.xlsx. Tras la carga, se revisaron tipos de datos, valores faltantes, duplicados y coherencia de categorías.

3. Perfil cliente promedio:

	Variable	Media	Mediana
Edad		39.98	38.00
Duración de la llamada		257.74	179.00
Número de contactos		2.57	2.00
Días entre contactos		6.07	6.00
Renta (renta)		93241.20	93050.50
Niños en casa (Kidhome)		1.00	1.00
Adolescentes en casa (Teenhome)		1.00	1.00
Visitas web al mes (NumWebVisitsMonth)		16.59	17.00
Trabajo :	administrativo		
Estado civil :	casado		
Educación:	título universitario		
Incumplimiento de pagos:	no		
Hipoteca:	sí		
Préstamo personal:	no		

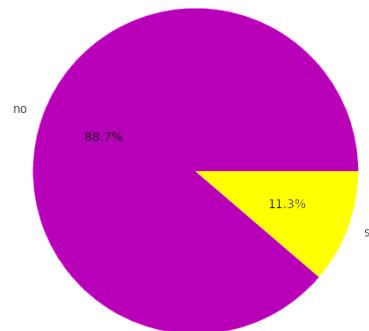
Análisis descriptivo

En primer lugar, y tras haber realizado un análisis de los bases de datos proporcionados, vemos en el gráfico que la campaña no ha tenido gran éxito: solo alrededor del 11% de los clientes han adquirido el producto. Aunque este tipo de marketing no suele ser el más efectivo, teniendo en cuenta que se trataba de los clientes habituales del banco, se podría esperar mejor respuesta. Por esta razón, me parece necesario hacer un análisis más profundo de la situación y entender que se podría mejorar para la próxima campaña.

Para el posterior análisis, me gustaría dividir la información en varias secciones:

- Datos de los clientes
- Datos de la campaña
- Factor tiempo
- Marcadores macroeconómicos

Distribución de Éxito de la Campaña

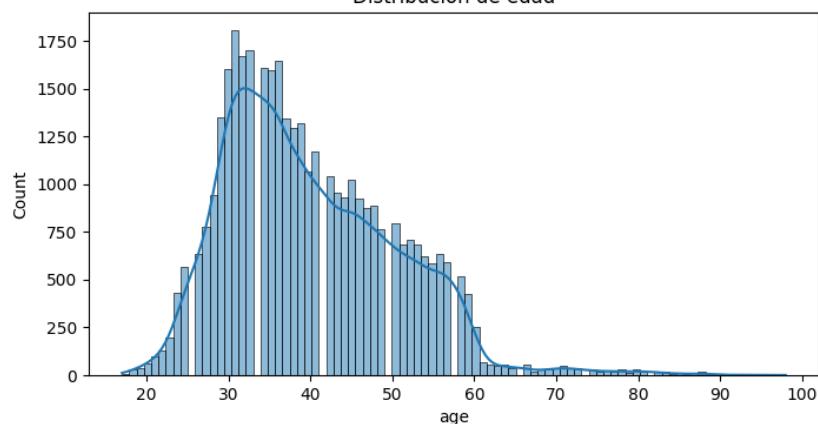


CLIENTES:

EDAD

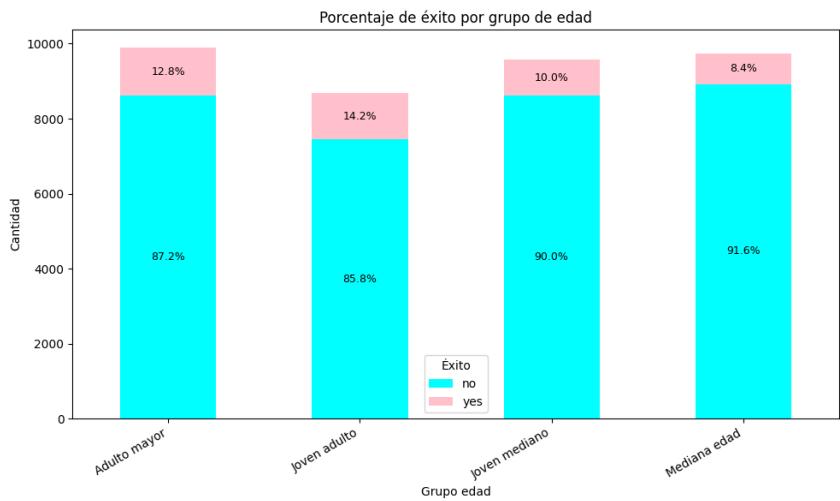
Nuestro cliente promedio tiene unos 40 años, pero vemos que el rango de edades va de 17 a 98 años, o sea, que tanto trabajamos con

Distribución de edad



gente muy joven como con las personas de avanzada edad. Me ha parecido interesante para el análisis dividir a los clientes en 4 grupos de edad según los cuartiles:

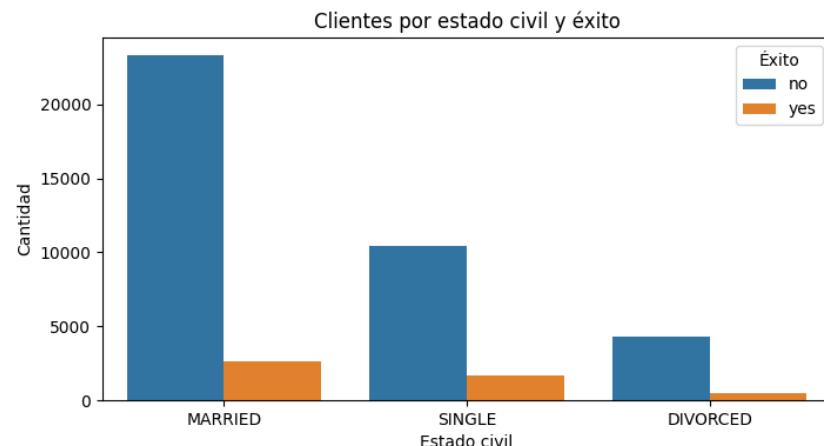
- Hasta 32 años – Joven Adulto
- De 32 a 38 – Joven Mediano
- De 38 a 47 – Mediana Edad
- A partir de 47 años – Adulto Mayor



Según vemos en el gráfico, los grupos donde tenemos la mayor posibilidad de obtener un "sí" son Joven Adulto (un 14,2%) y Adulto Mayor, (un 12,8%) mientras las edades de 32 a 47, es decir, la edad mediana, son los que más han rechazado nuestra oferta. Habrá que ver cómo mejoramos su respuesta en el futuro, puesto que es un grupo amplio y es de vital importancia capturar su atención.

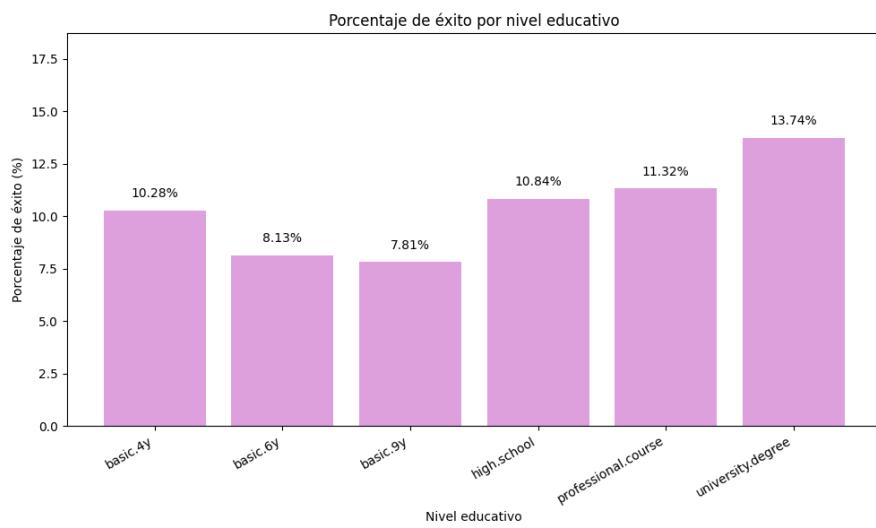
ESTADO CIVIL E HIJOS

La inmensa mayoría de nuestros clientes están casados (un hecho a tomar en cuenta para las próximas campañas), y



en general, se podría decir que hay muy poca diferencia en la aceptación del

producto según el estado civil. Se observa que los solteros suelen ser más positivos hacia la propuesta, pero la diferencia es muy poca, apenas un 1-2%. Asimismo, los hijos, independiente si son pequeños o adolescentes, no suelen ser un factor importante para la campaña, ya que los datos muestran que se observa el mismo porcentaje de éxito de un 11%.



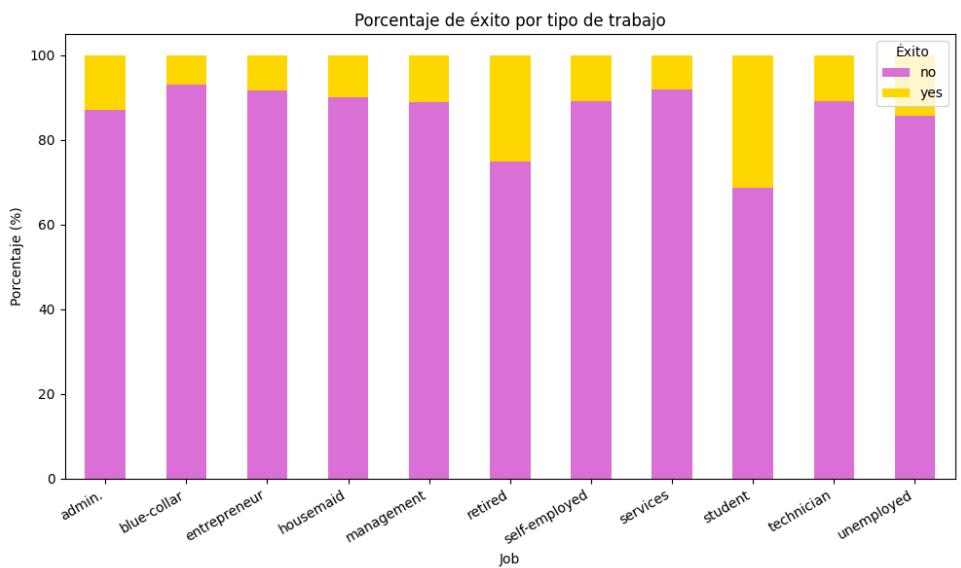
de que cuanto mayor es el nivel de estudios, mejor respuesta se obtiene para nuestra propuesta (13,7% para los universitarios frente a alrededor de 8% de educación básica). Es un factor a tener en cuenta. A lo mejor sería interesante ver otras maneras de motivar a las personas con menos estudios u otras formas de trasmitir nuestro mensaje.

PROFESIÓN

A primera vista, se podría concluir que los puestos de trabajo que han contratado el producto con mayor frecuencia son los jubilados (25%) y los

EDUCACIÓN

Para empezar, cabe destacar que he tomado la decisión de eliminar a los clientes analfabetos, ya que no es un grupo representativo (eran solo 4 personas). Claramente, podemos ver la tendencia



estudiantes (31%). No obstante, deberíamos tener en cuenta que no son grupos muy extensos, por lo tanto, los datos no mostrarán la situación exacta (aunque está muy bien saber que la campaña ha sido un éxito para estos dos grupos y podría ser interesante para la empresa buscar capturar más clientes en estos sectores de población).

Por otra parte, el porcentaje de aceptación entre los trabajadores manuales (“blue collar”) es realmente bajo (solo 6,9%) , seguidos de empleados del sector de servicios (un 8%) y empresarios (un 8,3% (aunque también es un grupo más pequeño)).

SITUACIÓN ECONÓMICA

Aquí vamos a incluir varios factores (renta, hipoteca, prestamos), ya que realmente han tenido muy poca relevancia para el éxito de la campaña.

	exito	no	yes	total	porcentaje_exito
grupo_renta					
renta alta	9371	1213	10584		11.460695
renta baja	9392	1200	10592		11.329305
renta mediana	9644	1248	10892		11.457951
renta muy alta	9749	1183	10932		10.821442

En cuanto a la renta, me ha parecido relevante dividir la gente en 4 grupos:

- Menos de 49000 euros - renta baja
- De 49000 a 93000 euros – renta mediana
- De 93000 a 136000 euros – renta alta
- Más de 136000 euros - renta muy alta

Como vemos en la tabla, no hay mucha diferencia entre las cuatro grupos y todos tienen el porcentaje de éxito medio de unos 11%.

Los factores como la hipoteca y préstamo tampoco han tenido influencia en la campaña, los porcentajes del éxito rondan unos 11% (los que tienen hipoteca han contratado el producto un 0,5% más, pero me parece muy poca la diferencia).

Hablando de los clientes con impagos, es un grupo tan pequeño (solo 3 personas) que no es un grupo para examinar para este estudio.

CONCLUSIONES CLIENTES

Tras analizar las características demográficas, educativas, profesionales y económicas de los clientes, se observa que el éxito de la campaña varía según determinados factores, mientras que otros muestran una influencia bastante limitada en la respuesta de los clientes.

En relación con la edad, se muestra que los clientes más jóvenes y los de mayor edad son los más propensos a aceptar la oferta, mientras que el grupo intermedio (32 a 47 años) presenta menor receptividad. Dado que este grupo representa una parte importante de la base de clientes, será necesario diseñar estrategias más enfocadas en captar su atención.

En cuanto al estado familiar, no vemos gran relevancia para la decisión final. A pesar de que los solteros muestran una ligera mayor aceptación, la diferencia es mínima. Asimismo, los hijos no cambian el resultado final.

El nivel educativo, por el contrario, sí marca diferencias. Hay una tendencia clara: a mayor nivel de estudios, mayor probabilidad de contratar el producto. Los universitarios muestran un rendimiento significativamente superior frente a aquellos con educación básica.

Respecto a la profesión, ciertos grupos (como jubilados y estudiantes) presentan tasas de éxito muy altas, aunque su representatividad dentro de la base de datos es baja, por lo que estos porcentajes deben interpretarse con cuidado. Por otra parte, profesiones como trabajadores manuales y empleados del sector servicios muestran bajos niveles de aceptación. Se necesitaría estudiar el porqué de su rechazo.

Por último, los factores relacionados con la situación económica (renta, hipoteca, préstamos) no parecen influir significativamente en la contratación.

Para resumir, el análisis afirma que el nivel educativo y la edad son los factores que mayor peso tienen en la aceptación del producto, mientras que la situación económica, el estado civil y la presencia de hijos apenas influyen en la respuesta. Esta información permite orientar futuras campañas

hacia segmentos con mayor potencial y ajustar la comunicación según las características reales del cliente.

CAMPAÑA:

CONTACTOS ANTERIORES

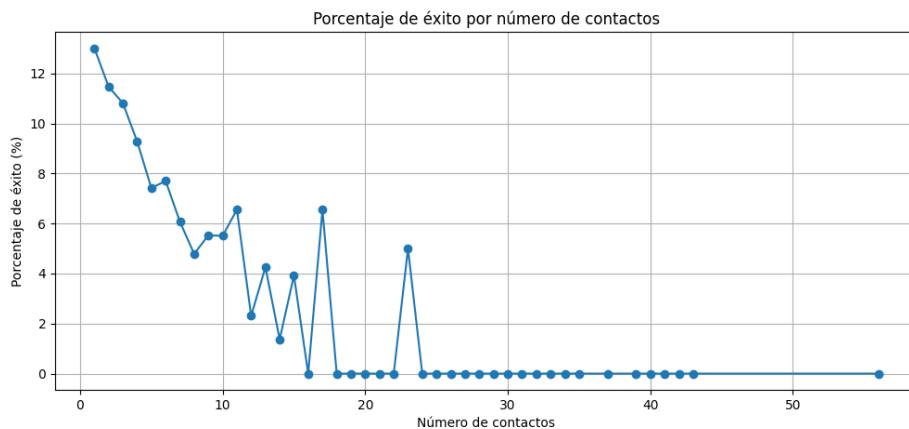
Como podemos contemplar en la tabla, si contactamos con un cliente por primera vez, la probabilidad del éxito es muy baja, solo unos 8%, pero a medida que vayamos conociéndonos, se ve

que el porcentaje del éxito sube mucho: a segundo contacto es de 21%, el tercero el 46% y el cuarto ya supera los 60%. Todo esto indica la gran importancia de la comunicación y de interacción con los clientes del banco. Cuanto mejor relación tengamos, mayor confianza generaremos.

ncontactosantes	exito	no	yes	total	porcentaje_exito
0	33832	3271	37103		8.815999
1	3770	1008	4778		21.096693
2	420	366	786		46.564885
3	92	140	232		60.344828
4	34	41	75		54.666667
5	5	15	20		75.000000
6	2	3	5		60.000000
7	1	0	1		0.000000

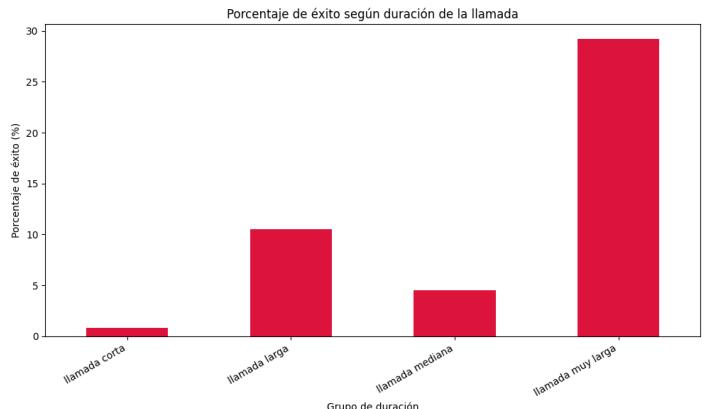
NÚMERO DE CONTACTOS

El gráfico muestra que el mayor porcentaje de éxitos se obtiene durante los 2-3 contactos en esta campaña y luego los números van empeorando. Es decir, aunque insistimos, no obtenemos mejores ventas del producto. Por lo tanto, se debería darles instrucciones al personal de no hacer más de 3 contactos la mayoría de veces, porque no suelen ser efectivos.



DÍAS ENTRE CONTACTOS

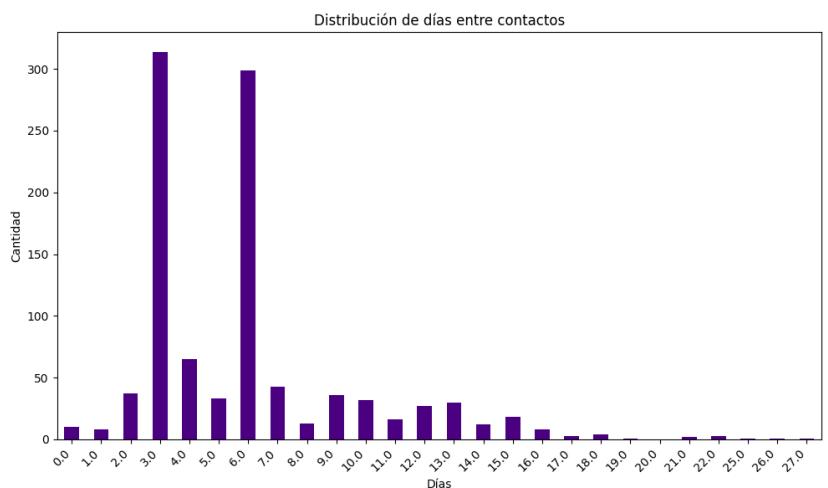
Cuando tememos que volver a llamar a los clientes, la tendencia es que la mejor opción es dejarles unos días para pensar (vemos que en los primeros días el porcentaje del éxito es muy bajo). Así notamos que los resultados más exitosos se obtienen si han pasado entre 3 y 6 días entre las llamadas. Esperar mucho más de una semana tampoco resulta efectivo.



DURACIÓN DE LA LLAMADA

Durante el análisis de datos, hemos observado que este factor sí que tiene una gran relevancia para el éxito. Para analizar el factor, he vuelto a dividir los datos en 4 grupos:

- Menos de 100 segundos – llamada corta
- De 100 a 180 segundos – llamada mediana
- De 180 a 320 segundos – llamada larga
- A partir de 320 segundos llamada muy larga



En primer lugar, la tendencia clara es que a mayor duración de la llamada, mayor es la probabilidad de obtener la respuesta afirmativa. En llamadas largas el éxito sube a 10.5%, y las llamadas muy largas dan fruto en 29% de los casos. Sin embargo, las llamadas de menos de 1 minuto y 40 segundos apenas dan resultado: solo 0.78% de "sí". Es decir, tenemos casi 30 veces más

posibilidades de éxito en una llamada muy larga, que permite responder las dudas de los clientes y dar mayor información sobre el producto.

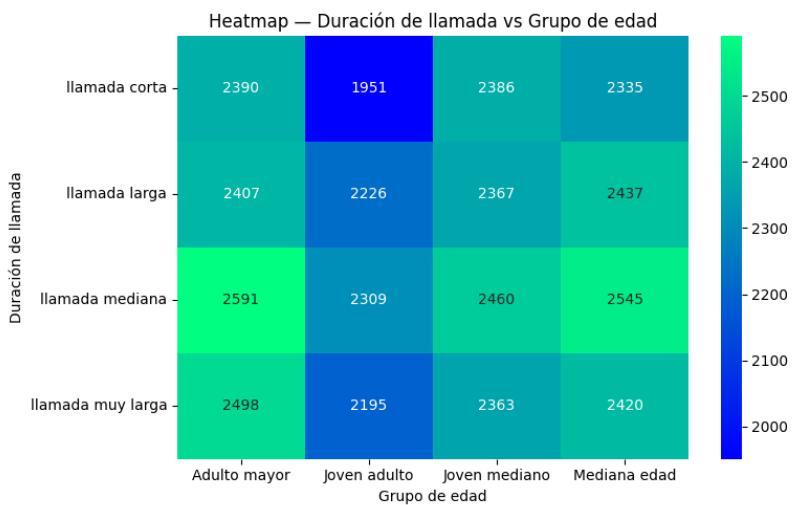
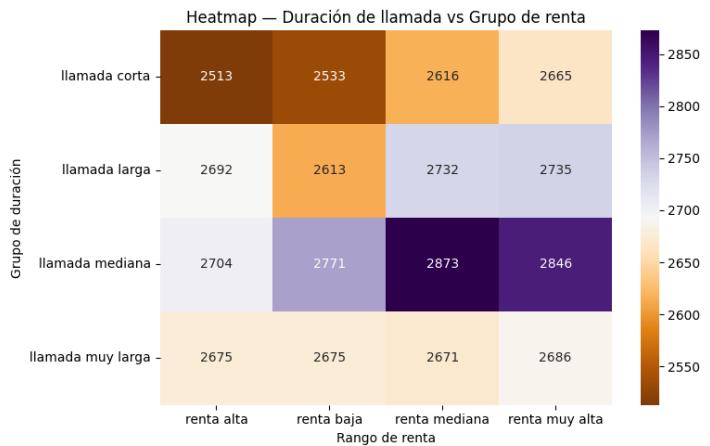
Teniendo en cuenta que se trata de un factor clave para el análisis, he decidido hacer varias comparaciones para entender mejor cómo obtenemos que las llamadas sean más largas.

DURACIÓN LLAMADA - EDAD

En el heatmap se ve que los de Mediana Edad son los que suelen tener conversaciones más largas (puede ser que sean los más acostumbrados a usar los teléfonos para recibir la información) mientras que los más jóvenes tienden a tener menos llamadas, porque, posiblemente, prefieren los métodos digitales (redes sociales y messengeres) para ser contactados. Es un dato interesante para tener en cuenta para las próximas campañas.

DURACIÓN LLAMADA – RENTA

En esta comparación vemos que la duración de las llamadas es prácticamente igual para todos los niveles de renta y se concentra mayoritariamente. Del mismo modo, tras llevar a cabo un análisis similar por profesiones o estado civil no se detecta una correlación entre la duración de la conversación telefónica y el factor analizado. Cabe destacar, que la duración más común es llamada mediana, que podría indicar que los operadores no suelen cambiar el mensaje según con quien hablan que sería una desventaja para la campaña.



CONCLUSIONES CAMPAÑA

Despues de ver los datos relacionados con la campaña, se observan varios patrones relevantes sobre el comportamiento de los clientes, que permite identificar varios factores que tienen mucha influencia en el éxito.

En primer lugar, observamos que la relación previa con el cliente es determinante: cuantos más contactos anteriores haya habido, mayor es la probabilidad de obtener una respuesta afirmativa. Esto demuestra que la confianza y la interacción continuada con el cliente son fundamentales para construir una relación que facilite la conversión.

No obstante, el número total de contactos dentro de una misma campaña tiene un punto óptimo. Los mejores resultados se concentran entre el segundo y el tercer contacto; a partir de ahí, insistir no mejora el rendimiento. Este hallazgo sugiere que es necesario optimizar los esfuerzos del equipo comercial y evitar un exceso de llamadas que no aportan valor.

En cuanto al tiempo entre contactos, los datos indican que las llamadas realizadas tras 3 a 6 días de espera ofrecen las mayores tasas de éxito. Llamar demasiado pronto reduce la efectividad, mientras que esperar más de una semana tampoco genera mejoras significativas.

El factor mde mayor peso ha sido la duración de la llamada. A mayor duración, mayor probabilidad de éxito: las llamadas muy largas multiplican por casi treinta la tasa de conversión respecto a las llamadas cortas. Esto evidencia la importancia de dedicar tiempo a explicar con detalle el producto, resolver dudas y generar confianza. Por tanto, la calidad de la conversación es un elemento clave para el cierre de ventas. La edad podría ayudar para determinar la conversación, mientras otros factores como renta o estado civil no muestran una relación importante.

En conjunto, el análisis permite concluir que las estrategias de contacto deben priorizar la relación continua, el momento adecuado y la profundidad de la conversación, enfocándose especialmente en los segmentos más abiertos a interacciones telefónicas. Estos hallazgos ofrecen una base sólida

para optimizar futuras campañas y mejorar las tasas de conversión del banco. Creo que sería buena idea preparar una pequeña guía para las próximas campañas.

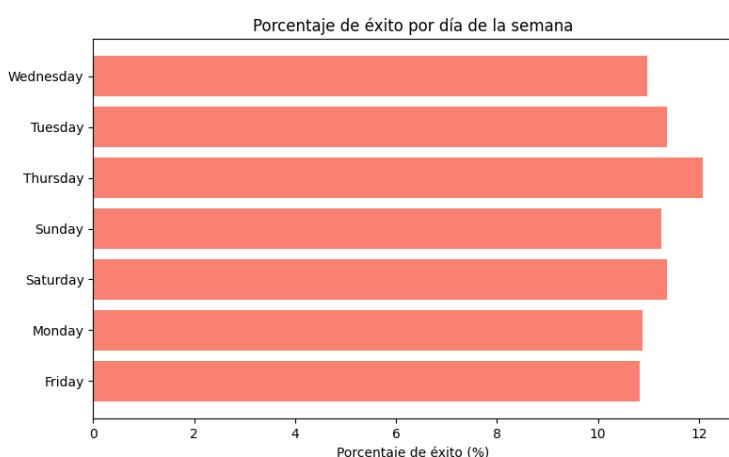
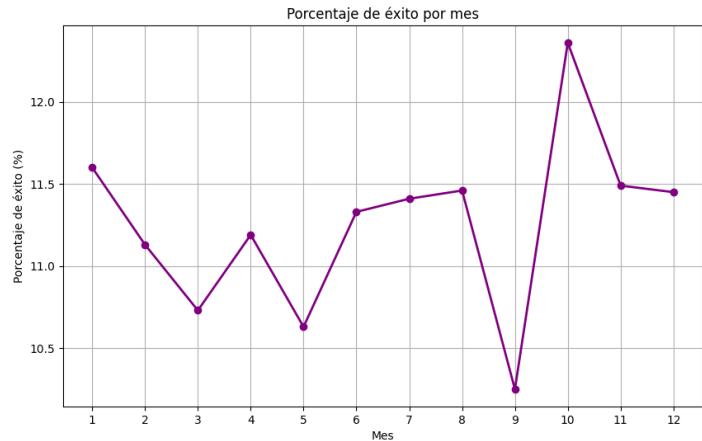
FACTOR TIEMPO

Otro factor que me ha parecido interesante estudiar es si el periodo exacto de la llamada tiene alguna influencia en el resultado final de la campaña.

Para empezar, los datos indican que el año en el que se producía la llamada no influía mucho en la respuesta de los clientes, ya que todos los años muestran la media de unos 11%.

ÉXITO POR MES

En cuanto hablamos de la diferencia por meses, tampoco se ve gran variación de los resultados. Sin embargo, hay una pequeña caída en septiembre (baja a 10,25%) que posteriormente repunta en octubre (sube a 12,36%). Podría ser un error en estadísticas, ya que se trata de dos meses consecutivos, pero se debería comprobar este punto.



ÉXITO POR DÍA DE SEMANA

Según nos muestran los datos, el mejor día para hacer campaña es el jueves (éxito más de 12%), seguido por martes (11,37%), sábado (11,37%) y domingo (11,25%). Estos últimos no suelen ser días muy habituales en las

campañas telefónicas, pero, sin embargo, vemos que la estrategia ha funcionado bastante bien.

Por otra parte, el día con los peores resultados ha sido el viernes (10,8%). Esto podría indicar que la gente suele estar ocupada el último día laborable terminando las tareas y planificando el fin de semana, por lo tanto, se debería evitar llamar el viernes.



CONCLUSIONES FACTOR TIEMPO

El análisis del factor tiempo demuestra que, en general, el tiempo en el que se realiza la llamada no influye de manera determinante en el éxito de la campaña, ya que las variaciones observadas entre años o meses son muy pequeñas y se mantienen en torno al 11% de éxito. Sin embargo, se podría identificar algunos patrones que pueden ayudar a optimizar la estrategia comercial.

Por un lado, el año y el mes parecen carecer de peso para determinar el resultado final de la campaña, puesto que casi no se ve variación.

No obstante, el día de la semana, sí se debería tomar en cuenta. El jueves destaca como el día más eficiente para contactar con los clientes, con un porcentaje de éxito superior al 12%. También sorprende el buen rendimiento de las llamadas realizadas en sábado, domingo y martes, situándose todas por encima del 11%, lo cual sugiere que los días no laborables pueden ser efectivos cuando la disponibilidad del cliente es mayor.

Por otro lado, el viernes es el día con peores resultados (10,8%), lo que indica que las personas probablemente estén más ocupadas cerrando asuntos laborales, disminuyendo así su receptividad. Este hallazgo señala que debería

evitarse programar llamadas comerciales en viernes, ya que reduce la eficiencia de la campaña.

En resumen, aunque el factor tiempo no es determinante a gran escala, optimizar los días de contacto podría mejorar el rendimiento, priorizando jueves, martes y fines de semana, y evitando el viernes. Este ajuste podría aumentar la efectividad global de la campaña.

MARCADORES MACROECONÓMICOS

El último pero no menos importante factor que vamos a analizar en este trabajo son los marcadores macroeconómicos que muestran una influencia bastante evidente en el comportamiento de los clientes.

exito	emp.var.rate	cons.price.idx	cons.conf.idx	euribor3m	nr.employed
no	0.244499	93.602360	-40.603748	3.810591	5176.032480
yes	-1.241247	93.352406	-39.763109	2.107420	5094.543064

emp.var.rate – Tasa de variación del empleo

Los clientes que dieron la respuesta favorable fueron contactados en momentos donde la tasa de variación del empleo era negativa (-1.24), o sea, cuando el mercado laboral estaba en peor situación.

cons.price.idx – Índice de precios

La diferencia es muy pequeña (93.60 vs. 93.35), por lo que se deduce que el nivel de precios no influyó en el éxito de la campaña.

cons.conf.idx – Confianza del consumidor

Los clientes que dijeron “sí” tienen una confianza del consumidor ligeramente mejor (-39.76 vs -40.60). Tiene logica que los “más optimistas” acepten mejor el producto. Aunque hay que decir que no es una tendencia muy fuerte, ya que los datos no varian mucho.

euribor3m – Interés a 3 meses

Se nota una diferencia muy clara (3,81% vs 2,1%). Esto nos indica el patrón de que cuando los tipos de interés son más bajos, los clientes están más abiertos a contratar productos bancarios.

nr.employed – Número de empleados

Los clientes con “sí” fueron contactados en contextos económicos con mayor paro. Cuando hay menos posibilidad de empleo, la gente suele reaccionar mejor a nuestro producto.

CONCLUSIONES FACTORES MACROECONÓMICOS

El análisis muestra que existe una relación entre muchas variables económicas y el éxito de la campaña. Tales factores como tasa de empleo, euribor y confianza del consumidor destacan que cuando el mercado es inestable, surge mayor interés hacia el producto que ofrecemos. Es decir, la campaña es más exitosa con el mercado con condiciones menos favorables.

Creo que es un hecho que se debería tomar muy en serio para las promociones futuras, ya que podría aumentar bastante el éxito de las campañas.

CONCLUSIONES FINALES

Tras analizar los datos referentes a la campaña realizada y teniendo en cuenta que los resultados no han sido los más favorables, se podría llegar a las siguientes conclusiones.

El éxito de la campaña depende mucho de la edad y el nivel educativo de los clientes, siendo los más jóvenes, mayores y aquellos con estudios universitarios los más receptivos. Factores como estado civil, presencia de hijos y situación económica han tenido poca influencia.

La relación previa con el cliente es clave: los contactos continuos generan confianza y aumentan la probabilidad de aceptación. El número óptimo de contactos es entre dos o tres, y las llamadas más largas incrementan significativamente la conversión.

Respecto al tiempo, el día de la semana importa: jueves, martes y fines de semana muestran mejores resultados, mientras que los viernes son menos efectivos. La fecha, mes o año no afectan de manera significativa.

Finalmente, se observa que el interés por el producto aumenta en contextos económicos menos favorables, lo que sugiere que adaptar la campaña según la situación del mercado puede mejorar los resultados.

Para resumir, si queremos maximizar el éxito, las próximas campañas deben centrarse en los segmentos más receptivos, priorizar la relación continua, optimizar el número y la duración de los contactos, y considerar la situación económica y los días de contacto. Se podría elaborar una guía para los empleados que vayan a dedicarse a las campañas por venir para que tengan en cuenta todos los factores y adapten mejor el discurso a la persona según su situación. Además, habrá que estudiar cómo hacer la oferta más personalizada e interesante para diferentes grupos de clientes.