

MEMORIA

easyMoney

**Estrategia de upsell y
crosssell al cliente ya adquirido**

Grupo PISA - Edición Marzo 2024

Integrantes:

Fernando Arroyo

Iván Sánchez

Carles Tisaire

Cristina Torrens

ÍNDICE

01	Introducción
	Antecedentes del proyecto
	Alcance y objetivos del proyecto
	Descripción general de la plataforma EasyMoney
	Desafíos y objetivos clave
02	Objetivos del proyecto
	Identificar y comprender las personas compradoras
	Desarrollar estrategias de venta adicional y venta cruzada
	Mejorar la retención y el compromiso del cliente
	Alinear las estrategias de ventas con segmentos clave de clientes
03	Metodología
	Recopilación y limpieza de datos
	Proceso de segmentación de clientes
	Herramientas y técnicas analíticas utilizadas
	Supuestos y limitaciones clave
04	Segmentación de clientes
	Descripción general de los segmentos clave
	Análisis demográfico y conductual
	Preferencias de productos y correlaciones
05	Estrategías de venta
	Enfoques de venta adicional y venta cruzada por segmento de clientes
	Mensajería personalizada y segmentación
	Ejemplos de ofertas y paquetes de productos
	Acciones de ventas basadas en datos

06**Implementación de tácticas de ventas para clientes adquiridos**

Acciones clave basadas en el análisis de datos

Canales de comunicación y técnicas

07**KPIs y medición de desempeño**

Descripción general de los indicadores clave de rendimiento (KPIs)

Métricas de retención de clientes

Tasas de conversión de ventas

ROI e impacto financiero

08**Desafíos y lecciones aprendidas**

Problemas enfrentados durante el procesamiento y análisis de datos

Desafíos internos (por ejemplo, colaboración con TI, Marketing y Finanzas)

Lecciones para futuras implementaciones

09**Conclusión**

Resumen de los logros del proyecto

Perspectivas futuras para las estrategias de ventas y marketing de EasyMoney

ANNEXO - Explicación Notebooks

01 - Introducción

Antecedentes del proyecto

El presente proyecto se desarrolla en el contexto de la compañía easyMoney, una plataforma multi-canal que ofrece productos financieros como ahorro, inversión y financiación. Fundada hace cuatro años por Carol Denver, easyMoney nació con la visión de simplificar la gestión financiera para sus clientes mediante productos accesibles y fáciles de utilizar. Sin embargo, con el crecimiento y la expansión de su oferta, la compañía ha enfrentado ciertos desafíos que ponen en riesgo su estabilidad financiera. Uno de los principales problemas ha sido la necesidad de rentabilizar su base actual de clientes, debido a la presión ejercida por su principal socio, Lion Global Management. En este escenario, la incorporación de un Data Scientist se convierte en una estrategia clave para mejorar la rentabilidad a través de la segmentación de clientes y la personalización de estrategias comerciales basadas en datos.

Alcance y objetivos del proyecto

El objetivo principal del proyecto es desarrollar un modelo de segmentación de clientes basado en técnicas avanzadas de machine learning. Este modelo busca identificar patrones de comportamiento entre los clientes de easyMoney, para agruparlos en diferentes segmentos y, en base a esto, diseñar campañas de marketing personalizadas. A través de este enfoque, se pretende maximizar la rentabilidad de la compañía mediante una mayor venta cruzada, una mejora en la retención de clientes y la optimización de los recursos invertidos en campañas comerciales.

Entre los objetivos específicos del proyecto destacan:

1. Segmentación de clientes: Identificar grupos de clientes con características comunes que permitan adaptar las estrategias de marketing y ventas a sus necesidades.
2. Optimización de productos: Diseñar modelos predictivos que determinen qué productos financieros son más adecuados para cada segmento de clientes.
3. Personalización de campañas de marketing: Implementar estrategias comerciales dirigidas a mejorar la relación con los clientes y aumentar su compromiso con los productos de easyMoney.
4. Medición del impacto financiero: Evaluar el retorno de inversión (ROI) de las campañas aplicadas, utilizando métricas como la tasa de conversión y la retención de clientes.

Descripción general de la plataforma EasyMoney

easyMoney se ha posicionado como una plataforma líder en el mercado financiero, con una oferta variada que incluye productos de ahorro, inversión y financiación. A través de alianzas estratégicas con entidades como easyBanking S.A., la compañía ha logrado captar una base de clientes significativa. No obstante, debido a la falta de inversión en tecnología y la necesidad de mejorar su EBITDA, easyMoney busca implementar una estrategia "Data-Driven" que le permita optimizar su propuesta de valor. El foco principal de esta nueva etapa es aprovechar su base actual de clientes mediante técnicas de analítica avanzada que faciliten el desarrollo de campañas comerciales más eficientes y dirigidas.

02 - Objetivos del proyecto

03 - Metodología

Recopilación y limpieza de datos

El proceso comenzó con la recopilación de tres conjuntos de datos proporcionados por easyMoney: **df_commercial_activity**, **df_products**, y **df_sociodemographic**. Estos datasets contenían información relevante sobre las actividades comerciales de los clientes, los productos financieros que poseen, y sus características sociodemográficas. Una vez obtenidos los datos, el siguiente paso fue realizar un exhaustivo proceso de limpieza y preprocesamiento para asegurar su calidad.

Durante la limpieza de datos, se identificaron múltiples duplicados, especialmente en clientes que aparecían hasta 17 veces. Para solventar esta situación, se decidió conservar la fila con la información más reciente de cada cliente. Posteriormente, se unificaron los tres datasets utilizando las claves **PK_CID** y **PK_PARTITION**, consolidando toda la información en un único conjunto de datos. Se eliminaron las columnas irrelevantes y se manejaron los valores nulos mediante técnicas de imputación. Por ejemplo, en el caso de la variable **salary**, se emplearon métodos de **backfill** y **frontfill** para completar los valores faltantes, agrupándolos según la edad para una mayor precisión.

Proceso de segmentación de clientes

Para la segmentación de clientes se aplicó un enfoque de clustering, utilizando tanto datos sociodemográficos como información sobre la posesión de productos financieros.

Se seleccionaron variables clave como salary, age, segment, y active_customer, además de nuevas variables derivadas, como number_of_products, account_balance, y loan_amount, para enriquecer el análisis y capturar patrones en el comportamiento de los clientes.

Se empleó el algoritmo de K-Means para segmentar a los clientes en varios grupos, primero en 7 clusters, y luego, para una campaña más específica, en 4 clusters, los cuales fueron utilizados para personalizar las campañas de marketing. Cada cluster fue interpretado y descrito, asignándoles nombres en función de las características dominantes de los clientes. Esto permitió identificar perfiles claros y diseñar estrategias de venta más efectivas.

Herramientas y técnicas analíticas utilizadas

El análisis y modelado se realizaron utilizando diversas herramientas y técnicas, todas basadas en Python. Las bibliotecas principales incluyeron pandas para la manipulación de datos, matplotlib y seaborn para la visualización, y scikit-learn para la implementación de algoritmos de machine learning como K-Means para clustering y modelos de clasificación como RandomForest y CatBoost.

Además, se utilizaron técnicas como One-Hot Encoding y Ordinal Encoding para transformar las variables categóricas, y se crearon nuevas variables a partir de datos de fecha para mejorar la capacidad predictiva de los modelos. El balanceo de clases fue una técnica clave, ya que los datos estaban altamente desbalanceados en dos categorías de productos (Inversión y Financiación), lo que requirió una cuidadosa consideración para asegurar que los modelos no se sesgaran.

Supuestos y limitaciones clave

Uno de los supuestos clave fue que los datos proporcionados eran representativos de la totalidad de los clientes de easyMoney. Sin embargo, es importante destacar que ciertos grupos de clientes podían estar sobrerrepresentados, lo que afectaría la generalización de los resultados. Asimismo, la gran cantidad de valores nulos en columnas importantes como salary supuso un desafío, ya que la imputación de estos datos, aunque necesaria, pudo introducir sesgos en el análisis.

En cuanto a las limitaciones, se identificó un desbalance significativo en la distribución de los clientes en relación con los productos financieros. Este desbalance afectó la capacidad de los modelos para predecir adecuadamente la adquisición de productos en las categorías de Inversión y Financiación, lo que requirió la implementación de técnicas de balanceo para mejorar el rendimiento de los modelos. Además, el hecho de que los datos financieros de los clientes tuvieran una distribución sesgada hacia la clase "no comprador" pudo limitar la capacidad predictiva de los modelos para identificar compradores potenciales.

04 - Segmentación de los clientes

Descripción general de los segmentos clave

La segmentación de los clientes se realizó agrupando a los clientes de easyMoney en siete clusters distintos. Cada cluster representa un grupo de clientes con características y comportamientos financieros comunes. La descripción detallada de cada uno de estos segmentos se encuentra en el documento titulado "Interpretación 7 Clusters".

Análisis demográfico y conductual

El análisis demográfico reveló una clara distinción en términos de edad, ingresos y actividad bancaria entre los diferentes clusters. Algunos clusters, como el grupo de Jóvenes Prometedores, están compuestos mayormente por clientes jóvenes con ingresos moderados y un bajo nivel de actividad en productos financieros. En contraste, otros clusters, como el de Clientes Complejos y Activos, agrupan a clientes con altos ingresos, múltiples productos bancarios y un alto compromiso con los servicios de easyMoney.

Los datos muestran que la mayoría de los clientes jóvenes tienden a tener una menor diversificación de productos y un nivel bajo de compromiso, lo que sugiere una menor relación a largo plazo con la plataforma. Por otro lado, los clientes de mayor edad y con mayores ingresos tienden a utilizar una mayor cantidad de productos financieros, manteniendo balances más elevados en sus cuentas, lo que representa un perfil más rentable para la empresa.

Preferencias de productos y correlaciones

En términos de preferencias de productos, los análisis identificaron correlaciones significativas entre el nivel de ingresos de los clientes y los productos que suelen adquirir. Por ejemplo, los clientes del segmento de Alta Capacidad, Baja Relación muestran una clara tendencia hacia productos de inversión y ahorro, mientras que aquellos con ingresos más bajos tienden a centrarse en productos básicos, como cuentas de ahorro o tarjetas de débito.

Además, los clusters con mayor compromiso con los productos de easyMoney tienden a generar una mayor rentabilidad, debido al uso intensivo de productos bancarios diversificados. Por ejemplo, los Inversores Élite tienen una fuerte correlación entre su

nivel de ingresos y la cantidad de productos que manejan, lo que los convierte en un segmento clave para las estrategias de venta cruzada.

05 - Estrategias de ventas

Enfoques de venta adicional y venta cruzada por segmento de clientes

Las estrategias de venta adicional (up-selling) y venta cruzada (cross-selling) se basaron en el análisis de cada uno de los segmentos definidos a partir del segundo clustering. En función de las características sociodemográficas y el comportamiento financiero de los clientes, se diseñaron enfoques específicos para maximizar la rentabilidad de la plataforma easyMoney:

1. Clientes Estables y Diversificados: Para este grupo, el enfoque principal fue la venta cruzada de productos como tarjetas de crédito y cuentas de nómina, que les permitirían una mayor centralización de sus operaciones financieras. Además, se promovieron fondos de inversión de bajo riesgo que se alinearan con su perfil financiero.
2. Clientes de Alto Valor y Oportunidad de Expansión: Se implementó una estrategia de venta adicional, centrada en la promoción de planes de pensiones y productos hipotecarios. Este grupo tiene una alta capacidad adquisitiva y está interesado en diversificar su portafolio con productos de inversión a largo plazo y bienes inmuebles.
3. Clientes de Patrimonio Elevado y Conservador: En este segmento, la estrategia se centró en la oferta de depósitos a largo plazo y cuentas premium. Dada su aversión al riesgo, estos clientes valoran productos que ofrecen seguridad financiera. Para aumentar su engagement, también se ofreció contenido exclusivo de figuras destacadas en el mundo de las finanzas, con el fin de fidelizar a este segmento.
4. Profesionales en Crecimiento: Para este grupo, se apostó por productos fáciles de gestionar, como criptomonedas y cuentas de ahorro a corto plazo. La venta cruzada también incluyó productos que fomentan la inversión inicial, como depósitos a corto plazo. La estrategia de ventas buscó captar su interés mediante productos llamativos, aprovechando tendencias financieras recientes como las criptomonedas.

La explicación más ampliada y detallada de los segmentos se encuentra en el documento llamado "Interpretación Segundo Clustering"

Mensajería personalizada y segmentación

La campaña de marketing incluyó la creación de mensajes personalizados para cada segmento, basados en los resultados del clustering y las preferencias detectadas. Cada email se diseñó con un enfoque específico, resaltando los beneficios que easyMoney podía ofrecer a cada grupo:

- ***Clientes Estables y Diversificados***: El mensaje principal para este segmento se centró en la seguridad y la flexibilidad financiera. Además, se destacaron productos como cuentas de nómina y tarjetas de crédito con beneficios especiales.
- ***Clientes de Alto Valor y Oportunidad de Expansión***: La mensajería para este segmento se enfocó en la inversión a largo plazo, resaltando la posibilidad de acceder a productos premium y personalizados, como planes de pensiones y servicios hipotecarios exclusivos.
- ***Clientes de Patrimonio Elevado y Conservador***: Este segmento no solo recibió ofertas de productos financieros, sino también contenido exclusivo de personalidades financieras de renombre, con el fin de fidelizarlos a través de una relación de confianza y autoridad.
- ***Profesionales en Crecimiento***: El mensaje para este grupo fue más dinámico, apelando a sus intereses con una campaña innovadora que incluyó a figuras de la cultura popular, como C. Tangana. Este enfoque creó polémica, pero logró captar la atención de este segmento joven, ansioso por seguir tendencias. Se ofrecieron productos como criptomonedas y herramientas de gestión financiera básicas.

Ejemplos de ofertas y paquetes de productos

Cada email incluyó una oferta específica que se ajustaba a las características del segmento correspondiente:

1. Clientes Estables y Diversificados: Se ofrecieron paquetes que incluían una tarjeta de crédito con tasas preferenciales, junto con una cuenta de nómina con beneficios exclusivos como la devolución de un porcentaje de los gastos.
2. Clientes de Alto Valor y Oportunidad de Expansión: Los paquetes ofrecidos incluían la posibilidad de combinar productos de inversión a largo plazo, como planes de pensiones, con préstamos hipotecarios a tasas reducidas.

3. Cientes de Patrimonio Elevado y Conservador: En este segmento, se ofrecieron productos como depósitos a largo plazo con tasas competitivas, junto con cuentas premium que incluían asesoría financiera personalizada.
4. Profesionales en Crecimiento: Se promocionaron productos financieros accesibles, como cuentas de ahorro y criptomonedas, con campañas llamativas que incluyeron la imagen de influencers y figuras relevantes en el mundo juvenil.

Acciones de ventas basadas en datos

Todas las decisiones de venta estuvieron respaldadas por datos obtenidos del clustering y del análisis de probabilidades de compra. Los segmentos se diseñaron en función de las características específicas de los clientes, y las ofertas fueron personalizadas según sus probabilidades de adquirir productos financieros específicos.

Por ejemplo, los Clientes de Alto Valor mostraron una alta propensión a adquirir productos de inversión a largo plazo, por lo que las estrategias de venta cruzada se centraron en este tipo de productos. Del mismo modo, para los Jóvenes Profesionales, la campaña se diseñó para captar su atención mediante productos innovadores, aprovechando tendencias del mercado como las criptomonedas.

06 - Implementación de tácticas de ventas para clientes adquiridos

Acciones Clave Basadas en el Análisis de Datos

La implementación de tácticas de ventas para los clientes adquiridos de EasyMoney se basa en la segmentación realizada y el análisis detallado de los comportamientos y preferencias de cada segmento de cliente identificado. Para maximizar la rentabilidad y el engagement, se adoptarán las siguientes acciones:

- **Segmentación Personalizada:** La segmentación permitió identificar cuatro grupos de clientes clave con características demográficas y comportamientos financieros únicos. Cada grupo tiene una estrategia de ventas y comunicación personalizada para asegurar la máxima efectividad de las campañas de upselling y cross-selling.
- **Upselling y Cross-Selling:** Basándonos en el comportamiento y preferencias de cada segmento, se ofrecerán productos adicionales que se adapten a sus necesidades y perfil financiero. Por ejemplo, para los "Clientes de Alto Valor y Oportunidad de Expansión", se promoverán productos de inversión a largo plazo como planes de pensiones y productos hipotecarios.

Canales de Comunicación y Técnicas

Para asegurar que las tácticas de ventas sean efectivas, se utilizarán los siguientes canales de comunicación y técnicas:

- **Emails Personalizados:** En base a la segmentación, se enviarán emails dirigidos que destaquen los beneficios de los productos ofrecidos. Los mensajes estarán adaptados a cada grupo, enfatizando las ventajas de los productos financieros que se ajustan a sus perfiles. Por ejemplo, el uso de contenido exclusivo para los "Clientes de Patrimonio Elevado y Conservador" ayudará a fidelizar este segmento.
- **Mensajería Específica:** Cada mensaje de correo electrónico será personalizado de acuerdo con el grupo al que se dirige. Por ejemplo:
 - **Para Clientes Estables y Diversificados:** Destacar la seguridad y flexibilidad de los productos.
 - **Para Profesionales en Crecimiento:** Utilizar un tono más dinámico y tendencias como las criptomonedas para captar su atención.
- **Campañas de Retargeting:** Se implementarán campañas de retargeting para impactar a aquellos clientes que han mostrado interés previo en productos específicos, aumentando la probabilidad de conversión.
- **Contenido Exclusivo y Beneficios Adicionales:** Se ofrecerán incentivos exclusivos como tasas de interés preferenciales, acceso a contenido financiero de expertos y bonificaciones, dependiendo del segmento al que pertenezcan los clientes.

07 - KPIs y medición del desempeño

Descripción general de los indicadores clave de rendimiento (KPIs)

Para evaluar el éxito de las campañas de marketing dirigidas a los cuatro segmentos de clientes identificados, se establecieron una serie de KPIs que permitieron medir tanto la eficacia de las campañas como su impacto financiero. Estos indicadores fueron los siguientes:

1. **Tasa de apertura:** Mide el porcentaje de destinatarios que abrieron el correo electrónico. Este KPI fue crucial para entender qué tan atractivo resultaba el asunto del correo para cada segmento.
2. **Tasa de clics (CTR):** Mide el porcentaje de personas que hicieron clic en los enlaces dentro del correo, como los botones de llamada a la acción (CTA). Este indicador ayudó a evaluar qué tan relevante era el contenido del email para los clientes.

3. **Tasa de conversión:** Indica el porcentaje de clientes que completaron la acción deseada (por ejemplo, comprar un producto o registrarse en un servicio) después de haber interactuado con el correo.
4. **Tasa de rebote:** Mide el porcentaje de correos que no pudieron ser entregados, ayudando a identificar problemas con la lista de contactos o el proceso de envío.
5. **Tasa de cancelación de suscripción:** Mide el porcentaje de clientes que optaron por darse de baja de la lista de correo, lo cual proporciona información sobre la relevancia del contenido y su adecuación a los intereses de los destinatarios.
6. **Retorno de la inversión (ROI):** Mide los ingresos generados por la campaña en comparación con los costos de ejecución, incluidos incentivos y comisiones.

Métricas de retención de clientes

La retención de clientes fue medida a través de la tasa de reenvío y el compromiso con el contenido. En el caso de los clientes de mayor valor (por ejemplo, los Clientes de Patrimonio Elevado y Conservador), una de las estrategias clave fue el envío de contenido exclusivo de personas influyentes en el mundo financiero. La tasa de reenvío de este segmento se utilizó como métrica para medir su interés y fidelización hacia el contenido recibido.

Adicionalmente, la tasa de cancelación de suscripción fue un indicador importante para medir la retención. Las campañas con una baja tasa de cancelación indicaron que los clientes encontraron útil y relevante el contenido que recibieron, lo que impulsó su permanencia en la lista de correo y su compromiso con easyMoney.

Tasas de conversión de ventas

Las tasas de conversión fueron clave para medir el éxito de la campaña en términos de generación de ingresos. Cada segmento de clientes mostró diferentes tasas de conversión en función de los productos ofrecidos:

- **Clientes Estables y Diversificados:** Este grupo mostró una tasa de conversión moderada en productos de inversión básica, como fondos de bajo riesgo y cuentas de nómina.
- **Clientes de Alto Valor y Oportunidad de Expansión:** Se observó una alta tasa de conversión en productos premium, como planes de pensiones y productos hipotecarios.
- **Clientes de Patrimonio Elevado y Conservador:** Este segmento mostró una alta tasa de conversión en productos de depósito a largo plazo y cuentas premium.

- **Profesionales en Crecimiento:** Este segmento, aunque mostró una menor conversión general en productos financieros tradicionales, respondió positivamente a la oferta de criptomonedas y productos financieros innovadores.

En total, las tasas de conversión variaron entre un 2% y un 20%, dependiendo del segmento y el producto ofrecido.

ROI e impacto financiero

El retorno de la inversión (ROI) fue calculado considerando los ingresos generados por cada segmento frente a los costos asociados con la campaña de marketing. En este análisis, los productos ofrecidos y las tasas de conversión fueron determinantes para evaluar el impacto financiero.

Por ejemplo, los Clientes de Alto Valor y Oportunidad de Expansión generaron altos ingresos debido a su propensión a adquirir productos premium, mientras que los Profesionales en Crecimiento, aunque generaron menores ingresos por producto individual, contribuyeron de manera significativa al ROI global gracias a su interés en productos como criptomonedas, cuya oferta requería menores costos operativos.

El análisis financiero final mostró que los segmentos de clientes con mayor compromiso y con productos de mayor valor agregado generaron el mayor ROI, contribuyendo de manera significativa a la mejora del EBITDA de easyMoney. En total, los ingresos estimados por cliente se calcularon en base a la probabilidad de compra y los márgenes de ingresos de cada producto, logrando un ROI positivo para cada una de las campañas implementadas.

08 - Desafíos y lecciones aprendidas

Problemas enfrentados durante el procesamiento y análisis de datos

Uno de los primeros problemas enfrentados durante la fase de preprocesamiento fue la gran cantidad de valores nulos en la columna salary. Dada la importancia de esta variable para el análisis sociodemográfico y financiero, su imputación representó un desafío significativo. Se consideraron varias estrategias, como eliminar las filas con valores faltantes o imputar los datos a partir de otras características, pero ninguna opción resultaba completamente satisfactoria. La eliminación masiva de filas habría reducido de manera drástica el tamaño del dataset, mientras que la imputación podría haber introducido sesgos en el análisis. Finalmente, se optó por conservar los datos más recientes, aunque la proporción de valores nulos no disminuyó considerablemente.

Otro reto importante fue la duplicidad de clientes, algunos de los cuales aparecían hasta 17 veces en los datos. Aunque se recomendó quedarse con los registros más recientes, surgió la preocupación de perder información relevante, ya que algunos clientes mostraban una mayor variedad de productos en fechas anteriores. La eliminación de duplicados permitió reducir el tamaño del dataset, que inicialmente alcanzaba casi 11 GB, a un volumen manejable. Sin embargo, este proceso también presentó el riesgo de perder datos valiosos sobre la cartera de productos de los clientes, lo cual pudo haber impactado en la calidad de los modelos predictivos.

Por último, se detectó que algunos clientes ya poseían productos financieros adquiridos de otros bancos antes de que se comenzara a recolectar datos sobre ellos. Aunque esto podría haber afectado la validez de ciertos análisis, se decidió incluirlos en el dataset, pues aún representaban oportunidades de marketing para easyMoney.

Desafíos internos

El proceso de colaboración entre los distintos equipos de easyMoney presentó sus propios desafíos. La interacción con los departamentos de TI, Marketing y Finanzas fue clave para garantizar el éxito del proyecto, pero surgieron dificultades al alinear las expectativas y los objetivos de cada área.

Colaboración con TI: El tamaño inicial del dataset (aproximadamente 11 GB) representó un reto logístico para los recursos computacionales disponibles. El equipo de TI proporcionó el soporte necesario para manejar grandes volúmenes de datos, pero se enfrentaron a limitaciones en términos de procesamiento, lo que retrasó algunas fases del análisis. Además, la unificación de las diferentes bases de datos mediante un proceso de merge resultó en tiempos de ejecución más largos de lo esperado.

Colaboración con Marketing: El equipo de Marketing tuvo un rol clave en la definición de los segmentos de clientes y las campañas asociadas. Sin embargo, la implementación de las estrategias de venta cruzada y venta adicional dependió en gran medida de la interpretación correcta de los clusters, lo que generó fricciones en cuanto a la segmentación adecuada. El desafío principal fue alinear los resultados del análisis de datos con las necesidades comerciales de Marketing.

Colaboración con Finanzas: El equipo de Finanzas se mostró muy interesado en la precisión de los modelos predictivos, en especial en lo relacionado con las proyecciones de ingresos por cliente. La dificultad aquí radicó en encontrar un equilibrio entre modelos que generaban tasas de acierto muy altas, lo que sugería sobreajuste, y modelos más equilibrados que reflejaban resultados más realistas pero con tasas de acierto más bajas.

Lecciones para futuras implementaciones

A lo largo del proyecto, se extrajeron varias lecciones valiosas que pueden servir de referencia para futuras implementaciones:

1. **Mejora en la gestión de datos faltantes:** La presencia de valores nulos, en especial en columnas críticas como salary, requiere una planificación más rigurosa desde el inicio del proyecto. La posibilidad de recurrir a imputaciones más sofisticadas, como algoritmos de machine learning para estimar valores, debe considerarse para futuros análisis.
2. **Reducción de duplicados sin pérdida de información:** El manejo de clientes duplicados fue un desafío que afectó tanto el tamaño del dataset como la calidad del análisis. En futuros proyectos, sería recomendable desarrollar una metodología más refinada para identificar qué versión de los datos de un cliente es más representativa, sin perder información valiosa sobre su comportamiento en el tiempo.
3. **Equilibrio en la creación de nuevas features:** La creación de nuevas variables puede mejorar los modelos, pero también puede llevar a problemas de sobreajuste. El proyecto mostró que es importante evaluar constantemente el impacto de las features nuevas en los modelos, ajustándolas o eliminándolas si generan resultados demasiado optimistas.
4. **Colaboración multidisciplinaria desde el inicio:** La interacción entre los equipos de TI, Marketing y Finanzas fue fundamental, pero también presentó algunos retos. Para futuras implementaciones, sería ideal integrar a todos los equipos desde las primeras fases del proyecto, estableciendo objetivos claros y expectativas realistas en cuanto a la interpretación y uso de los datos.
5. **Interpretación de resultados y acciones basadas en datos:** Uno de los desafíos más complejos fue interpretar los resultados obtenidos de los modelos y tomar decisiones basadas en ellos. La lección aquí es la importancia de trabajar estrechamente con expertos de las áreas comerciales para alinear la interpretación de los datos con las estrategias empresariales.

09 - Conclusión

Resumen de los logros del proyecto

El proyecto realizado para easyMoney ha permitido alcanzar varios objetivos clave, con resultados significativos en la segmentación de clientes y en la implementación de estrategias de marketing personalizadas. A través del uso de técnicas avanzadas de machine learning, como el clustering y los modelos de predicción, se logró identificar grupos de clientes con comportamientos y preferencias similares. Estos segmentos

fueron utilizados para diseñar campañas de marketing dirigidas, optimizando así las oportunidades de venta cruzada y venta adicional.

Uno de los logros más destacados fue la creación de cuatro campañas de marketing personalizadas, adaptadas a los segmentos de clientes más prometedores, lo que permitió maximizar la probabilidad de conversión. Además, se implementaron soluciones efectivas para manejar los desafíos presentados en el procesamiento y análisis de datos, lo que resultó en modelos predictivos más precisos y realistas.

La combinación de técnicas de segmentación, personalización de mensajes y análisis de rendimiento a través de KPIs permitió a easyMoney obtener una visión clara del impacto de las campañas de marketing en sus ingresos. El proyecto también proporcionó una base sólida para futuras iniciativas de marketing basadas en datos, mejorando la toma de decisiones en tiempo real y el retorno de la inversión (ROI).

Perspectivas futuras para las estrategias de ventas y marketing de EasyMoney

Mirando hacia el futuro, las estrategias de ventas y marketing de easyMoney pueden continuar evolucionando sobre la base del trabajo realizado en este proyecto. En particular, se destacan tres áreas de enfoque clave:

1. **Seguimiento continuo y actualización de la base de datos (BBDD):** Para lograr resultados estadísticos más robustos y fiables, es fundamental seguir complementando y actualizando la base de datos. Esto permitirá refinar los perfiles de los clientes y ajustar las campañas de marketing en función de los cambios en sus comportamientos y necesidades. Además, se deberán identificar nuevos Buyer Personas y desarrollar datos adicionales que ofrezcan insights más profundos sobre los usuarios.
2. **Validación de las acciones:** Las estrategias de ventas implementadas deben ser sometidas a un seguimiento continuo para validar su efectividad. Las campañas personalizadas deben revisarse regularmente para asegurarse de que siguen siendo relevantes para los diferentes segmentos, adaptándose a las nuevas dinámicas del mercado. Las futuras campañas pueden beneficiarse de este enfoque dinámico, maximizando tanto la retención de clientes como las tasas de conversión.
3. **Validación de los datos cuantitativos y cualitativos:** Es crucial llevar a cabo una validación tanto cualitativa como cuantitativa de los perfiles de los clientes y las campañas de marketing. A medida que se realicen reuniones con los clientes, se podrán obtener datos adicionales que refinen las estrategias, permitiendo una diferenciación aún mayor en cuanto a las preferencias de los canales de comunicación y el contenido ofrecido. Esta validación permitirá a easyMoney implementar mensajes más efectivos y acciones de activación personalizadas.

ANNEXO - Explicación Notebooks

1. PREPROCESSING

En este primer notebook se realiza un análisis exhaustivo de preprocesamiento y limpieza de tres datasets (`df_commercial_activity`, `df_products` y `df_sociodemographic`) con el objetivo de preparar los datos para su análisis y modelado posterior.

1.1. Análisis exploratorio de los datasets:

Se inicia el proceso explorando la distribución de las variables en los tres datasets, identificando patrones y anomalías. Durante este análisis, se detecta que algunos clientes están repetidos hasta 17 veces. Para corregir esto, se decide quedarse con la fila que contiene la información más reciente de cada cliente en los tres datasets, utilizando la fecha como criterio de selección.

1.2. Unificación de los datasets:

Se realiza un `*merge*` entre los datasets `df_commercial_activity` y `df_products` utilizando la columna `PK_CID` (clave de cliente). Posteriormente, se unen estos datos con `df_sociodemographic` usando la columna `PK_PARTITION`, asegurándose de que las fechas coincidan entre los datasets. Este paso es crucial para consolidar toda la información en un solo conjunto de datos coherente.

1.3. Análisis de la distribución por producto:

Al analizar la distribución de productos, se observa que más del 90% de las variables tienen el valor 0, lo que sugiere un desbalance importante en los datos. Este descubrimiento es clave para ajustar los enfoques en el modelado futuro.

1.4. Limpieza de duplicados y columnas innecesarias:

Se eliminan los duplicados y las columnas irrelevantes como `'Unnamed'`, presentes en los tres datasets. Esto ayuda a reducir la cantidad de ruido y facilitar un análisis más eficiente.

1.5. Manejo de valores nulos:

En cuanto a los valores nulos, se eliminan las filas con nulos en las columnas ``entry_channel``, ``entry_date``, y ``segment``, ya que representan una proporción insignificante de los datos.

En el dataset ``df_sociodemographic``, se agrupan las categorías de la columna ``country_id`` que no son "ES" (España) bajo la categoría "Otros". Durante este proceso, se observa que las filas clasificadas como "Otros" no tienen valores en la columna ``region_code``. Esto sugiere que las observaciones bajo "Otros" no necesitan un código de región, por lo que se decide eliminar dichas filas.

En la columna ``salary``, que contiene aproximadamente un millón y medio de valores nulos, se aplica una técnica de imputación con métodos de `*backfill*` y `*frontfill*`. Además, el salario se agrupa en función de la variable ``age``, con el fin de obtener una imputación más precisa.

1.6. Agrupación de categorías poco representadas:

En la columna ``entry_channel``, las categorías que representan menos del 7% del total se agrupan bajo una nueva categoría, lo que simplifica el análisis y evita problemas de alta cardinalidad en el modelado.

1.7. Ordenación y consolidación de datos:

Al finalizar el proceso, los datos se ordenan cronológicamente de la fecha más antigua a la más reciente, lo que facilita el análisis temporal. Finalmente, se realiza un `*merge*` final de los tres datasets, consolidando toda la información en un único conjunto de datos.

1.8. Reducción del tamaño del dataset:

Aunque los datasets originales contenían casi 6 millones de filas cada uno, tras eliminar duplicados y realizar el `*merge*`, el número de filas se reduce considerablemente a alrededor de 450.000. Esta reducción permite un manejo más eficiente de los datos sin perder calidad en la información.

2. ENCODING

En este segundo notebook se lleva a cabo el proceso de *encoding* y transformación de las variables categóricas y de fecha, con el fin de que los datos puedan ser utilizados en modelos de machine learning.

2.1. Conversión de columnas de fecha:

- Tras importar los datos preprocesados del notebook anterior, se detecta que las columnas de fechas como `pk_partition` y `entry_date`, que

previamente estaban en formato datetime, se habían convertido en objetos (object). Por lo tanto, se procede a reconvertirlas a su formato correcto de datetime.

- Dado que los modelos de machine learning no pueden trabajar directamente con datos en formato de fecha, se crean nuevas columnas a partir de estas fechas: año, mes, día y día de la semana. Esto permite capturar la información temporal en forma de variables numéricas. El día de la semana se codifica como números del 0 (lunes) al 6 (domingo).

2.2. Encoding de variables categóricas:

- Se procede a transformar las columnas categóricas que aún no estaban en formato numérico. Para ello, se utilizan dos métodos principales de encoding:
 - **One-Hot Encoding:** Se aplica a las columnas entry_channel, country_id, gender y deceased, creando nuevas columnas binarias para cada categoría en estas variables.
 - **Ordinal Encoding:** Se aplica a la columna segment, ya que sus categorías tienen una numeración implícita del 1 al 3, que indica un nivel de importancia o jerarquía entre ellas. Por lo tanto, el ordinal encoding preserva esta relación numérica.

2.3. Eliminación de columnas redundantes:

- Tras aplicar el One-Hot Encoding, se generan columnas binarias, y se identifican columnas redundantes o duales como gender_H y deceased_N, que son eliminadas para evitar duplicidades innecesarias en los datos y optimizar el rendimiento del modelo.

3. SEGMENTACIÓN

En este tercer notebook se realiza el proceso de segmentación de clientes del banco EasyMoney utilizando técnicas de *clustering*, con el objetivo de agrupar a los clientes en 7 clusters diferentes en función de sus características.

3.1. Selección de características clave para la segmentación:

- El análisis inicial determinó que las características más importantes para segmentar a los clientes eran el salario (salary) y la edad (age), pero también se identificaron otras variables relevantes como segment, region_code y active_customer. Estas variables aportan información valiosa sobre el perfil del cliente y su comportamiento.

3.2. Creación de nuevas features:

- Con el fin de incluir más información sobre los productos financieros de los clientes, se crean nuevas features que representan de manera resumida y cuantitativa su interacción con los productos del banco, sin tener que utilizar todas las columnas de productos de forma directa. Estas nuevas features incluyen:
 - **number_of_products**: Número total de productos financieros que el cliente posee con el banco.
 - **account_balance**: Saldo promedio de la cuenta (inferido a partir de ciertos productos financieros).
 - **loan_amount**: Monto total de préstamos o deuda que el cliente tiene con el banco.
 - **product_engagement_score**: Una métrica compuesta que mide el nivel de compromiso del cliente con los productos financieros, asignando diferentes pesos a distintos tipos de productos.
 - **debt_to_income_ratio**: Relación entre la deuda total del cliente (préstamos + hipotecas) y su salario. Esta métrica es fundamental para evaluar el riesgo financiero de un cliente.

3.3. Aplicación del algoritmo de *clustering*:

- Una vez que las nuevas features están creadas, se procede a aplicar un algoritmo de *clustering* para agrupar a los clientes en 7 clusters, tal como se requiere en el proyecto. Las características seleccionadas y las nuevas features permiten identificar patrones en el comportamiento de los clientes que facilitan la segmentación.

3.4. Interpretación de los clusters:

- Después de realizar el *clustering*, se lleva a cabo un análisis de los clusters para asignarles nombres y definir los perfiles de los clientes en cada uno. Esta etapa es clave para entender cómo se diferencian los grupos de clientes y para identificar sus necesidades específicas.

3.5. Cálculo de estadísticas clave por cluster:

- Para proporcionar una visión general de cada cluster, se calculan estadísticas como la media del salario, la edad y el número de productos por cluster. Estas estadísticas ayudan a obtener una mejor comprensión de las diferencias entre los clusters, permitiendo una interpretación más profunda de los resultados.

DESPUÉS DE ESTO SE AÑADE EL DOCUMENTO DE LA SEGMENTACIÓN

4. AGRUPACIÓN

En este notebook se realiza una agrupación de los clientes del banco en función de los productos financieros que poseen. El objetivo es preparar los datos para un modelo de clasificación que permita predecir la probabilidad de que un cliente adquiriera uno o más productos dentro de diferentes categorías.

4.1. Creación de categorías de productos financieros:

- **Categoría de Ahorro e Inversión:** Esta categoría agrupa productos como depósitos a corto y largo plazo, fondos de inversión, valores (securities) y planes de pensiones.
- **Categoría de Financiación:** Incluye productos de financiamiento como préstamos personales (loans), hipotecas (mortgage) y tarjetas de crédito.
- **Categoría de Cuentas:** Agrupa productos relacionados con cuentas bancarias y otros productos transaccionales como cuentas de ahorro, nómina (payroll) y tarjetas de débito.

4.2. Conversión a formato binario:

- Para cada categoría, se creó una columna binaria que indica si un cliente tiene al menos un producto en esa categoría. Se asigna un valor de 1 si el cliente posee al menos un producto en la categoría, y un valor de 0 si no tiene ninguno. Este enfoque simplifica el conjunto de datos al reducir múltiples columnas de productos individuales en tres columnas resumen.

4.3. Eliminación de columnas de productos individuales:

- Tras la creación de las nuevas columnas de agrupación, se eliminaron las columnas originales de productos financieros. Esto ayuda a reducir la dimensionalidad del dataset y a centrarse en las variables agregadas que son más fáciles de manejar en el posterior análisis de clasificación.

4.4. Revisión de la distribución de las categorías:

- Se revisaron las distribuciones de los valores 0 y 1 dentro de cada categoría. Como se muestra en la imagen que compartiste, se observó que las distribuciones son bastante desiguales:
 - En la **Categoría de Ahorro e Inversión**, un 94.35% de los clientes no tiene productos de esta categoría, mientras que solo un 5.64% sí los tiene.
 - En la **Categoría de Financiación**, la disparidad es aún mayor, con un 98.92% de los clientes sin productos de financiación y solo un 1.07% con productos de este tipo.
 - La **Categoría de Cuentas** tiene una distribución más equilibrada, con un 73.58% de los clientes teniendo al menos un producto de esta categoría y un 26.41% sin ninguno.

4.5. Consideraciones para el modelo de clasificación:

- Esta revisión inicial de la distribución de los productos financieros revela un fuerte desequilibrio en dos de las tres categorías (Inversión y Financiación). Esto podría tener implicaciones importantes en el rendimiento del modelo de clasificación, ya que los datos desbalanceados suelen requerir técnicas específicas para asegurar que el modelo no se sesgue hacia la clase mayoritaria (en este caso, el valor 0).

5. BALANCEO

En este notebook, se aborda el problema del desbalanceo de clases en los datos de los tres grupos de productos financieros creados previamente: **Categoría de Ahorro e Inversión**, **Categoría de Financiación**, y **Categoría de Cuentas**. El objetivo es evitar problemas de sobreajuste (*overfitting*) en los modelos de clasificación, los cuales podrían surgir debido a que la gran mayoría de los clientes no poseen productos bancarios.

5.1. Identificación del desbalanceo:

- Se detecta que la mayoría de los usuarios no poseen productos bancarios, lo que se refleja en las clases con valor "0". Este desequilibrio es muy marcado, especialmente en las categorías de Inversión y Financiación, donde la mayoría de los clientes no tienen productos en esas áreas. Este desbalanceo crea un problema significativo en los datos, haciendo que los modelos de clasificación sean propensos a predecir siempre la clase mayoritaria ("0").

5.2. Problema de sobreajuste (*overfitting*):

- Cuando se entrena un modelo de clasificación en datos desbalanceados, el modelo puede mostrar tasas de acierto superiores al 98%. Sin embargo, este alto porcentaje es engañoso, ya que el modelo se sesga hacia la clase mayoritaria ("0"), ignorando la clase minoritaria ("1"). Esto da como resultado un modelo que no es capaz de predecir adecuadamente a los clientes que tienen productos bancarios, lo que es crucial para el análisis.

5.3. Balanceo de los datos:

- Para corregir este desbalanceo, se aplican técnicas específicas para asegurar que el número de clientes con productos (valor "1") sea igual al número de clientes sin productos (valor "0"). El procedimiento seguido es el siguiente:
 - Se separan las observaciones de la clase minoritaria (usuarios con productos, valor "1").

- Se toma una muestra aleatoria de los usuarios que no tienen productos (valor "0") para igualar el número de observaciones en ambas clases.
- Ambas muestras se combinan, asegurando que el conjunto de datos resultante tenga un número igual de "0" y "1". Este proceso equilibra el dataset y permite que los modelos aprendan de manera adecuada tanto de los usuarios con productos como de aquellos sin ellos.

5.4. Verificación del balanceo:

- Tras el proceso de balanceo, se verifica la distribución de las clases "0" y "1" para asegurar que ahora ambas clases tienen una proporción del 50%. Esto garantiza que los modelos de clasificación no estarán sesgados y podrán hacer predicciones justas y equilibradas sobre ambas clases.

5.5. Preparación para el modelado:

- Con los datos ya balanceados, el conjunto de datos está listo para ser utilizado en modelos de clasificación. El balanceo de clases asegura que los modelos podrán predecir con precisión tanto los usuarios que no tienen productos bancarios como aquellos que sí los tienen, evitando los problemas de sobreajuste observados anteriormente.

6. MODELO PARA GRUPO DE INVESTMENTS

Este Notebook sigue el mismo enfoque que el resto de los Notebooks con Modelos, pero enfocado en los productos del grupo "Investment". Los pasos realizados son los siguientes:

6.1. Preparación de los datos:

- Se eliminan las columnas relacionadas con los otros grupos de productos.
- Los datos se dividen en variables predictoras y la variable objetivo (category_investment).
- Se realiza una división en conjuntos de entrenamiento y prueba, para poder entrenar el modelo y evaluar su rendimiento.

6.2. Torneo de modelos:

- Se prueban varios modelos de clasificación, como LogisticRegression, RandomForest, y CatBoostClassifier, entre otros, para seleccionar el modelo con mejor rendimiento.
- Se selecciona el mejor modelo basado en el porcentaje de acierto y la métrica roc_auc.

6.3. Predicción:

- El modelo seleccionado se utiliza para predecir si los clientes comprarán productos del grupo "Investment".
- Se filtran los clientes que tienen una probabilidad de compra superior al 70%. En este caso, se identifican 4,991 clientes.

6.4. Exportación de resultados:

- Los clientes con alta probabilidad de compra (superior al 70%) se exportan a un archivo CSV. Estos resultados serán utilizados más adelante para conocer los ingresos potenciales y realizar una segmentación de clientes para una campaña de marketing.

7. MODELO PARA GRUPO DE FINANCING

Este Notebook tiene como objetivo crear un modelo de clasificación para predecir la probabilidad de que los clientes compren productos del grupo "Financing". Aquí está el resumen de lo que se realiza:

7.1. Preparación de datos:

- Se eliminan las columnas relacionadas con los otros dos grupos de productos ("investment" y "accounts") para reducir la dimensionalidad y centrarse únicamente en "Financing".
- Los datos se dividen en conjuntos de entrenamiento y prueba para poder entrenar el modelo y evaluar su rendimiento.

7.2. Torneo de modelos:

- Se realizan pruebas con varios algoritmos de clasificación, evaluando cuál ofrece el mejor rendimiento en cuanto a precisión.
- El mejor modelo es seleccionado para realizar las predicciones, basado en el mayor porcentaje de acierto.

7.3. Predicción:

- Se utiliza el modelo seleccionado para predecir si los clientes comprarán productos de "Financing" y se obtiene también la probabilidad de compra.
- Se seleccionan únicamente aquellos clientes con una probabilidad de compra superior al 70%, ya que son considerados los más propensos a adquirir estos productos.

7.4. Exportación de resultados:

- Los clientes con probabilidad de compra mayor al 70% (933 en total) se exportan a un archivo CSV, ya que serán utilizados más adelante en el proyecto para conocer los ingresos potenciales de la empresa y para realizar una segmentación de clientes

8. MODELO PARA GRUPO DE ACCOUNTS

En este Notebook, se sigue el mismo proceso que en los Notebooks de "Modelo Financing" y "Modelo Investment", pero centrado en el grupo de productos "Accounts". A continuación, se detalla lo que se realiza:

8.1. Preparación de los datos:

- Se eliminan las columnas relacionadas con los otros grupos de productos.
- Los datos se dividen en variables predictoras y la variable objetivo (category_accounts).
- Se realiza la división en conjuntos de entrenamiento y prueba para entrenar el modelo y evaluar su rendimiento.

8.2. Torneo de modelos:

- Se prueban múltiples algoritmos de clasificación y se selecciona el modelo con mejor rendimiento según las métricas de precisión y roc_auc.

8.3. Predicción:

- Se predice si los clientes comprarán productos del grupo "Accounts".
- Inicialmente, se identifica que casi 40,000 clientes tienen una probabilidad de compra superior al 70%.

- Debido a que solo se requieren 4076 clientes para completar el total de 10,000 en el proyecto, se decide seleccionar los clientes con las mejores probabilidades de compra por encima del 96%.

8.4. Exportación de resultados:

- Se exportan los 4076 clientes seleccionados con las mayores probabilidades de compra en un archivo CSV, ya que serán utilizados en la segmentación y la campaña de marketing.

9. UNIÓN DE PREDICCIONES

En este Notebook, se consolidan las predicciones de los tres modelos previamente entrenados (Financing, Investment, y Accounts). El objetivo es crear un único dataframe que agrupe las probabilidades de compra de cada cliente para los distintos grupos de productos.

9.1. Carga de predicciones:

- Se cargan las predicciones de los tres modelos entrenados, cada una en un dataframe con las columnas `pk_cid` (identificador de cliente), la predicción (0 o 1) de si el cliente comprará productos de ese grupo, y la probabilidad de compra.

9.2. Unión de dataframes:

- Los tres dataframes de predicciones se unen en uno solo, utilizando `pk_cid` como clave. En el dataframe resultante, se mantiene la columna `pk_cid` y las columnas de probabilidades de compra para cada uno de los tres grupos de productos.

9.3. Imputación de valores nulos:

- Los clientes que solo tienen una probabilidad de compra en un grupo de productos específico, tienen valores nulos en las otras dos columnas de probabilidades. Para facilitar los cálculos posteriores, se imputan estos valores nulos con 0, lo que permite tratar los datos de manera uniforme.

9.4. Preparación para el cálculo de ingresos:

- El dataframe resultante se exporta para ser utilizado en un documento Excel. En este documento, se calcula el ingreso esperado de cada cliente en función de las probabilidades de

compra, el ingreso estimado por la venta de cada grupo de productos, y la precisión de los modelos entrenados. Así, se obtiene una estimación aproximada del ingreso total que cada cliente generaría al comprar productos de los diferentes grupos.

10. SEGMENTACIÓN DE CLIENTES POTENCIALES

En este Notebook, el objetivo es realizar una nueva segmentación de los 10,000 clientes con la mayor probabilidad de compra, utilizando un modelo de clustering para identificar grupos que permitan personalizar una campaña de marketing.

10.1. Unión de información:

- Se cargan los datos con la información del encoding (df_encoding) y las predicciones de los modelos anteriores (df_predicciones).
- Se realiza una unión de ambos dataframes usando la columna pk_cid para identificar a los 10,000 clientes con la mejor probabilidad de compra.
- Se filtran los clientes seleccionados y se crea un dataframe con la información de cada cliente (características demográficas y financieras del df_encoding) junto con sus probabilidades de compra.

10.2. Exportación del dataframe:

- El dataframe resultante, con los 10,000 clientes y su información completa, se exporta a un archivo CSV para futuros análisis.

10.3. Segmentación (Clustering):

- Una vez preparado el dataframe, se aplica un algoritmo de clustering (K-Means) para segmentar a los clientes en 4 clusters.
- Estos clusters se interpretan para identificar perfiles de clientes basados en las características demográficas, financieras y las probabilidades de compra.

10.4. Interpretación de los clusters:

- Se realiza un análisis detallado de los 4 clusters, identificando patrones que permitan diseñar campañas de marketing personalizadas. Esto incluye determinar qué tipo de productos (de los tres grupos: Financing, Investment y Accounts) son más adecuados para cada grupo de clientes.

10.5. Uso para campañas de marketing:

- Con la información de los clusters, se pueden crear estrategias de marketing específicas, utilizando correos electrónicos personalizados para ofrecer productos financieros específicos a cada uno de los grupos.

DESPUÉS DE ESTO HAY QUE AÑADIR LA INTERPRETACIÓN DEL 2º CLUSTERING, LA EXPLICACIÓN DE LOS MAILS, LOS KPIS Y NEXT STEPS.