

**UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

REDISEÑO DE PROCESO DE VENTAS DE LA EMPRESA ALFAPEOPLE

**PROYECTO DE GRADO PARA OPTAR AL GRADO DE MAGÍSTER EN
INGENIERÍA DE NEGOCIOS CON TECNOLOGÍAS DE INFORMACIÓN**

ALFONSO JAVIER CASANOVA VILLA

**PROFESOR GUÍA:
EZEQUIEL MUÑOZ KRSULOVIC**

**MIEMBROS DE LA COMISION:
EDUARDO CONTRERAS VILLABLANCA
MONICA CORTES HIDALGO**

**SANTIAGO DE CHILE
2021**

RESUMEN EJECUTIVO

Existe una premisa importante en el mundo de los negocios que indica la dificultad de adquirir un nuevo cliente en comparación de mantener uno existente. Es un concepto utilizado en blogs, conferencias, videos, entre otros. Sin embargo, existe una tendencia que siguen la mayoría de las empresas exitosas: basarse en adquirir un mayor segmento del mercado, intentando transformarse en el más importante o en el único. Este pensamiento genera un importante descuido en la base instalada, elemento que es posible de comprobar sólo con revisar los procesos de cada organización. Lo descrito fue comprobado en la organización para la cual el autor trabaja y aplico cabalmente, identificando no tan solo que no existe trabajo dedicado a la base instalada, sino que tampoco fue pensado en los procesos de venta.

Este proyecto se basa en proponer un nuevo proceso de negocio basado en la metodología de rediseño de procesos utilizando patrones, que permite identificar puntos de mejora basados en buenas prácticas de industria y en la revisión de los datos actuales de la empresa AlfaPeople para en búsqueda de generación de nuevos negocios en clientes existentes, aplicando técnicas de segmentación de clientes basados en RFM.

Palabras claves: RFM, Procesos, Base Instalada

*Para mi hijo Ismael quien me llena de alegría y ganas de ser mejor cada día.
A mis padres, quienes me educaron con esfuerzo y amor.*

AGRADECIMIENTOS

Primero quisiera agradecer a mi familia por su paciencia y cariño.

A mis compañeros del magister, quienes fueron un pilar fundamental cuando iniciamos este camino. Tuve la suerte de conocer a grandes personas y profesionales.

A mi profesor guía Ezequiel Muñoz, siempre con gran disposición y apoyo. Quien dio grandes consejos y guías para ser mejor estudiante/profesional.

A mis “vecinos”, quienes constantemente con risas me empujaban a finalizar este proceso.

A Laura y Ana María, a quienes les agradezco mucho su paciencia y entendimiento. Y para todas aquellas personas que, en este largo tiempo, se han encargado de dar sus fuerzas para finalizar de la mejor manera posible.

TABLA DE CONTENIDO

CAPITULO 1. INTRODUCCIÓN Y CONTEXTO	1
1.1 Antecedentes de la Industria	1
1.1.1 Industria de CRM a nivel mundial.....	1
1.1.2 Industria de CRM a nivel local.....	2
1.2 Descripción General de la Empresa.....	3
1.2.1 Visión.....	4
1.2.2 Misión.....	5
CAPITULO 2. JUSTIFICACION Y CONTEXTO.....	7
2.1 Acerca del problema y su justificación.....	7
2.2 Objetivos del Proyecto de Investigación y resultados esperados	7
2.2.1 Objetivo General.....	7
2.2.2 Objetivos específicos y resultados esperados.....	7
2.3 Alcance	8
2.3.1 Dentro de alcance	8
2.3.2 Fuera de alcance.....	8
2.4 Riesgos Potenciales	9
CAPITULO 3. MARCO TEÓRICO	10
3.1 Metodología de Ingeniería de Negocios.....	10
3.2 Descubrimiento de información	14
3.3 Clustering	15
3.3.1 KMeans	16
3.3.2 DBSCAN	18
3.4 RFM.....	18
CAPITULO 4. INVESTIGACIÓN / PROYECTO APLICADO	23
4.1 Levantamiento de la Situación Actual.....	23
4.1.1 Posicionamiento Estratégico	23
4.1.2 Balanced Scorecard	24
4.1.3 Modelo de Negocios.....	26
4.1.4 Arquitectura de Procesos	27
4.1.5 Modelamiento Detallado de Procesos (AS – IS)	27
4.2 Diagnóstico de la Situación Actual	30

4.3	Dirección de cambio	34
4.4	BPMN del nuevo proceso (Tercer nivel)	35
CAPITULO 5. PROPUESTA DE APOYO TECNOLÓGICO.....		39
5.1	Especificación de Requerimientos.....	39
5.1.1	Requerimientos Funcionales	39
5.1.2	Requerimientos No Funcionales	39
5.2	Arquitectura Tecnológica.....	39
5.2.1	Descarga de datos	40
5.2.2	Aplicación de RFM	41
CAPITULO 6. EVALUACIÓN DEL PROYECTO		53
6.1	Flujo de Caja.....	53
6.2	Análisis de Sensibilidad	56
CAPITULO 7. CONCLUSIONES		58
CAPITULO 8. BIBLIOGRAFIA		59
CAPITULO 9. ANEXOS.....		60
ANEXO A: Código de RFM.....		60

ÍNDICE DE ILUSTRACIONES

Figura 1. Localización de AlfaPeople nivel mundial.....	3
Figura 2. Ontología de diseño de negocios	11
Figura 3. Esquema de ingeniería de negocios	12
Figura 4. Arquitectura de macro procesos y sus relaciones	13
Figura 5. Proceso de KDD.....	14
Figura 6. Ejemplo de dos aplicaciones de KMeans en dos fuentes de datos. Tomada de (MacKay, 2003)	17
Figura 7. Ejemplo método del codo	18
Figura 8. Posicionamiento estratégico.....	24
Figura 9. Modelo de negocios	26
Figura 10. Nivel 0	27
Figura 11. Nivel 1	28
Figura 12. Nivel 2	29
Figura 13. Cierre de proyecto estratégico CRM	30
Figura 14. Cantidad de oportunidades por cliente	32
Figura 15. BPMN del nuevo proceso.....	35
Figura 16. Modelo de entidad relación del archivo en Excel	41
Figura 17. Lectura de datos en Python.....	41
Figura 18. Análisis de datos de tipos de oportunidades	42
Figura 19. Cálculo de R en Python.....	42
Figura 20. Obtención de última fecha de cierre del set de datos.....	43
Figura 21. Diferencia entre último cierre y fecha de cierre del registro.....	43
Figura 22. Días de cierre de registros	43
Figura 23. Análisis de distribución de los datos.....	44
Figura 24. Aplicación del método del codo para R	44
Figura 25. División de clusters para R.....	45
Figura 26. Generación de nuevo Data frame para cálculo de F	45
Figura 27. Agrupación de cantidad de oportunidades ganadas para cálculo de F	45
Figura 28. Cambio de nombre de columna a F	46
Figura 29. Análisis de datos de F	46
Figura 30. Aplicación del método del codo para F.....	47
Figura 31. Suma de ingresos para generar M	48
Figura 32. Análisis de los datos de M.....	48
Figura 33. Cálculo de los puntajes de RFM.....	49
Figura 34. Análisis de comportamiento de Ingresos VS Frecuencia	50
Figura 35. Análisis de comportamiento de Ingresos VS Reciente	50
Figura 36. Análisis de comportamiento de Frecuencia VS Reciente	51
Figura 37. Gráfico de tipo de segmento	51

Figura 38. Análisis de VAN y TIR	57
---	----

ÍNDICE DE TABLAS

Tabla 1: Marketshare y crecimiento según Gartner (mayo 2015)	2
Tabla 2: Listado de Partners CRM definidos por Microsoft.....	3
Tabla 3: Objetivos generales y específicos.....	7
Tabla 4: Riesgos potenciales	9
Tabla 5: Inicio de cálculo de RFM.....	20
Tabla 6: Cálculo de R	21
Tabla 7: Cálculo de F y M.....	21
Tabla 8: Ejemplo de score de RFM	22
Tabla 9: Balanced Scorecard.....	24
Tabla 10: Técnica de los 5 por qué.....	31
Tabla 11: Análisis de clientes con un negocio	33
Tabla 12: Dirección de cambio	34
Tabla 13: Análisis de sueldos	53
Tabla 14: Tasa de descuento	54
Tabla 15: Flujo de caja	55
Tabla 16: Indicadores de proyecto.....	55
Tabla 17: Análisis de sensibilidad.....	56

CAPITULO 1. INTRODUCCIÓN Y CONTEXTO

AlfaPeople se desempeña en el ámbito de la consultoría de negocios e implementador de software, entregando valor a las empresas en la implementación de un Enterprise Resource Planning (ERP), Customer Relationship Management (CRM) u otras soluciones que puedan ser utilizadas como Software as a Services (SaaS). Es partner categoría GOLD de Microsoft para su producto Dynamics 365, siendo uno de los principales implementadores a nivel mundial.

El año 2016 se decidió potenciar el área de CRM que poseía la organización y se contrató a profesionales que contaban con experiencia de implementación del producto, por los que aumentaron los costos del equipo. Fue necesario aumentar la meta de ventas y los proyectos que se deberían ejecutar. En el año 2017 las ventas sólo llegaron al 50% de lo proyectado para el producto, por lo que existen elementos que no están permitiendo que AlfaPeople consiga su cuota de ventas.

Puesto que la modificación a nivel organizacional requiere que se potencie el área señalada, este documento contendrá información de negocios y de mercado relacionadas con esta división junto con la información de la industria de uno de productos que ofrece Microsoft para soluciones empresariales, llamado Microsoft Dynamics 365 y su concepto de CRM.

1.1 Antecedentes de la Industria

1.1.1 Industria de CRM a nivel mundial

El mercado de CRM a nivel mundial se ha encontrado en continuo crecimiento desde el año 2013 en adelante, debido a la alta demanda de obtener SaaS que permitan mejorar el desempeño empresarial.

De acuerdo a Gartner¹, el análisis realizado el año 2015 refleja aproximadamente un crecimiento de un 13% o USD 14 billones de dólares distribuidos en los 10 principales vendedores de CRM del mundo, destacando principalmente a Salesforce y Microsoft en términos de crecimiento de mercado.

¹ Rivera, J.& Meuler, R. *Gartner Says Customer Relationship Management Software Market Grew 13.3 Percent*. Mayo 19, 2015 < <http://www.gartner.com/newsroom/id/3056118>>

Tabla 1: Marketshare y crecimiento según Gartner (mayo 2015)

Company	2014 Revenue	2013 Revenue	2014 Market Share (%)	2013 Market Share (%)
Salesforce	4,268.5	3,330.2	18.4	16.3
SAP	2,809.4	2,621.3	12.1	12.8
Oracle	2,115.2	2,060.8	9.1	10.1
Microsoft	1,438.6	1,181.8	6.2	5.8
IBM	873.1	792.1	3.8	3.9
Others	11,681.9	10,474.7	50.4	51.1
Total	23,186.7	20,460.9	100	100

La tendencia de crecimiento de Microsoft se ha mantenido hasta el año 2016 ², debido al aumento de los catálogos de productos que han sido llevados a la nube y la modificación de los ya existentes, como es el caso de Dynamics 365 (Los productos de CRM y ERP principalmente fusionados como servicio en la nube, los cuales se explicarán en este documento más adelante).

1.1.2 Industria de CRM a nivel local

En Chile, las tendencias mundiales también se ven reflejadas al momento de preferir un CRM. El mercado se ve compartido entre las marcas, los partners y las competencias de cada partner, elemento relevante que cada cliente debe tener en consideración para elegir su próxima especialización.

Sólo por el área de Microsoft, es posible tener una de las 3 competencias de acuerdo con su nivel de especialización y producto, correspondientes a “Partner”, “Partner Silver” y “Partner Gold”, siendo la última la de mayor especialización.

De acuerdo a Microsoft, el ecosistema de partners que se puede identificar son los siguientes (Ordenados de acuerdo a relevancia), tal como se señala en la .

² *Pulso*, enero 26, 2017 < <http://www.pulso.cl/empresas-mercados/ganancias-de-microsoft-suben-36-impulsados-por-mayor-demanda-de-servicios-cloud-en-4t16/>>

Tabla 2: Listado de Partners CRM definidos por Microsoft

Partner	Competencia
CJSC Softline Trade	Gold
AlfaPeople	Gold
Axxon Consulting	Gold
Zenith Consulting	Silver
Beacon42	Silver
Defontana Chile	Silver

1.2 Descripción General de la Empresa

AlfaPeople una empresa de origen danés fundada en el año 2006, dedicada principalmente a la implementación de Microsoft Dynamics AX 4.0 (Software de ERP). Debido al continuo crecimiento del producto, de Microsoft en el área de soluciones empresariales y de la empresa misma, al día de hoy es un importante partner de Microsoft en implementación de proyectos de Microsoft Dynamics 365 (principalmente los módulos relacionados a CRM y ERP) y soluciones Cloud (Cloud Solution Provider o CSP en adelante). Al día de hoy se encuentra ubicado en más de 15 países y con más de 500 empleados alrededor del mundo.

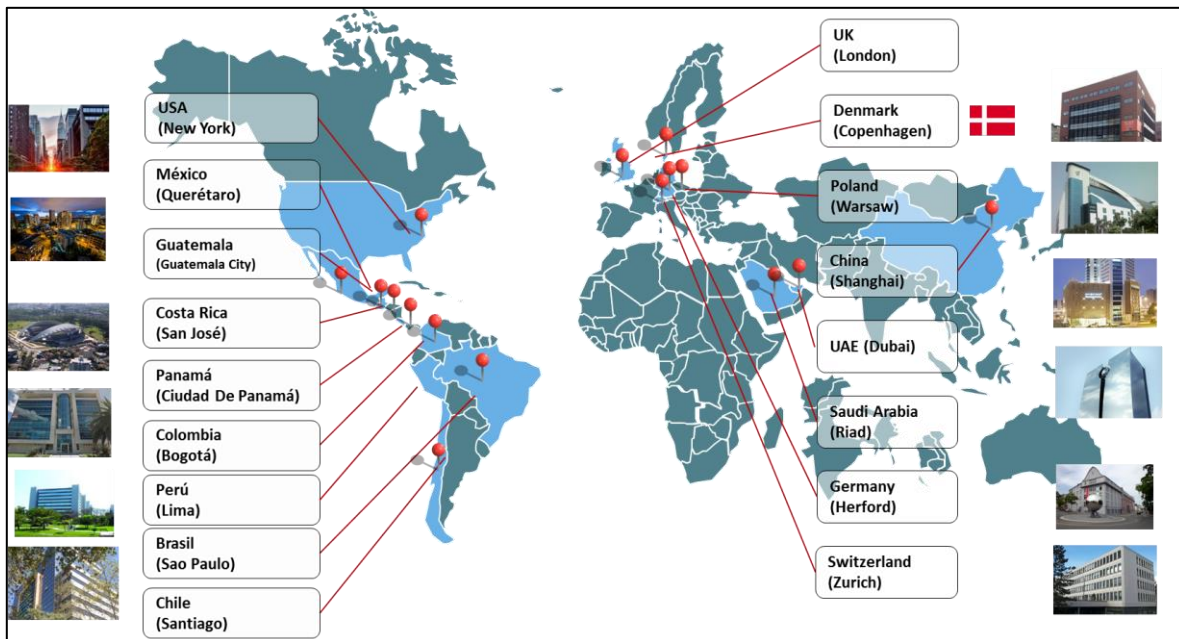


Figura 1. Localización de AlfaPeople nivel mundial

1.2.1 Visión

La visión fue obtenida desde la página web de AlfaPeople. la cual fue generada desde Dinamarca y fue traducida para ser presentada en múltiples idiomas. En un simple análisis, la visión no refleja la visión de la compañía, sino más bien una explicación de las funcionales aplicadas a implementaciones remotas.

“Creemos que el software como servicio es el modelo de negocio para el futuro que hoy en día ya se encuentra disponible. Para la mayoría de los clientes, este modelo de licencia y el modelo de despliegue en línea es la base óptima para su organización. Proporcionando un equilibrio entre precio, calidad, funcionalidad, escalabilidad y seguridad. Si la flexibilidad, la intimidad o la continuidad son las prioridades, también entregamos soluciones de soporte que puede desplegarse en instalaciones fijas. Sin embargo, creemos que con el tiempo estas prioridades serán mejor con software como servicio.

También se beneficiará de nuestros servicios de consultoría en línea. Todos nuestros procesos de servicio se han optimizado para ofrecer el mejor de los servicios en consultoría, con un equipo global de consultores y expertos a la mano para asegurar de que recibe el servicio que merece. De esta manera, nuestros servicios de consultoría en línea no sólo son económicamente competitivos y convincentes, sino también más seguros

Algunos de nuestros proyectos de implementación implican trabajar en el lugar del cliente el 20% del tiempo. Sin embargo, tener consultores de TI en la oficina de un cliente durante días y semanas es ineficiente y costoso, por tanto, debe ser limitado. Los costos adicionales de viaje, alojamiento y tiempo de viaje son extras muy visibles. Estos costos ocultos se acumulan con el tiempo y son contraproducentes para su organización. Por ejemplo, la ineficiencia en el lugar se produce debido a los desafíos implicados en la planificación y coordinación con el horario de trabajo de un cliente, y añade costes innecesarios en muchas situaciones. Más importante aún, lo que limita el trabajo in situ permite el uso completo de especialistas globales que pueden abordar las tareas más exigentes.

Creemos que hay una mejor manera si comparte nuestra visión para el software en línea y servicios en línea.”³

³ AlfaPeople. *Visión*. < <https://alfapeople.com/cl/acerca-de-nosotros/>>

1.2.2 Misión

La misión sigue el mismo ejemplo que la visión y también se obtuvo de la página web. Sólo se tradujeron de acuerdo a lo definido en Dinamarca. La misión tiende a carecer de simplicidad y no pareciera estar enfocada al cliente, sino más bien parece explicación de lo que se va a presentar en una posible solución de implementación.

“Cualquier organización necesita seleccionar cuidadosamente los procesos de negocio, a la medida de las necesidades específicas, siendo estos únicos y competitivos. Si una solución de TI estándar 100% es su elección, o bien no tiene procesos de negocio específicos o su solución no podrá apoyar los procesos en el futuro.

Nuestra misión es permitir el crecimiento y proporcionar procesos de negocio específicos y competitivos con el apoyo de soluciones de Microsoft Dynamics flexibles y robustas de la industria, junto con los complementos y mejoras adicionales.

Si usted tiene un negocio único y una solución de TI estándar, habrá lagunas. Esto significa ya sea el cambio de la solución o frenar el progreso de su organización, y en cualquiera de los dos casos los procesos son largos y costosos.

Creemos que nada va a proporcionar un mejor retorno de la inversión que construir o mejorar sus procesos de negocio específicos con un ajuste perfecto soportado en una solución TI. Ninguna empresa se convierte en un líder en el mercado sin una solución única. Permítanos empoderarlo: vamos a diseñar, construir y entregar.

También trabajamos para consolidar los sistemas locales en una misma solución global y estamos comprometidos a ofrecer a nuestros clientes lo que llamamos ‘una organización, una solución’. Ningún sistema de TI debe ser una isla, tendrá que ser parte de la organización. Una solución global puede integrar varias de las mejores aplicaciones integradas. Nuestra prioridad fundamental es eliminar todos los sistemas redundantes y reconocer los datos de los procesos empresariales de misión crítica en toda la organización, para asegurar que sea lo más eficaz y eficiente posible.

El diseño de la solución se origina a partir de los requerimientos del cliente. Nuestra misión es identificar, construir y apoyar en las mejores prácticas corporativas a cada cliente, delimitando los requisitos reales o percibidos de la organización o la necesidad de sus personalizaciones, al

mismo tiempo que se logra una excelente adopción del sistema por parte de los usuarios.”⁴

⁴ AlfaPeople. *Misión*. < <https://alfapeople.com/cl/acerca-de-nosotros/>>

CAPITULO 2. JUSTIFICACION Y CONTEXTO

2.1 Acerca del problema y su justificación

Según Amy Galio de Harvard Business Review (2014), “adquirir un nuevo cliente cuesta entre 5 y 25 veces más que retener a un cliente existente y con sólo retener a un 5% de los clientes se podrían aumentar los ingresos entre un 25% a un 95%” (párrafo 1).

AlfaPeople no tiene un proceso definido para los clientes existentes para generar nuevos negocios. Los puntos de contacto en todo proceso de un proyecto suelen ser más reactivos que proactivos para generar nuevas iniciativas.

Este proyecto tiene como objetivo trabajar en un plan que permita mejorar la rentabilidad de los clientes existentes, identificando los que podrían generar nuevos negocios y seleccionarlos para múltiples campañas de marketing con ofrecimientos de distintas nuevas soluciones de negocios.

2.2 Objetivos del Proyecto de Investigación y resultados esperados

2.2.1 Objetivo General

Rediseñar el proceso de marketing/ventas de la organización incorporando una clasificación de los clientes, utilizando datos de industria a la que pertenece y su comportamiento de cierre de negocios con AlfaPeople.

2.2.2 Objetivos específicos y resultados esperados

Tabla 3: Objetivos generales y específicos

Objetivo específico	Resultado esperado
Obtener y diagramar el proceso actual de la organización	Diagrama de situación actual en la organización
Rediseñar el proceso de marketing/ventas en el que afectará la implementación de este proyecto	Diagrama de rediseño de proceso y su comunicación con la situación actual
Analizar los datos del actual sistema y generar segmentación de clientes	Aplicación de RFM en datos del cliente para categorizarlos basado en el comportamiento registrado en CRM.

2.3 Alcance

2.3.1 Dentro de alcance

- Este proyecto será aplicable sólo a nivel Chile, pero utilizará los datos de Latam.
- Los resultados serán un análisis que permitirá segmentar los clientes de la base de marketing, identificando los criterios con los que se deberán segmentar.

2.3.2 Fuera de alcance

- No considera implementación en otras unidades organizacionales.
- No considera inyectar los resultados a las oportunidades de negocio
- No considera desarrollos sobre la plataforma que involucren código directamente no soportado.
- No considera múltiples recursos de trabajo en el proyecto.

2.4 Riesgos Potenciales

Tabla 4: Riesgos potenciales

Nº	Tipo	Riesgo	Probabilidad	Impacto	Estrategia de Mitigación
1	Gestión de Proyecto	Baja cooperación entre áreas	Alto	6	Comunicar oportunamente a las áreas involucradas la existencia de este proyecto y programar reuniones de revisiones en conjunto
2	Gestión de Proyectos	Atrasos en Gantt	Medio	4	Evaluar estrategias de apoyo por parte de la organización
3	Técnico	No contar con disponibilidad de obtención de datos	Medio	6	Investigar el desarrollo de un elemento de exportación masivo no nativo del producto
4	Técnico	Dificultades de integración de herramientas	Bajo	3	Entregar un vínculo con un acceso directo a las recomendaciones de cliente
5	Proyecto	Bajo interés en la organización de utilizar el proyecto	Medio	3	Generar estrategias de gestión del cambio para los usuarios con más posibilidad de uso
6	Proyecto	No contar con el tiempo necesario para la evaluación del proyecto	Alto	5	Declarar previamente los tiempos necesarios para la construcción del proyecto

CAPITULO 3. MARCO TEÓRICO

Este capítulo consiste en la explicación de los elementos teóricos que serán utilizados para la elaboración de esta tesis. Primero se describirá la metodología explicada por el Departamento de Ingeniería Industrial de la Universidad de Chile en el Magister de Ingeniería de Negocios con Tecnologías de Información. Luego, se abordarán conceptos teóricos sobre los procesos de descubrimiento de la información y minería de datos.

3.1 Metodología de Ingeniería de Negocios

La metodología por aplicar en este proyecto corresponde a la que se estudia en Master in Business Engineering (MBE), llamada “Ingeniería de Negocios”. Es explicada en distintos cursos en el transcurso del magister, complementada con el libro “Business Engineering and Service Design” (Barros, 2017).

Ésta tiene como objetivo definir y entregar herramientas para el diseño de negocios. Para lograrlo, cuenta con un enfoque multidisciplinario que involucra el uso de elementos como modelos de negocio, arquitectura empresarial, definición de procesos, analítica, tecnología de software, entre otros, los que integrados permiten la definición de la organización como un todo y no de elementos aislados.

En la Figura 2. Ontología de diseño de negocios, se puede visualizar un modelo conceptual de definición de empresas, del cual se generó una metodología jerárquica de diseño de una organización, partiendo desde detalles globales hasta componentes específicos.

Los elementos de la metodología de diseño de una empresa son los que se visualizan en la La utilización de la metodología, significa la aplicación de modelos de referencia que poseen la gran mayoría de organizaciones que se dedican a un rubro en particular y que fueron exitosas en diseñar sus organizaciones. Esto permite que cualquier empresa cuente con un punto de partida de componentes y así puedan definir sólo elementos propios de su organización, mediante la definición de los procesos particulares. Los modelos señalados son llamados Patrones de Negocio (PN) y Patrones de Proceso de Negocio (PPN).

Los Patrones de Negocio indican las diferentes componentes, estructuras y relaciones que una empresa podría seguir para la prestación de servicios para sus clientes, mientras que los Patrones de Proceso de Negocio indican cómo se puede implementar operativamente lo diseñado.

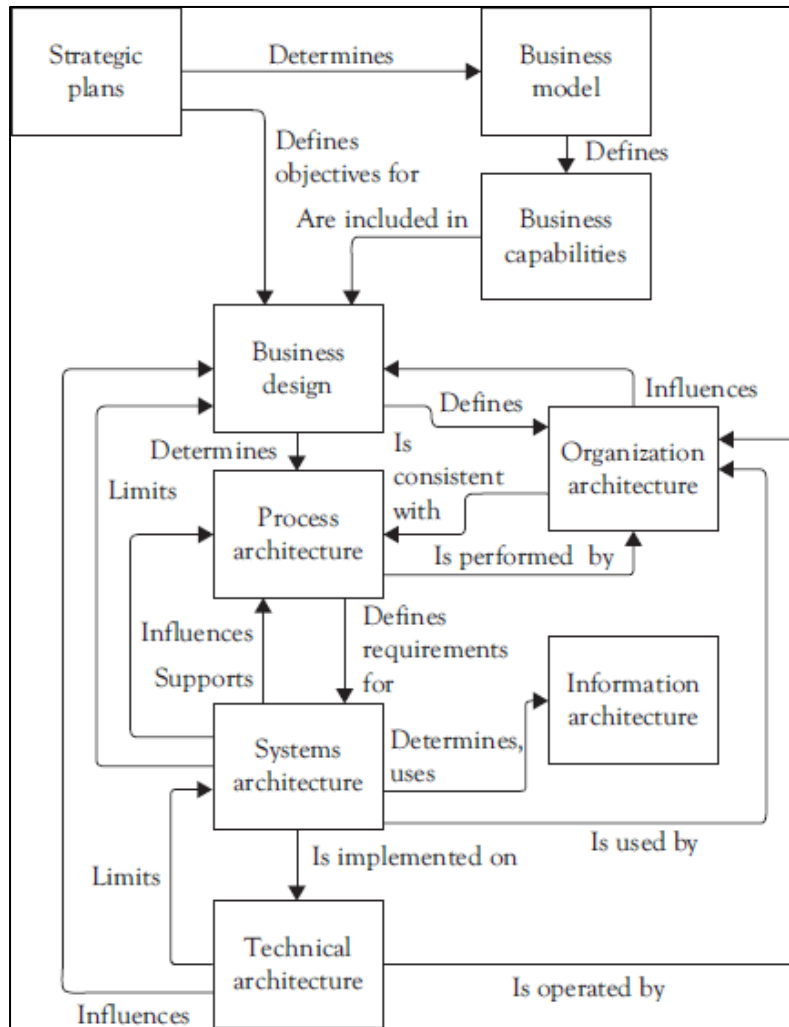


Figura 2. Ontología de diseño de negocios

En síntesis, la metodología de Ingeniería de Negocios permite el diseño de negocios de una manera sistémica e innovadora, permitiendo que las organizaciones sean más competitivas, eficaces y eficientes.

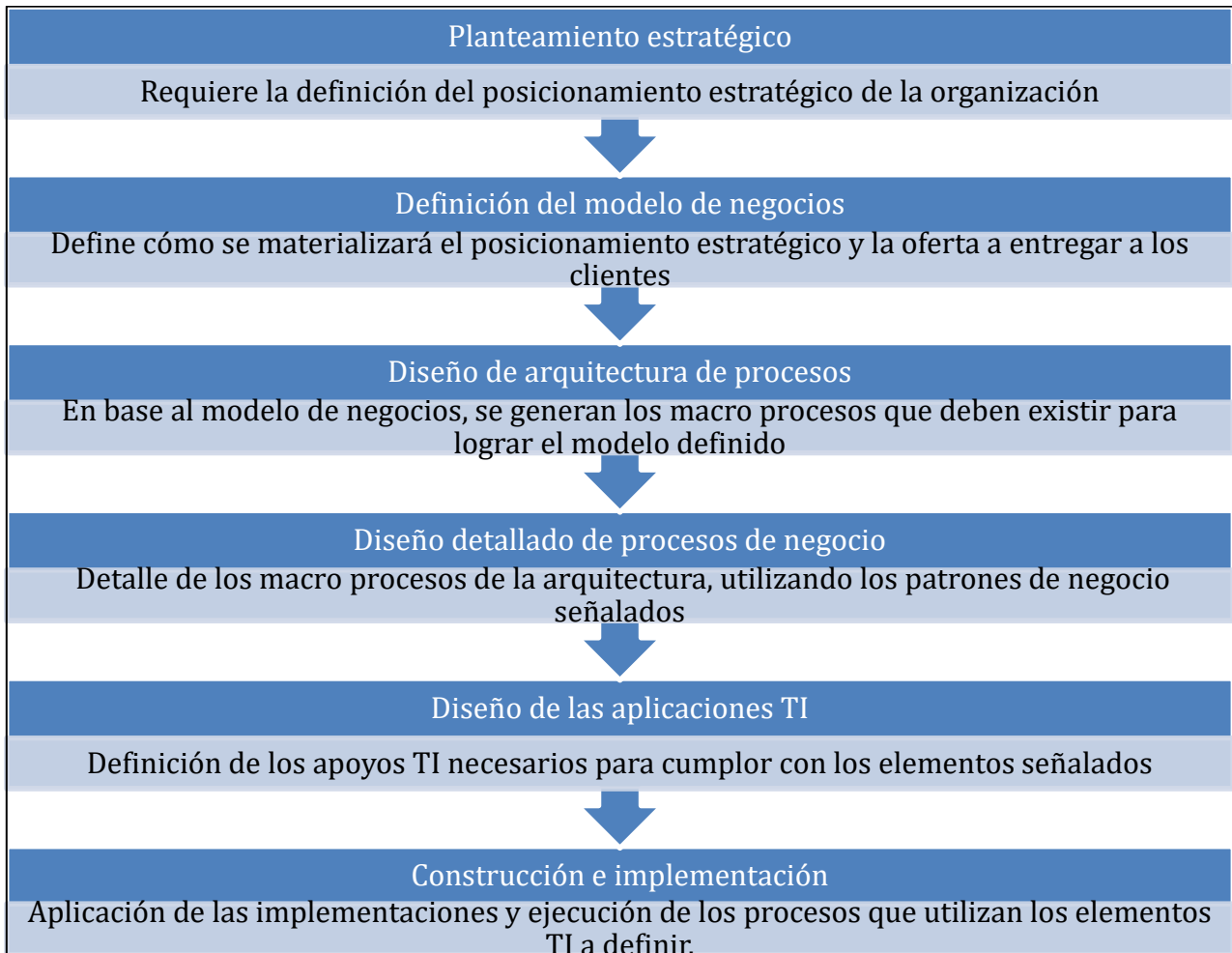


Figura 3. Esquema de ingeniería de negocios

La arquitectura de procesos de negocio es la definición de las relaciones coordinadas entre los procesos y lógicas de negocios que permitan la automatización de los procesos, utilizando componentes de TI. Esto permite la implementación de los componentes necesarios para cubrir las capacidades de un negocio.

Para el diseño de la arquitectura de negocios, se utilizan patrones, tal como fue mencionado en el punto anterior. Estos patrones corresponden a agrupaciones de procesos, basados en buenas prácticas y experiencia de organizaciones. Estas agrupaciones son llamadas macroprocesos (o macros) y se dividen en cuatro tipos:

- **Macro 1:** Conjunto de procesos que permiten la ejecución de los bienes y/o servicios de una organización. Estos procesos contienen desde las primeras

comunicaciones con un cliente hasta que se le entregan los elementos requeridos. Este macroproceso tiene por nombre *cadena de valor*.

- **Macro 2:** Corresponde a los procesos que crean nuevas capacidades para la organización, de tal manera que permitan que se mantenga competitiva. Permite la creación de elementos tales como nuevos productos, servicios o modelos de negocio. Principalmente este macroproceso permite enfocarse en las capacidades de innovación que tiene una organización
- **Macro 3:** Contiene los procesos de planificación de un negocio, es decir, lo que define hacia donde se mueve la organización al futuro, mediante la definición de programas y planes.
- **Macro 4:** Corresponden a los procesos de apoyo que permiten que las otras macros señaladas operen. Estos se dividen en cuatro grupos: Recursos financieros, humanos, infraestructura y materiales.

Estos macroprocesos permiten el modelamiento de cualquier organización que tenga una estructura integrada y coherente. En la Figura 4. Arquitectura de macro procesos y sus relaciones se puede visualizar las relaciones que poseen los procesos, sus flujos de información y requerimientos entre ellos.

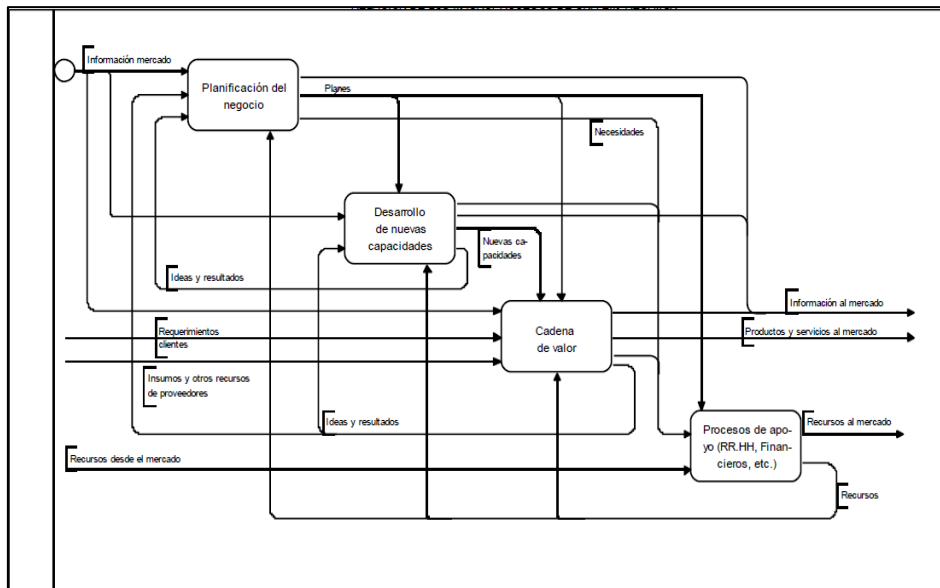


Figura 4. Arquitectura de macro procesos y sus relaciones

3.2 Descubrimiento de información

Las organizaciones frecuentemente almacenan información en sus sistemas con el único fin de dejar registro de que “algo” fue realizado en un periodo en particular. Pasado el tiempo, siguen manteniendo la información y no realizan mayores acciones con ella, siendo que pueden ser elementos ricos de obtención de información en particular.

Bajo esa premisa, nace un concepto llamado *Knowledge Discovery in Databases* (KDD), el cual permite que, al realizar una serie de actividades iterativas, se pueda obtener información que podría ser relevante para la toma de decisiones.

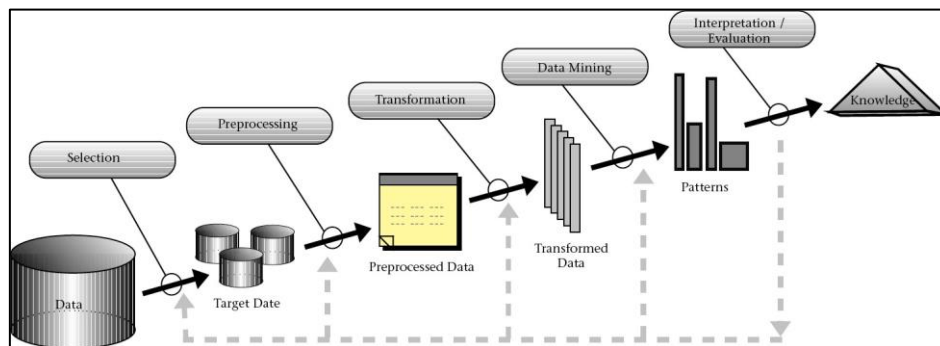


Figura 5. Proceso de KDD

Este proceso consta de 5 fases, tales como se muestran en la Figura 5. Proceso de KDD correspondientes a:

1. Selección de datos: Corresponde a la etapa en la que se determina de dónde se obtendrá y el tipo de información con el que se trabajará.
2. Pre procesamiento: Es la etapa en la que se preparan y limpian los datos obtenidos desde las distintas fuentes para que sean manejables por los elementos de las etapas posteriores. Se utilizan distintas estrategias para manejar los datos faltantes.
3. Transformación: Corresponde al tratamiento inicial de los datos, transformación y generación de nuevas variables de acuerdo con la estructura de datos con la que se trabajará. En esta etapa se normalizan los datos de una manera necesaria para la siguiente fase.
4. Data Mining: Fase de modelamiento y aplicación de los algoritmos inteligentes con el objetivo de obtener los patrones de datos previamente desconocidos y

potencialmente útiles, ocultos en la información. Para este proyecto, se realizará un descubrimiento de grupos con la información disponible en el sistema CRM, por lo que se aplicarán algoritmos de Clustering. (A explicar en el punto 3.3 Clustering)

5. Interpretación y evaluación: Etapa en la que se identifican los patrones obtenidos y se evalúa su utilidad.

3.3 Clustering

Clustering corresponde a una técnica de aprendizaje no supervisado utilizado para buscar estructuras con los datos que generan agrupamientos naturales o “clústeres”. Un clúster generalmente es un área de densidad en un espacio, en donde los datos de un dominio están más cercanos a un núcleo que otros. Con eso se señala que un clúster puede tener un centroide, correspondiente a una característica en el espacio.

Los tipos de clustering que se pueden realizar son:

- Exclusivo (Particionamiento): Agrupamiento de datos que sólo pueden pertenecer a un grupo.
- Aglomerativo: En este método, todos los datos son un clúster. Se busca medir la distancia entre dos clústeres cercanos, disminuyendo la cantidad de clústeres.
- Overlapping: Se utiliza un método “fuzzy” para agrupar los datos. Cada punto es de un grupo de dos o más clústeres separados por grados de afinidad.
- Probabilístico: Esta técnica usa distribución probabilística para definir los grupos

Existen múltiples tipos de algoritmos para ejecutar esta técnica. Algunos miden similitud o miden distancias entre las características en el espacio definido, identificando alguna densidad entre sectores de observación. Puesto a que esta técnica puede trabajar con múltiples valores de datos numéricos, suele ser una buena práctica el escalar los datos antes de ejecutar los algoritmos.

Algunos de los algoritmos de clustering es necesario especificar el número de segmentos que se descubrirán en los datos, en otros es necesario especificar un mínimo de distancia entre las distintas observaciones.

De acuerdo con una publicación “10 Clustering Algorithms With Python. Machine Learning Mastery” (Brownlee, J., 2020), se ejemplifican 10 de los algoritmos más populares:

- Affinity Propagation
- Agglomerative Clustering

- BIRCH
- DBSCAN
- K-Means
- Mini-Batch K-Means
- Mean Shift
- OPTICS
- Spectral Clustering
- Mixture of Gaussians

Cada algoritmo otorga una aproximación diferente para encontrar grupos naturales en los datos. No existe un mejor algoritmo que otro y la mejor manera de identificar la utilidad de cada uno es ejecutando experimentos controlados en los datos.

Para buscar la información en los datos del sistema, se utilizará un tipo de clustering exclusivo para identificar los grupos a los que pertenecen los distintos clientes. Se utilizará el algoritmo de k-means y adicionalmente se segmentarán los clientes utilizando RFM

3.3.1 KMeans

KMeans es un algoritmo que se basa en intentar minimizar la varianza de puntos entre datos relacionados a un centroide.

Realiza iteraciones constantes basadas en una definición de clases/grupos y aleatoriamente busca su punto central (centroide), el cual se define como vectores de igual distancia entre grupos de puntos. Por cada grupo definido, se calcula nuevamente el centro entre estos grupos tomando la media de todos los vectores definidos en los grupos. Esta actividad se repite por una cantidad de iteraciones definidas o hasta que el grupo de centros no cambie mucho entre iteraciones.

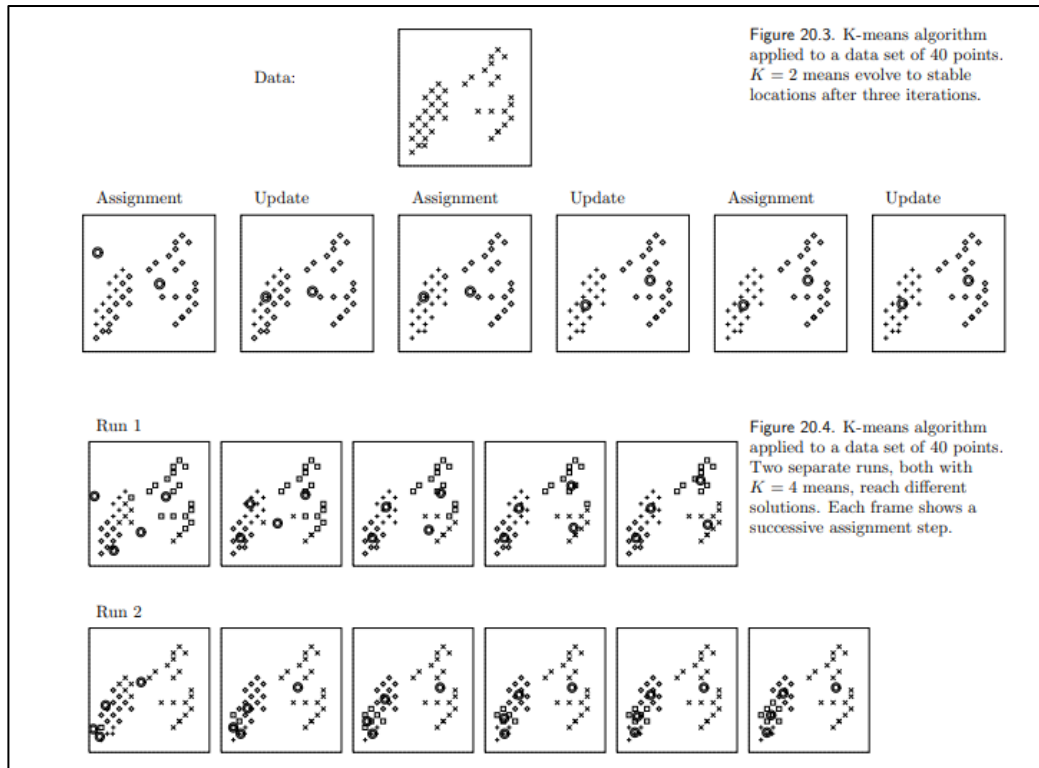


Figura 6. Ejemplo de dos aplicaciones de KMeans en dos fuentes de datos. Tomada de (MacKay, 2003)

La aplicación de este algoritmo busca minimizar la varianza intra-cluster y a su vez maximizar la varianza inter-clúster. Puesto que se requiere definir un número de grupos (K) en el algoritmo, es necesario contar con algún método que permita recomendar previamente, siendo el método del codo uno de los más populares.

3.3.1.1 Método del codo

Este método toma la distancia media de cada observación con su centroide. Clústeres más compactos significa que poseen menor distancia intra-cluster.

El método del codo se encarga de identificar un valor K que después de ese valor, no cambie sustancialmente la distancia media intra-cluster.

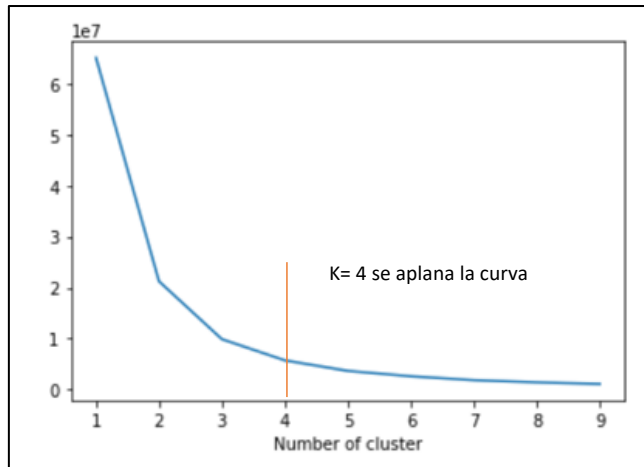


Figura 7. Ejemplo método del codo

La definición del K se basa en ejemplos visuales. Como se aprecia en la imagen anterior, desde K=4 ya no existe una amplia variación en la curva, recomendando ese valor en el algoritmo de KMeans.

3.3.2 DBSCAN

Es un algoritmo de clustering basado en densidad. Su propósito es calcular en cada vecindario los puntos más cercanos, basado en un punto en particular.

Funciona basado en 2 parámetros:

- Radio: Determina la longitud específica que podría incluir la cantidad de puntos mínima
- Puntos mínimos: Corresponde a la cantidad de datos mínimos que se requieren alrededor de una observación

3.4 RFM

RFM corresponde a un análisis como método para identificar respuestas a clientes en promociones de marketing y mejorar las respuestas ante comunicaciones. Es un modelo que tiene más de 40 años, su valor es asignar una evaluación o “score” a una

base de datos de clientes para generar acciones y ver su comportamiento con la organización durante el tiempo.

Este método funciona bajo tres premisas principales:

- Los clientes que han comprado recientemente tienen una mayor tendencia a responder a una promoción que aquellos que han comprado hace más tiempo. Esta es una premisa universal que es cierta en la mayoría de las industrias.
- Los clientes frecuentes tienen mayor tendencia a comprar nuevamente que clientes menos frecuentes.
- Clientes que gastan más, generalmente responden mejor que aquellos que gastan menos.

Estos principios son los que se encuentran en análisis detrás del método de RFM. Se codifican a los clientes en regiones o celdas, examinando sus características para generar una calificación.

Cada una de las premisas se definen en los siguientes conceptos:

- Recency (Actualidad): Determina la cantidad de tiempo que pasó desde una fecha determinada hasta la última compra
- Frequency (Frecuencia): Determina la cantidad de compras que ha realizado un cliente en un periodo de tiempo
- Monetary (Valor monetario): Determina el valor monetario de las compras que ha realizado el cliente en el periodo de tiempo.

Con estos valores, se genera un puntaje por cada cliente basado en las variables. Una de las técnicas más usadas es dividir los valores entregadas en cada una de las

variables, se dividen en porcentajes iguales (cuartiles o quintiles) y se suman/promedian/interpretan generando un puntaje de RFM.

A modo de ejemplo:

Tabla 5: Inicio de cálculo de RFM

Cliente	Días última compra	Cantidad de compras en x periodo	Promedio de compra	R	F	M
1	5	4	30			
2	7	2	10			
3	2	7	20			
4	4	6	40			

El primer paso consiste en identificar la compra reciente. Para ello, se dividirá el total de clientes en cuartiles puesto que en el ejemplo sólo se cuenta con 4 clientes. A esta división se le dará un puntaje de 1 a 4, en donde 1 es menos reciente y 4 es que ha comprado más recientemente. Según el análisis, el cliente 3 recibiría un puntaje de 4 mientras que el que recibe menor puntaje es el cliente 2. La completitud de la tabla es como se indica:

Tabla 6: Cálculo de R

Cliente	Días última compra	Cantidad de compras en x periodo	Promedio de compra	R	F	M
1	5	4	30	2		
2	7	2	10	1		
3	2	7	20	4		
4	4	6	40	3		

Posteriormente se continúa con la frecuencia y monto siguiendo la misma lógica que R:

- Se asignará 4 a los clientes con mayor frecuencia de compra y 1 a los clientes con peores valores
- Se asignará 4 a los clientes con mayor valor monetario y 1 a los clientes de menor valor monetario

Con esta premisa, la tabla queda según se indica:

Tabla 7: Cálculo de F y M

Cliente	Días última compra	Cantidad de compras en x periodo	Promedio de compra	R	F	M
1	5	4	30	2	2	3
2	7	2	10	1	1	1
3	2	7	20	4	4	2
4	4	6	40	3	3	4

Al finalizar, se puede sumar, ponderar o generar un score uniendo los valores de RFM. Para este ejemplo se sumará el puntaje:

Tabla 8: Ejemplo de score de RFM

Cliente	Días última compra	Cantidad de compras en x periodo	Promedio de compra	R	F	M	RFM
1	5	4	30	2	2	3	7
2	7	2	10	1	1	1	3
3	2	7	20	4	4	2	10
4	4	6	40	3	3	4	10

Con este análisis es posible concluir que los clientes 3 y 4 son más valiosos para la organización, identificando que:

- El cliente 3 realizó una compra hace poco y compra frecuentemente sin un valor monetario muy alto
- El cliente 4 no compra tan reciente ni frecuentemente con el cliente 3, pero cuando compra lo hace por montos más elevados

Sin embargo, el cliente 2 es un mal cliente para la organización, porque realiza compras largas, poco frecuentes y de bajo valor monetario.

Con este análisis se puede segmentar de mejor manera los clientes y dirigir más efectivamente las campañas a los clientes de interés para la organización.

CAPITULO 4. INVESTIGACIÓN / PROYECTO APLICADO

4.1 Levantamiento de la Situación Actual

4.1.1 Posicionamiento Estratégico

Para la identificación del posicionamiento estratégico, se utilizará el modelo Delta propuesto por Arnoldo Hax.

AlfaPeople como tal se refleja como una empresa consultora de negocios, que posee conocimiento en la implementación de sistemas de ERP y CRM de Microsoft.

El producto se ha ido renovando cada 6 meses y la marca va capacitando a sus Partners en cómo ofrecer el producto y cómo realizar su implementación. Microsoft ha invertido en adquirir nuevas empresas y mejorar los servicios que tiene online para simplificar a los clientes en obtener su plataforma empresarial.

Al observar la misión y visión, se puede obtener lo siguiente:

Misión

*“Nuestra misión es permitir el crecimiento y proporcionar procesos de negocios específicos y competitivos con el apoyo **de soluciones de Microsoft Dynamics flexibles y robustas de la industria, junto con los complementos y mejoras adicionales...**”*

Visión

*“Vamos a ser los líderes del mercado en **servicios de Microsoft Dynamics** porque nuestra gente entrega los mejores y más competitivos servicios en línea, no en el sitio...”*

Es decir, ambas apuntan a un posicionamiento cercano a la esquina de “Mejor producto” que señala Hax.

Al analizar el comportamiento organizacional, se puede obtener que siempre la conversación con distintos clientes termina siendo una conversación de precios. Al momento de cotizar, los clientes suelen preferir la propuesta con un menor costo.

Por tanto, el posicionamiento estratégico corresponde a **Low Cost**.

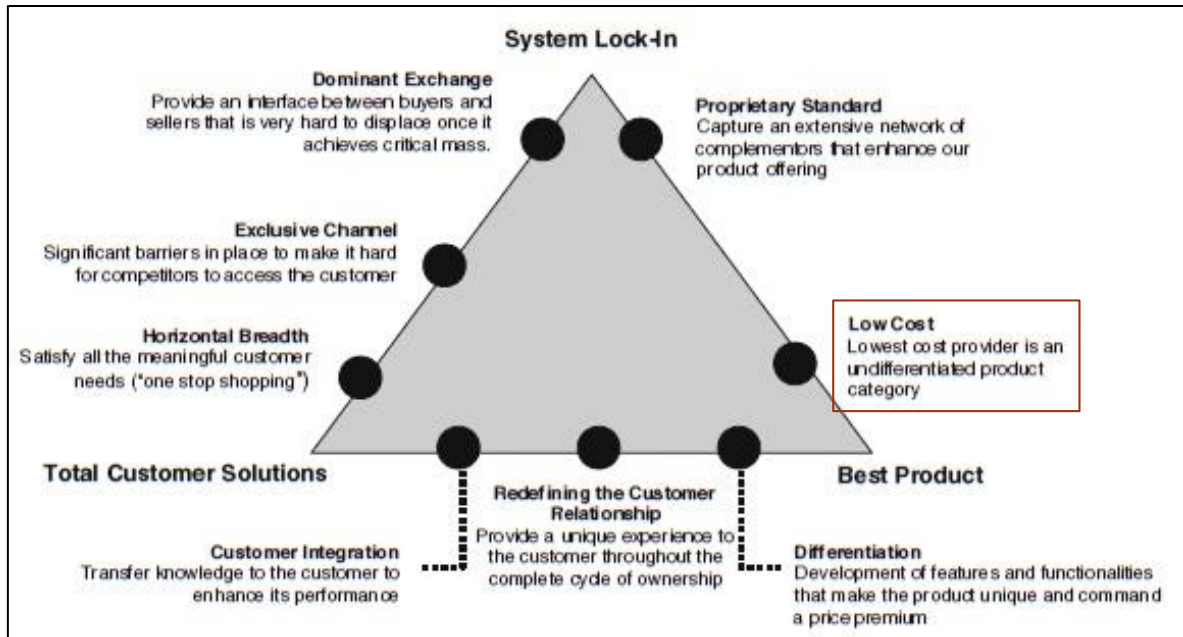


Figura 8. Posicionamiento estratégico

4.1.2 *Balanced Scorecard*

La organización no tiene un Balanced Scorecard escrito, por lo que se ha definido en base a los datos investigados, generando la siguiente tabla:

Tabla 9: Balanced Scorecard

Perspectiva	Objetivos	Indicadores	Responsable
Financiera	Cumplir con la meta de ventas de USD \$3,5M a final del año 2018 venta de proyectos	Suma montos Ingresos anuales	Área comercial CRM y dependientes
	Cumplir con la meta de ventas de licenciamiento del producto de USD \$14.000 a final del año 2018	Suma montos Suscripciones anuales	Área comercial CRM y dependientes
Cliente	Mejorar los tiempos de entrega de propuestas técnico – económicas en un 20%	Tiempo mejorado / tiempo actual	Preventa funcional

	Generar campañas de captación de clientes nuevos al menos una vez al mes	Cantidad de campañas por mes	Marketing CRM
	Realizar al menos dos presentaciones de los módulos de CRM de Dynamics 365 a clientes de la base instalada en un mes	Cantidad de presentaciones por mes	Preventa funcional y KAM
Procesos internos	Definir un proceso de ventas en Chile basado en el proceso de Solution Selling antes de diciembre del 2018	Existencia del proceso antes de mayo	Área comercial CRM y dependientes
	Generar una herramienta de estimación de proyectos que se base en las actividades del proceso de ventas a definir antes de septiembre de 2018	Existencia del proceso antes de septiembre	Preventa funcional
Aprendizaje y crecimiento	Generar planes de capacitación internos que permitan obtener al menos una nueva certificación al finalizar el año 2019	Cantidad de capacitaciones anuales por persona	Área comercial CRM y dependientes

La implementación de este proyecto impacta principalmente en las siguientes perspectivas del BSC:

- Financiera: La clasificación de los clientes permitirá ejecutar una estrategia comunicacional en clientes que entreguen mayor valor a la organización, lo que permitirá generar una mayor cantidad de oportunidades de negocios, aumentando la probabilidad de cierre de negocios nuevos. En AlfaPeople, se gana un negocio de cada 10 oportunidades, por lo que un mayor flujo impacta financieramente

- Cliente: La implementación de RFM permite ir monitoreando los movimientos de los clientes con la organización. Esta perspectiva se ve impactada puesto permitirá identificar tanto a los buenos como a los malos clientes y generar acciones para mantener/mejorar su percepción con la organización.

4.1.3 Modelo de Negocios

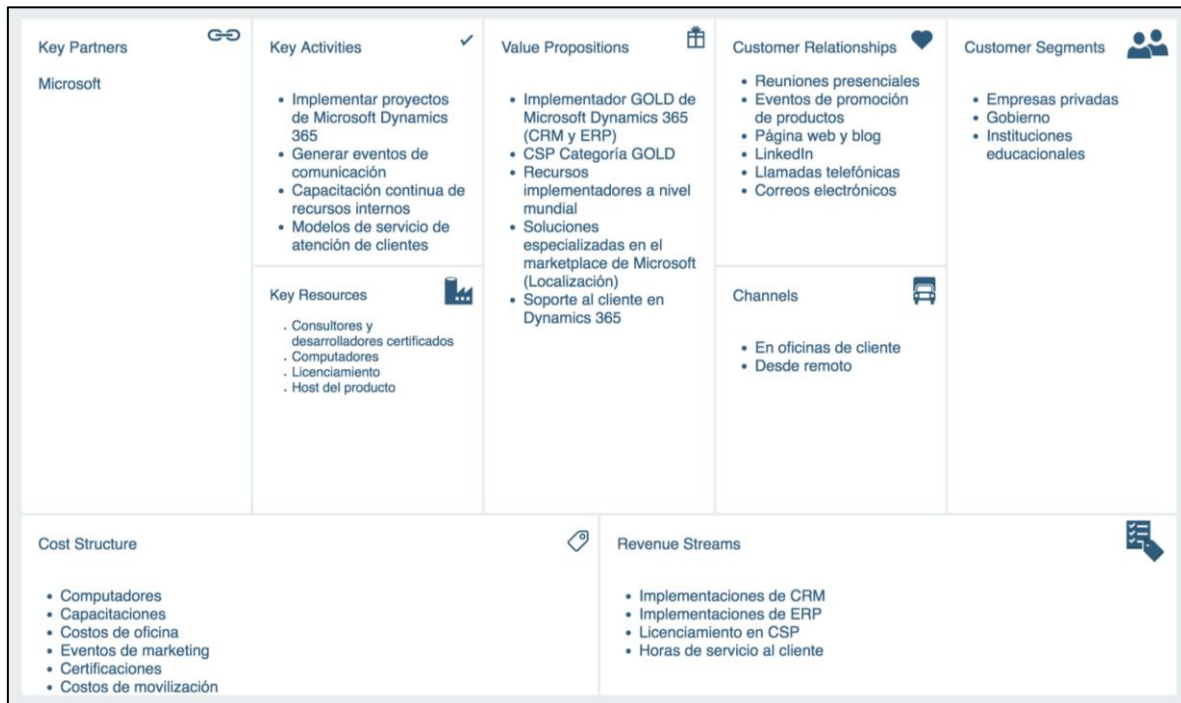


Figura 9. Modelo de negocios

La implementación de este proyecto impacta en los siguientes aspectos del modelo de negocios:

- Segmentos de cliente: El implementar RFM con los datos de compra, permitirá generar nuevos segmentos de clientes basados en el comportamiento con la organización, independiente de la industria a la que pertenezcan
- Relación con los clientes: Permitirá mejorar la relación al momento de generar acciones personalizadas para cada cliente de acuerdo con sus segmentos, buscando moverlos de segmentos dentro del score de RFM
- Actividades clave: Parte de las actividades que se realizan son de comunicación respecto a los productos Microsoft. Las segmentaciones por RFM permiten mejorar las comunicaciones de acuerdo con la categoría del cliente, permitiendo focalizar éstas según el cliente.

4.1.4 Arquitectura de Procesos

Actualmente no existe un detalle de cómo ejecutar las ventas a nivel Chile. Sólo se completan campos en el CRM de la compañía a modo de “obligación” más que seguir con un proceso definido, por tanto, se identifica como parte de la problemática de los bajos cierres el no contar con un proceso de ventas.

4.1.5 Modelamiento Detallado de Procesos (AS – IS)

Las siguientes arquitecturas por describir, utilizarán los patrones de procesos y reflejarán los elementos en los que se considera trabajar.

4.1.5.1 Nivel 0: Macroprocesos

Nivel 0 corresponde al primer nivel de la empresa. Este nivel es la representación de la cadena más genérica de la organización. En la Figura 10. Nivel 0 se reflejan los procesos que reemplazan la “Cadena de valor”, correspondientes a “Proceso de venta de soluciones” y “Proceso de implementación de soluciones”.

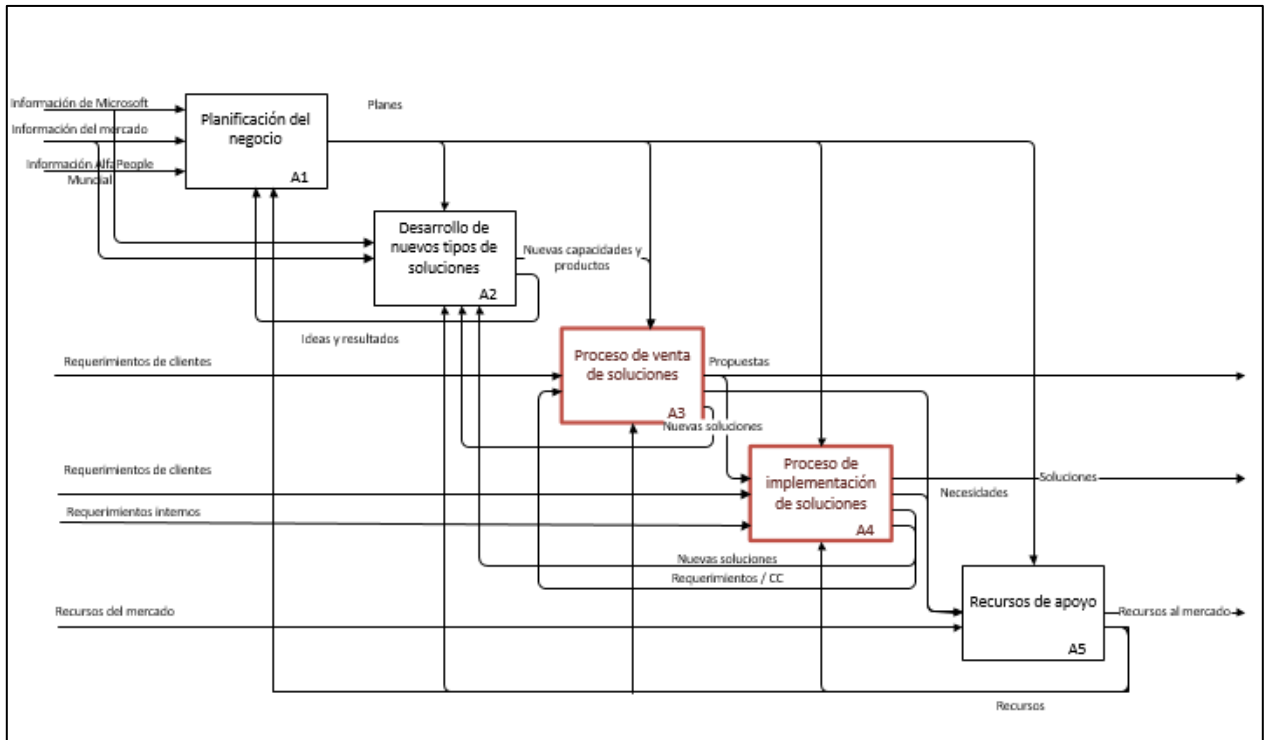


Figura 10. Nivel 0

Se realiza una separación de las cadenas de valor de “Proceso de venta de soluciones” y el “Proceso de implementación de soluciones” debido a que cada una puede optar

con servicios de la otra. Es decir, no sólo se cuenta con un proceso lineal en el que se vende y luego se implemente, sino que consultores en la implementación pueden realizar ventas como parte de su proceso (por medio de controles de cambio) y parte del equipo de preventa participar en apoyo de consultoría.

El problema por abordar se encuentra en la cadena de valor de “Proceso de venta de soluciones”.

4.1.5.2 Nivel 1: Proceso de venta de soluciones (Cadena de valor)

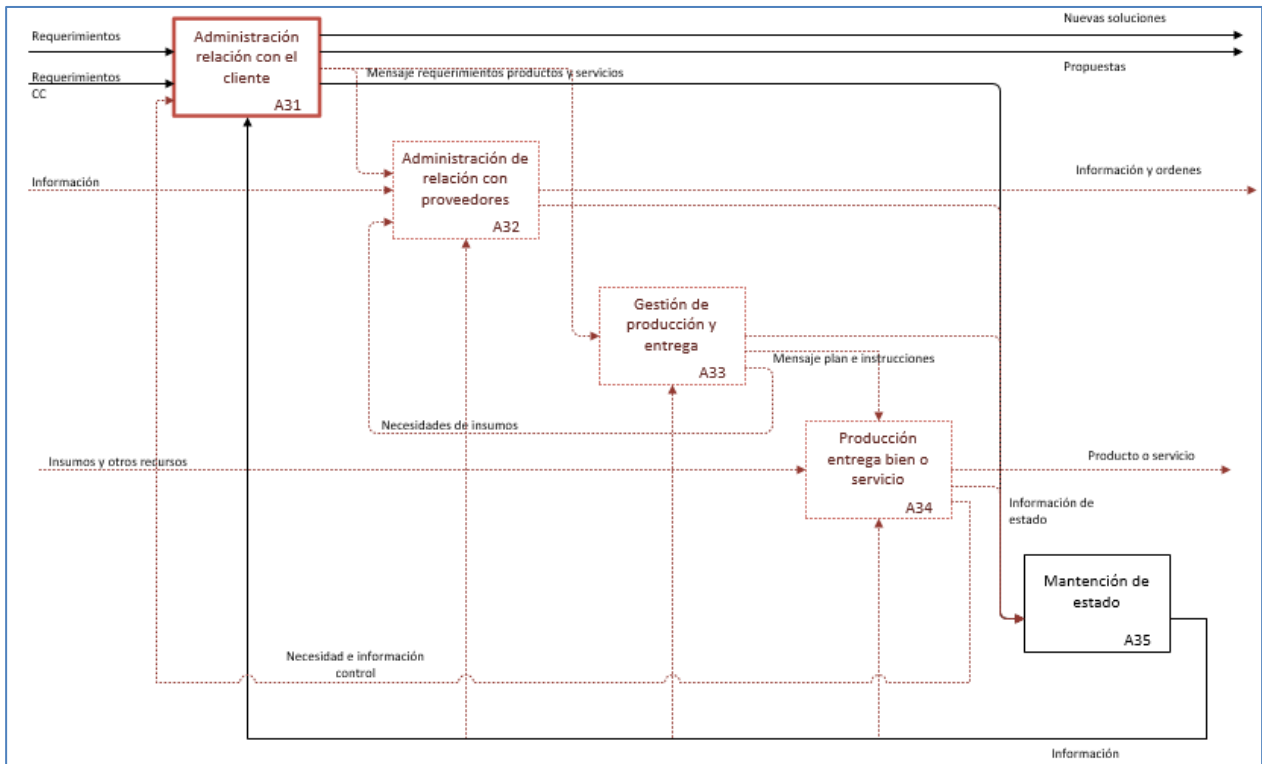


Figura 11. Nivel 1

Del Nivel 1 de AlfaPeople y basado en el patrón de procesos de negocio para la cadena de valor se explica lo siguiente:

- El “Producción entrega bien o servicio” podría corresponder al macro proceso de Nivel 0 llamado “Proceso de implementación de soluciones”. Debe ser retirado de Nivel 1 debido a que cada una de las áreas pueden ser independientes entre sí; Se pueden entregar propuestas que no pasan por una entrega (por ejemplo, vender sólo licenciamiento) y consultoría podría realizar las ventas de controles de cambio sin depender del área de ventas de la organización
- Si se almacena en mantención de estados desde la administración con el cliente. Esto correspondería a la acción que realiza el Account Manager sobre el CRM interno de la organización

- Se mantuvo el nombre de “Administración relación con el cliente” debido a que no existe una instancia de procesos en Chile desde esta parte. Por tanto, la propuesta de este proyecto es contar con un macro proceso que permite este macro proceso.
- Existen relaciones que no son posibles dentro de lo propuesto en el Nivel 1, como por ejemplo la relación con proveedores. Generalmente los elementos necesarios para realizar la implementación son provistos por el cliente o adquiridos directamente a AlfaPeople (Por su característica de ser CSP), evitando la participación de algún tercero.
- Este proceso AS IS solo tiene definido el trabajo de A31 y A35 del diagrama, permitiendo identificar mejoras o implementación de procesos en A32, A33 y A34
- La salida de este proceso puede ser generada desde el macro proceso de “Administración relación con el cliente”, debido que es desde esa área en la que se generan las propuestas y/o las nuevas soluciones.

4.1.5.3 Nivel 2: Administración de relación con el cliente

El punto que se abordará en Nivel 2 corresponde al detalle de la “Administración de relación con el cliente” para continuar con el proceso de ventas.

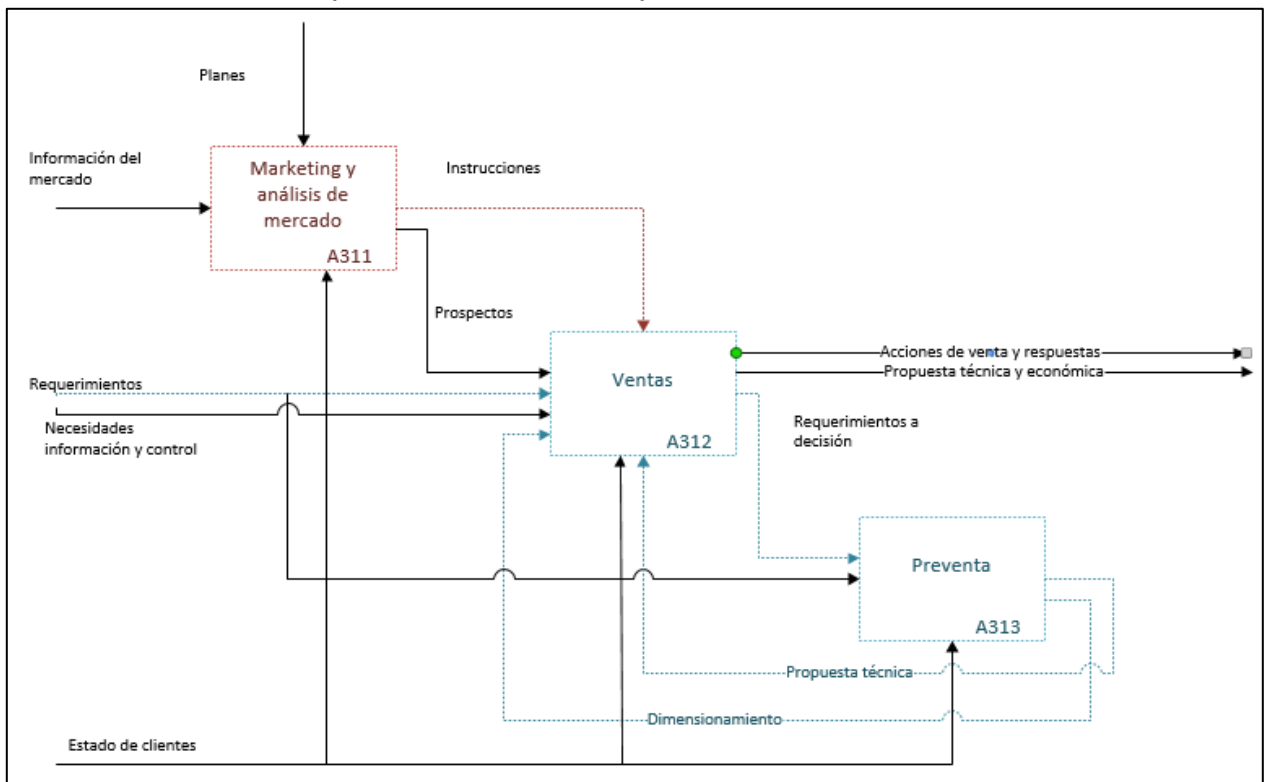


Figura 12. Nivel 2

De este diagrama se puede concluir:

- El problema se encuentra entre los procesos de Ventas y Preventas. No existe una formalidad de trabajo ni tampoco una correcta identificación del cliente para trabajar con él.
- El proceso de Preventa reemplaza lo señalado en el macro proceso de “Decidir satisfacción de requerimientos”. Esto es debido a que preventa es quien propone la solución y el cumplimiento de lo solicitado por distintos clientes
- Marketing no se encuentra dentro del proceso actual. En AlfaPeople es más un recurso compartido que se puede utilizar por distintas áreas de la organización.
- Marketing no entrega instrucciones, entrega prospectos a la organización.
- Ventas no genera planes de cuentas ni tampoco puede planificar el uso de preventa sin claridad de qué ofrecer.

4.2 Diagnóstico de la Situación Actual

El objetivo de AlfaPeople es crecer como compañía y sostenerse en el tiempo en la implementación de los procesos de Microsoft Dynamics 365. El año 2016 se cerró un proyecto que tendrá una implementación de aproximadamente dos años de los módulos de Operaciones y CRM de Dynamics, lo involucró que la compañía tuviera que invertir en nuevos recursos para cumplir con la demanda del proyecto más la implementación de los proyectos existentes.

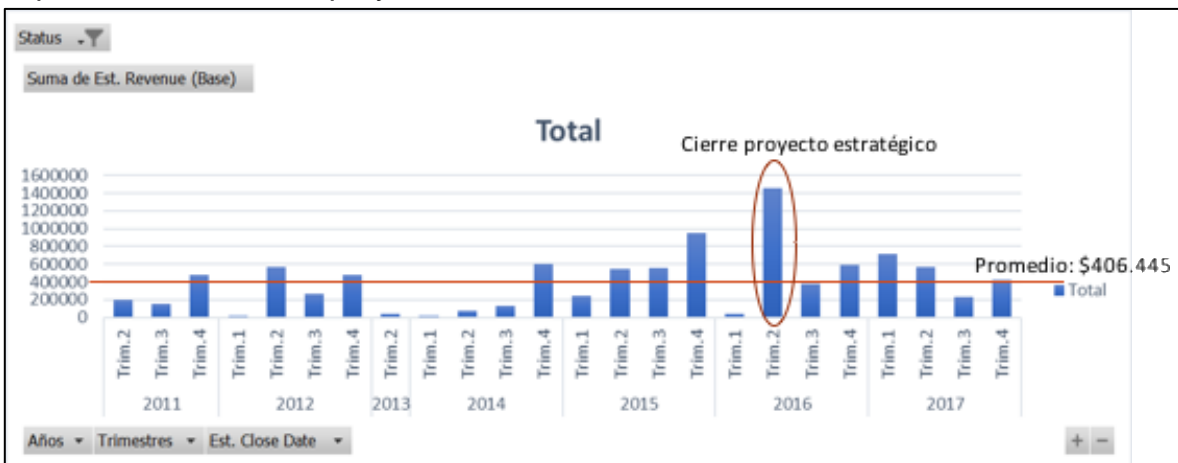


Figura 13. Cierre de proyecto estratégico CRM

La Figura 13. Cierre de proyecto estratégico CRM refleja los proyectos de CRM que se han ganado desde el 2011, el cual ha sido variado por los siguientes motivos:

- Para los años 2011 y 2012, AlfaPeople no contaba con el área de CRM a nivel local, sólo contaban con consultores que conocían el producto y principalmente sus implementaciones eran realizadas por un equipo de Brasil

- El año 2013 e inicios de 2014, AlfaPeople no comercializó el producto, cerrando el área en Chile. Sólo algunos proyectos que eran consecuencia de la implementación de ERP eran los que implementaron en CRM.
- El año 2015 se construyó nuevamente el área de CRM, adquiriendo a un ex partner dedicado en la implementación de proyectos de CRM.
- El año 2016 se ganó el proyecto señalado y se decidió armar el equipo de CRM más experimentado.
- El año 2017 no se logró cumplir la meta de ventas de aproximadamente USD \$1.000.000 en nuevos negocios.

Para identificar el problema, se utilizó la regla de los “5 por qué” para obtener las causas raíz del problema de las bajas ventas. Los argumentos fueron conversados con distintos actores del área comercial, llegando a un resumen como el que se indica en la siguiente tabla:

Tabla 10: Técnica de los 5 por qué

Problema a estudiar	1	2	3	4	5
¿Por qué no se están cerrando los negocios ?	Por falta del seguimiento o del proceso de ventas	¿Por qué no están siguiendo el proceso de ventas?: Porque no se les ha capacitado	¿Por qué no se les ha capacitado?: Porque nadie ha planificado la capacitación	¿Por qué nadie lo ha planificado?: Porque nadie quiere ser responsable de la planificación	¿Por qué nadie quiere planificar?: Porque se debe vender
	Por el cambio de gobierno	¿Por qué afecta?: Por incertidumbre de las empresas	¿Por qué hay incertidumbre?: Porque no se está seguro de realizar inversiones		
	Por malas referencias de los clientes	¿Por qué tenemos malas referencias?: Porque los proyectos no se están implementando bien	¿Por qué no se están implementando bien?: Porque los consultores les falta experiencia	¿Por qué les falta experiencia?: Porque no se han capacitado	¿Por qué no se han capacitado?: Porque siempre están ocupados
	Por dedicar mucho tiempo a clientes sin potencial	¿Por qué se les está dedicando tiempo?: Porque no	¿Por qué no están siendo bien calificados?: Porque no	Ir a proceso de ventas	

		están siendo bien calificados	están haciendo las preguntas que el proceso les indica		
	Por sólo dedicarnos a clientes nuevos	¿Por qué sólo nos dedicamos a nuevos clientes?: Porque nadie realiza seguimiento a los clientes existentes	¿Por qué nadie hace seguimiento?: Porque no existe un plan de tratamiento de clientes existente	¿Por qué no existe plan?: Porque no se le dio foco en la planificación del año	¿Por qué no se le dio foco?: Porque no ven valor del cliente existente

De los 5 puntos identificados, el punto que más llamó la atención fue el último. Se descubrió que la planificación comercial sólo está siendo enfocada en la captación de nuevos clientes para implementaciones desde cero y no se están potenciando nuevos negocios en los clientes existentes.

Al realizar un análisis de los negocios históricos de AlfaPeople Latam, se lograron identificar datos que reflejan la falta de interés en el cliente existente:

- El total de las oportunidades de negocio desde el año 2011 hasta julio del 2018 es de **1111**
- Esas oportunidades se encuentran distribuidas en **467 clientes**

De esos datos, se realizó un análisis de los clientes que cuentan con más de un negocio. Tal como indica la Figura 14. Cantidad de oportunidades por cliente, un **66% de los clientes sólo cuentan con un negocio.**

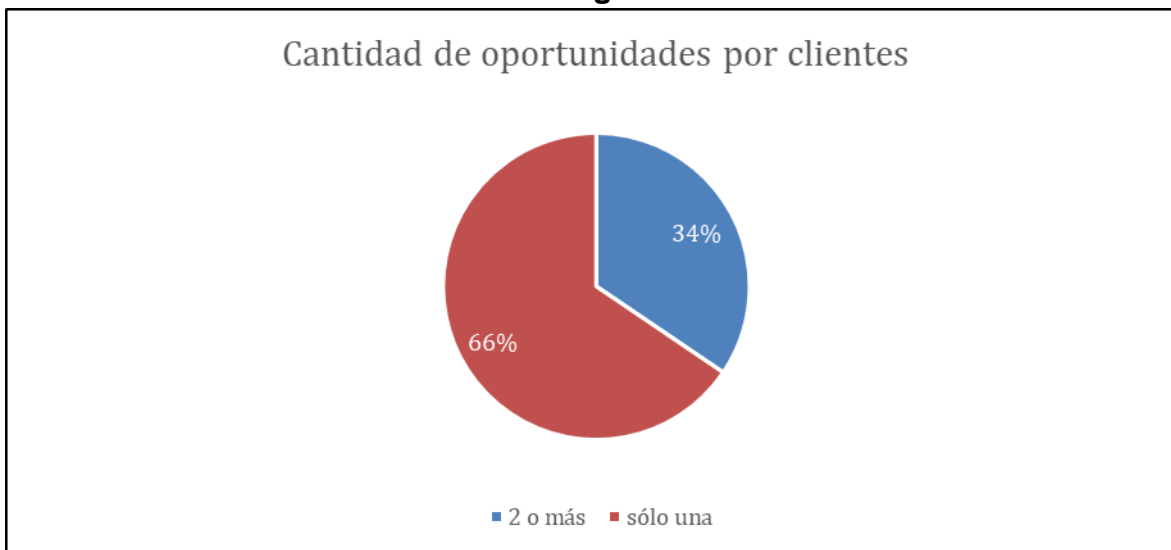


Figura 14. Cantidad de oportunidades por cliente

Al revisar los clientes que sólo cuentan con un negocio, se obtiene:

Tabla 11: Análisis de clientes con un negocio

<i>Clientes sólo un negocio</i>	163
Renovación licencias o servicios	87
Nuevos proyectos	76

Renovación de licencias equivale a que los clientes hicieron modificaciones del proveedor de servicios para cuando quieren agregar nuevos productos a su organización. Nuevos proyectos equivalen a proyectos de consultoría, lo que es el principal interés de la organización. Es decir, el foco de la organización es lo que ha dejado hasta el día de hoy sólo el 16% de los clientes desde el año 2011 a la fecha y no enfocando esfuerzos en mantener a los clientes existentes.

4.3 Dirección de cambio

Tabla 12: Dirección de cambio

Tipo	Situación actual	Propuesta
Coordinación	Actualmente no se utilizan los datos de los clientes para generar campañas focalizadas en base al comportamiento	Coordinación entre el área de Marketing y Comercial para establecer comunicaciones guiadas en las características de los clientes, ofreciendo productos específicos
Prácticas de trabajo	Comunicación de clientes por gerentes de cuentas o marketing basada en selección por decisión estratégica o posibles compras sin definición de una herramienta para almacenar la contactabilidad.	Comunicación basada en el comportamiento de compras con la organización, permitiendo mayor foco en la base instalada, revisando los resultados del RFM y generando las campañas en CRM
Integración de procesos conexos	Marketing no se encuentra incorporado al proceso de ventas para trabajo con base instalada, sólo gestiona nuevos prospectos	Integración en el proceso de venta de base instalada a Marketing, basado en los resultados de RFM
Apoyo computacional	Utilización de CRM sólo para la generación de campañas con los datos	Utilización de algoritmos en Python para generar RFM y luego seleccionar los clientes en CRM para las comunicaciones en base instalada

4.4 BPMN del nuevo proceso (Tercer nivel)

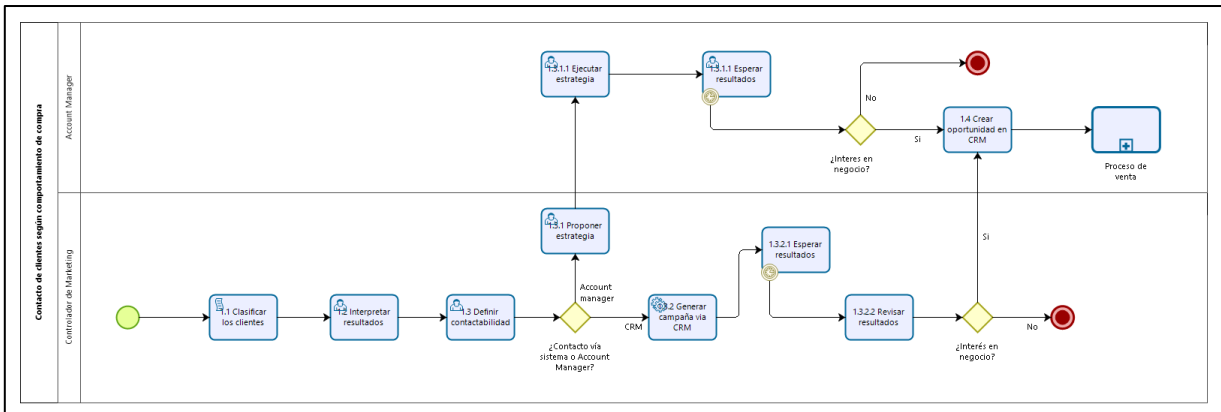


Figura 15. BPMN del nuevo proceso

En este BPMN se describe la estrategia a incorporar en el proceso de venta para trabajar con los clientes existentes y los datos que se poseen registrados. Se diagrama que este proceso parte desde el equipo de Marketing y dependiendo la acción que se desee tomar es si se contactará desde el ejecutivo de cuentas o continuará la campaña desde Marketing.

1.1 Clasificar a los clientes

Rol ejecutante:

Coordinadora de Marketing

Descripción:

Proceso en el que se ejecutará el análisis de RFM que se propone en este proyecto.

Pasos:

- Descargar los datos desde CRM y ubicarlo en una carpeta definida para estos análisis
- Ejecutar Jupyter Notebook con el script configurado

1.2 Interpretar resultados

Rol ejecutante:

Coordinadora de Marketing

Descripción:

Tarea que consiste en interpretar los resultados entregados por el Notebook

Pasos:

- a. Identificar los clientes según su valor

1.3 Definir contactabilidad

Rol ejecutante:

Coordinadora de Marketing

Descripción:

Tarea que consiste en definir cómo se contactará al cliente basado en la estrategia definida en el punto anterior

Pasos:

- a. Definir si existen clientes a contactar por Marketing
- b. Definir si existen clientes a contactar por los gerentes de cuentas.

Basado en quien los atiende, se decide qué realizar en el proceso

En caso de ser gerentes de cuentas:

1.3.1 Proponer estrategia

Rol ejecutante:

Coordinadora de Marketing

Descripción:

Tarea en la que se arma el plan y la estrategia a seguir con el cliente.

Pasos:

- a. Definir la estrategia basada en el valor para la organización.

1.3.1.1 Ejecutar estrategia

Rol ejecutante:

Gerente de cuenta

Descripción:

Tarea en la que el gerente de cuenta ejecuta la estrategia propuesta por Marketing

Pasos:

- a. Contactar al cliente según definición de estrategia

1.3.1.2 Esperar resultados

Rol ejecutante:

Gerente de cuenta

Descripción:

Espera de los resultados de la contactabilidad con el cliente

Pasos:

- a. Revisar si se obtuvieron respuestas
- b. En caso de no obtener, realizar seguimiento según de definición de estrategia

Con estos resultados, se debe evaluar si se continúa con el punto 1.4

1.3.2 Generar campaña vía CRM**Rol ejecutante:**

Coordinadora de Marketing

Descripción:

Actividad de generación de la actividad en CRM directamente con los datos entregados con la segmentación.

Pasos:

- a. Segmentar los clientes en CRM y crear una lista de marketing
- b. Generar campaña de contactabilidad en CRM y asociar la lista de marketing
- c. Generar actividades asociadas a la lista de marketing
- d. Ejecutar campaña en CRM

1.3.2.1 Esperar resultados**Rol ejecutante:**

Coordinadora de Marketing

Descripción:

Actividad de espera de resultados de la campaña y revisión de lo registrado en CRM. Realizar seguimiento según lo definido en estrategia

Pasos:

- a. Revisar las respuestas de campaña en CRM
- b. Al pasar el tiempo, realizar seguimiento según lo definido en estrategia

1.3.2.2 Revisar resultados

Rol ejecutante:

Coordinadora de Marketing

Descripción:

Actividad en la que se analizarán los resultados de la campaña según estrategia definida

Pasos:

- a. Realizar análisis de las respuestas de campaña y definir si poseen posibles oportunidades o no.

En caso de si ser potenciales nuevos negocios, se debe continuar con el punto 1.4, notificando al gerente de cuentas para que cree la nueva oportunidad de negocio

1.4 Crear oportunidad en CRM

Rol ejecutante:

Gerente de cuentas

Descripción:

Actividad de creación de una nueva oportunidad de negocio en el sistema

Pasos:

- a. En CRM, crear una nueva oportunidad con los clientes que poseen interés en un nuevo negocio

Al finalizar esta actividad, se continúa con el proceso de ventas que posee la organización.

CAPITULO 5. PROPUESTA DE APOYO TECNOLÓGICO

5.1 Especificación de Requerimientos

Los requerimientos por señalar sólo apoyarán parte de la solución que se debe realizar. El problema no se resuelve con agregar una nueva herramienta tecnológica, siendo uno de los últimos puntos que se deberán abordar en esta propuesta.

5.1.1 *Requerimientos Funcionales*

- REQ 001: Se requiere que el sistema permita utilizar los datos de los clientes, oportunidades y cierres de negocio
- REQ 002: Se requiere que el resultado de los segmentos de RFM quede visible en alguna de las visualizaciones corporativas
- REQ 003: Se requiere que se pueda ejecutar constantemente la clasificación de los clientes con cambios en los datos según las fechas

5.1.2 *Requerimientos No Funcionales*

- REQ 001: Se debe capacitar a las personas en el uso de las nuevas funcionalidades del sistema
- REQ 002: Se espera que mantenga el mismo uptime que el CRM actual de un 99%

5.2 Arquitectura Tecnológica

Este proyecto no tendrá un impacto directamente en la plataforma tecnológica para las comunicaciones de CRM, más bien entregará un primer canal de entrada para las segmentaciones de los clientes y la ejecución de las campañas.

A futuro, se podrán realizar mejoras en la implementación, tales como llevar la codificación de Python a Power BI y mantener los datos actualizados directamente desde CRM.

El objetivo será:

- Generar segmentaciones de clientes basados en la información que se encuentra a un periodo de tiempo en el sistema
- Almacenar datos suficientes para realizar análisis de datos
- Definir un modelo de segmentación de los datos

La realización del proyecto se basa en las siguientes actividades:

5.2.1 Descarga de datos

Contiene la consulta en CRM con la generación del archivo en Excel.

El dataset generado para este proyecto, contiene información de las siguientes entidades de CRM

- Cuentas: Información de los clientes empresa
- Contactos: Información de los clientes contacto (persona natural)
- Oportunidades: Información de los cierres de negocio
- Industrias: Información de la industria a la que pertenece la empresa
- Listas de precios: Lista de precio de los productos que se negociaron dentro de la oportunidad
- Productos: Mantenedor de productos
- Productos de oportunidad: Relación entre una oportunidad de negocios y sus productos
- Divisa: Mantenedor de divisa de una oportunidad
- Unidad organizacional: División de AlfaPeople con la que se está trabajando.

Esta fuente de datos tiene lineamientos del grupo. No se pueden trabajar con todos los datos debido a los cumplimientos gubernamentales que tienen las distintas divisiones de AlfaPeople a nivel global, por lo que se debe cumplir con las siguientes reglas:

- 1) Sólo debe tener datos hasta primer Q del 2020
- 2) Sólo debe considerar la región LATAM, excluyendo Brasil e incorporando México.

El modelo de entidad - relación simplificado del set de datos se resume en el siguiente diagrama:

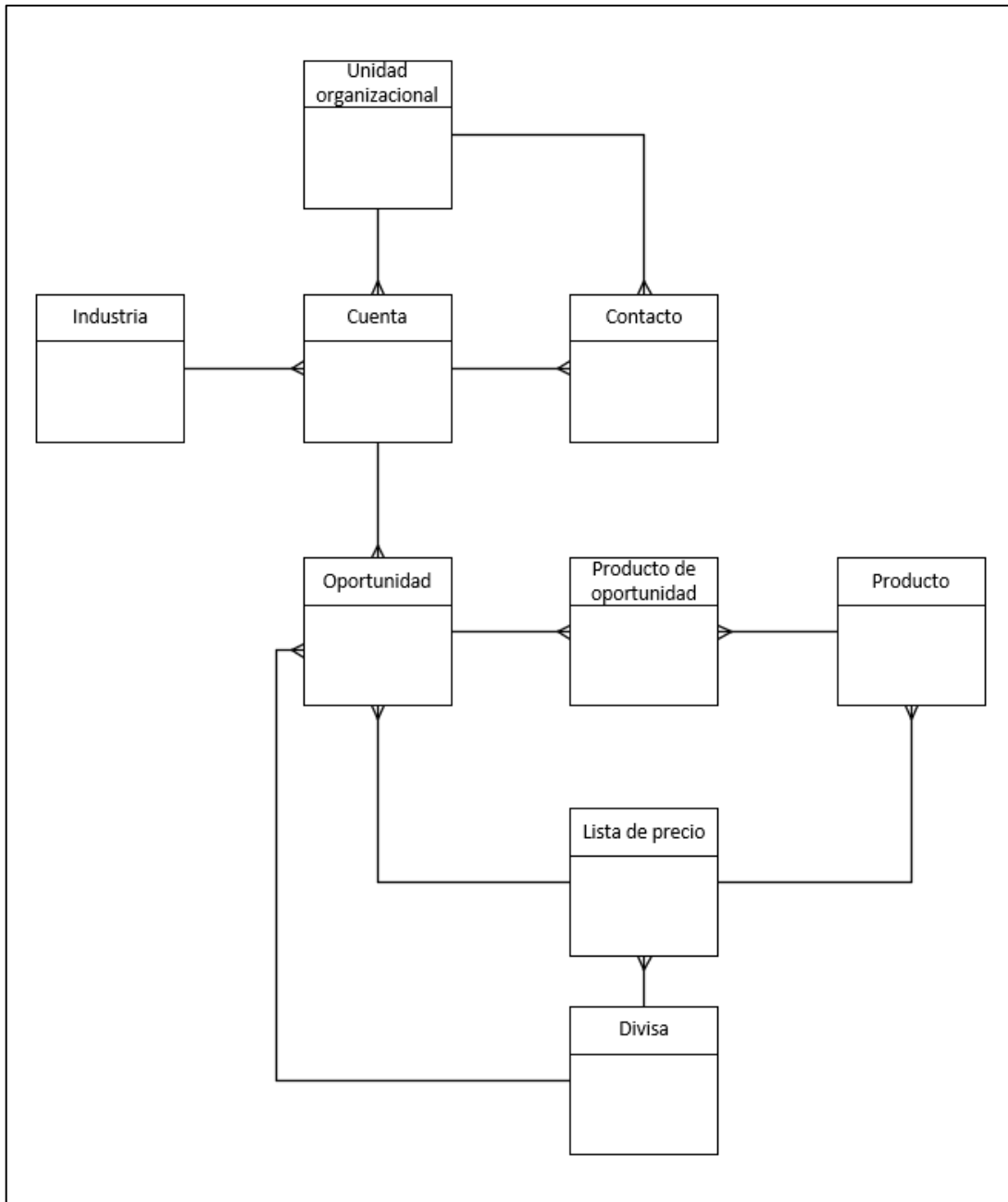


Figura 16. Modelo de entidad relación del archivo en Excel

El archivo se ubicará en una dirección a definir para que se pueda ejecutar el código desarrollado en este trabajo.

5.2.2 Aplicación de RFM

Lo primero a realizar es la lectura de los datos extraídos desde CRM.

```

#Cargar los datos. .
opp_df = pd.read_csv("DF2020.csv", low_memory=False,error_bad_lines=False, decimal = ',', sep = ';')
  
```

Figura 17. Lectura de datos en Python

Con esto se realiza un análisis exploratorio de los datos para ver el contenido. RFM se aplica solo con negocios cerrados con la organización, por lo que se debe evaluar cuanto del contenido importado será utilizado.

```
In [212]: opp_df['opportunitystatecode'].value_counts(dropna = False)
Out[212]: 2    996
          1    762
          Name: opportunitystatecode, dtype: int64
```

Figura 18. Análisis de datos de tipos de oportunidades

El código en las oportunidades indica el estado en el que se encuentra.

- 1 es ganada (oportunidades que serán utilizadas)
- 2 es perdida. Oportunidades que no se concretaron con la organización.

Al aplicar cualquiera de los puntos de RFM, se deberá realizar este filtro en los datos y así mantener la homologación de la información.

5.2.2.1 Recency

Para calcular el cliente que compró hace menos tiempo, se identificará cuál fue la última fecha de cierre de cualquier oportunidad en el set de datos. Se calculará la cantidad de días que pasaron entre esa fecha y la fecha de cierre del registro.

Para ello, primero se genera un segmento de datos:

```
In [296]: df_r = pd.DataFrame()
          df_r[['opportunityid.customerid_account.name', 'opportunityid.actualclosedate']] =
          df[['opportunityid.customerid_account.name', 'opportunityid.actualclosedate']]
```

Figura 19. Cálculo de R en Python

Luego se obtiene la última fecha de cierre:

```
In [297]: #Obtenemos la última fecha de cierre para calcular el recency
In [298]: df_r['MasReciente'] = df_r['opportunityid.actualclosedate'].max()
In [299]: df_r.head()
Out[299]:
```

	opportunityid.customerid_account.name	opportunityid.actualclosedate	MasReciente
0	Inmobiliaria LOGA	2018-10-17 18:17:00	2020-06-01 18:56:00
1	Inmobiliaria LOGA	2018-10-17 18:17:00	2020-06-01 18:56:00
2	Inmobiliaria LOGA	2018-10-17 18:17:00	2020-06-01 18:56:00
3	CALZADO CHICOTE	2018-12-13 16:59:00	2020-06-01 18:56:00
4	CALZADO CHICOTE	2018-12-13 16:59:00	2020-06-01 18:56:00

Figura 20. Obtención de última fecha de cierre del set de datos

Se puede identificar que la fecha más reciente de cierre de registros es el 01-06-2020 en los datos analizados.

Posteriormente se aplica la fórmula para calcular Recency (R)

```
In [300]: #Recency = última opp cerrada - fecha de cierre de la opp
In [301]: df_r['R'] = df_r['MasReciente'] - df_r['opportunityid.actualclosedate']
df_r['R'] = df_r['R']/np.timedelta64(1, 'D')
```

Figura 21. Diferencia entre último cierre y fecha de cierre del registro

Aplicada la formula, el data frame posee los días desde el cierre del registro:

```
In [308]: df_r.head()
Out[308]:
```

	opportunityid.customerid_account.name	R
7	472	303.352778
9	ABCDIN	973.116667
10	ACB Ingenieria	665.093750
24	ACHS	303.095833
30	ADVISICON	1273.901389

Figura 22. Días de cierre de registros

La distribución de los datos queda:

```
In [412]: df_r.R.describe()
Out[412]: count      646.000000
          mean       673.922853
          std        317.626677
          min         0.000000
          25%        430.863194
          50%        650.595833
          75%        904.374653
          max        1603.538889
          Name: R, dtype: float64
```

Figura 23. Análisis de distribución de los datos

Con esta descripción, se puede determinar que el promedio de cierre de oportunidades es 673 días, siendo explicado principalmente en que las oportunidades que involucran costos en productos de ERP toman más tiempo por los presupuestos elevados. Para calcular el score de R, se aplicará KMeans a los datos. Lo primero a realizar es aplicar el método del codo, para poder determinar el número óptimo de clústeres

```
In [364]: from sklearn.cluster import KMeans

sse={}
tx_recency = dfrfm[['Recency']]
for k in range(1, 10):
    kmeans = KMeans(n_clusters=k, max_iter=1000).fit(tx_recency)
    tx_recency["clusters"] = kmeans.labels_
    sse[k] = kmeans.inertia_
plt.figure()
plt.plot(list(sse.keys()), list(sse.values()))
plt.xlabel("Number of cluster")
plt.show()

/Users/alfonsocasanova/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py:7: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html#indexing-view-versus-copy
import sys
```

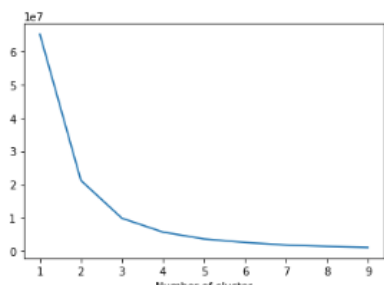


Figura 24. Aplicación del método del codo para R

De acuerdo al gráfico, es posible visualizar que el número óptimo es k=4 puesto que es donde se genera el mayor quiebre y se aplana la curva.

Con estos datos, se aplica KMeans con k= 4 y se visualizan los siguientes datos:

Out[367]:

	count	mean	std	min	25%	50%	75%	max
RecencyCluster								
0	81.0	1230.406147	139.750253	1058.156944	1104.538889	1197.179861	1329.087500	1603.538889
1	171.0	889.369274	88.725406	745.162500	809.860764	896.722917	955.918403	1055.960417
2	226.0	598.361388	82.491796	450.068750	536.358681	593.028819	663.132639	740.683333
3	168.0	287.972892	88.366594	0.000000	219.838368	300.930556	360.454514	441.187500

Figura 25. División de clusters para R

En este primer análisis de R, se puede distinguir que los clústeres están ordenados de menor a mayor, identificando que aquellos que pertenecen al clúster 0 son los que más tiempo demoran en cerrar un negocio, mientras que los que están en el clúster 3 son los que menos demoran, siendo más atractivos.

5.2.2.2 Frecuency

El cálculo de frecuency (frecuencia o F) indica la cantidad de oportunidades ganadas por cada cliente.

Para aplicarlo, se genera un nuevo dataframe en el que sólo se obtendrán los datos necesarios del dataframe original, correspondientes al nombre del cliente, la cantidad invertida en la oportunidad y el estado de la oportunidad. Posteriormente se mantienen sólo las oportunidades cuyo estado sea “ganado”.

```
In [56]: #ganadas
df_g = pd.DataFrame()
df_g[['opportunityid.customerid_account.name', 'baseamount_base', 'estadoOp_2']] =
df[['opportunityid.customerid_account.name', 'baseamount_base', 'estadoOp_2']]

In [57]: df_g = df_g.drop(df_g[df_g.estadoOp_2 == 1].index)
```

Figura 26. Generación de nuevo Data frame para cálculo de F

Con sólo las oportunidades ganadas, para obtener “F” se contará la cantidad de productos comprados en cada oportunidad y se agruparán por los clientes. (No es necesaria la cantidad en este punto y será utilizada en Monetary).

```
In [60]: df_g = df_g.groupby(['opportunityid.customerid_account.name']).count().reset_index()

In [61]: df_g.head()

Out[61]:
```

	opportunityid.customerid_account.name	estadoOp_2
0	ACHS	1
1	ADVISICON	1
2	AGROSAVIA	14
3	AIEP	8
4	AKVA group Chile	1

Figura 27. Agrupación de cantidad de oportunidades ganadas para cálculo de F

Para dar más orden, se cambiará la cuenta por “F” en el dataframe.

```
In [62]: df_g.rename(columns={'estadoOp_2': 'F'}, inplace=True)
In [63]: df_g.head()
Out[63]:
```

	opportunityid.customerid_account.name	F
0	ACHS	1
1	ADVISICON	1
2	AGROSAVIA	14
3	AIEP	8
4	AKVA group Chile	1

Figura 28. Cambio de nombre de columna a F

Al analizar los datos, se obtiene que el promedio de compra de los clientes de AlfaPeople es 3 productos. Esta explicación es debido a que cada proyecto requiere más de un producto al momento de iniciar. Por ejemplo, para un proyecto de ERP a implementar en Chile requiere a lo menos el licenciamiento del producto, el licenciamiento de la localización en Chile y el servicio de implementación. Por el lado de CRM, se requiere licenciamiento a nivel usuario, licenciamiento a nivel administrador e implementador del producto.

```
In [129]: df_g.F.describe()
Out[129]:
```

count	243.000000
mean	3.135802
std	3.119201
min	1.000000
25%	1.000000
50%	2.000000
75%	4.000000
max	26.000000
Name: F, dtype: float64	

Figura 29. Análisis de datos de F

Para calcular el score respectivo, se aplica igual lógica que lo implementado en Frecuency, con K= 4, entregando los siguientes datos al momento de describir la información:

```
In [109]: dfrfm = order_cluster('FrequencyCluster', 'Frecuency',dfrfm,True)
#see details of each cluster
dfrfm.groupby('FrequencyCluster')['Frecuency'].describe()
```

```
Out[109]:
```

	count	mean	std	min	25%	50%	75%	max
FrequencyCluster								
0	492.0	0.180894	0.385322	0.0	0.0	0.0	0.0	1.0
1	99.0	2.616162	0.710089	2.0	2.0	2.0	3.0	4.0
2	50.0	6.540000	1.875840	5.0	5.0	6.0	7.0	11.0
3	5.0	17.400000	5.079370	14.0	14.0	15.0	18.0	26.0

Figura 30. Aplicación del método del codo para F

Es posible identificar que nuevamente están ordenados de peor a mejor cliente según sus clústeres:

- Cluster 0: Posee sólo clientes que compran hasta un producto.
- Cluster 1: Posee clientes que han comprado entre 2 a 4 productos.
- Cluster 2: Clientes que han comprado entre 5 a 11 productos
- Cluster 3: Clientes que han comprado entre 14 a 26 productos.

Los que pertenecen al Cluster 0, suelen ser los que compran licencias y/o se cambiaron a AlfaPeople para la compra del licenciamiento, mientras que los que pertenecen al Cluster 2 y 3 corresponden principalmente a clientes que han implementado un proyecto con AlfaPeople y se han mantenido en el transcurso del tiempo renovando licenciamiento y soporte

5.2.2.3 Monetary (Revenue)

Este punto analiza la cantidad de dinero generada por cada cliente con AlfaPeople. Para ello, se calcula la suma de los ingresos de todos los productos de las oportunidades ganadas.

Para obtener “M”, se suman todos los ingresos de los productos por cada cliente. A nivel de código, se utiliza el mismo código que el utilizado en Frecuency (crear nuevo dataframe con los clientes y los montos), posteriormente se suman todos los ingresos y se agrupa por clientes

```

In [66]: df_m[['opportunityid.customerid_account.name', 'baseamount_base', 'estadoOp_2']] = df[['opportunityid.customerid
In [67]: df_m = df_m.drop(df_m[df_m.estadoOp_2 == 1].index)
In [68]: df_m = df_m.drop(['estadoOp_2'], axis= 1)
In [69]: df_m = df_m.groupby(['opportunityid.customerid_account.name']).sum().reset_index()
In [70]: df_m.head()
Out [70]:

```

	opportunityid.customerid_account.name	baseamount_base
0	ACHS	1146738
1	ADVISICON	917647
2	AGROSAVIA	52724254
3	AIEP	35403784
4	AKVA group Chile	353402

Figura 31. Suma de ingresos para generar M

Con estos datos, se puede aplicar KMeans con K=4 y analizar los resultados

```

Out [111]:

```

	count	mean	std	min	25%	50%	75%	max
MonetaryCluster								
0	586.0	5.866128e+05	1.448971e+06	0.0	0.000000e+00	0.0	225245.0	8698738.0
1	39.0	1.790935e+07	8.036407e+06	9516978.0	1.172503e+07	13653232.0	20859601.5	35586648.0
2	18.0	5.528263e+07	1.426896e+07	38553076.0	4.321666e+07	52428235.5	66903435.0	79260314.0
3	3.0	1.473939e+08	3.513652e+07	114606749.0	1.288490e+08	143091349.0	163787439.5	184483530.0

Figura 32. Análisis de los datos de M

Se puede apreciar que los clústeres se encuentran nuevamente ordenados, siendo Cluster 0 el que posee a los clientes que generan menos ingresos, mientras que los del Cluster 3 son los clientes que generan más ingresos.

5.2.2.4 Cálculo de puntaje de los clientes

Obtenido todos los clústeres de RFM, permite realizar el cálculo del puntaje del cliente. En esta implementación, se sumarán los valores de los clústeres de RFM y se obtendrá un promedio, lo que permitirá clasificar a cada cliente:


```

In [112]: #calculate overall score and use mean() to see details
dfrfm['OverallScore'] = dfrfm['RecencyCluster'] + dfrfm['FrequencyCluster'] + dfrfm['MonetaryCluster']
dfrfm.groupby('OverallScore')['Recency', 'Frequency', 'Monetary'].mean()

<ipython-input-112-951afdf92798>:3: FutureWarning:
Indexing with multiple keys (implicitly converted to a tuple of keys) will be deprecated, use a list instead.

Out[112]:

```

	Recency	Frequency	Monetary
OverallScore			
0	1240.396890	0.275362	4.550359e+05
1	916.347250	0.223684	2.317782e+05
2	630.186085	0.367021	2.357764e+05
3	393.100005	0.986755	1.204604e+06
4	438.417803	3.484848	6.218068e+06
5	444.328889	5.550000	1.709705e+07
6	333.938325	6.875000	3.425551e+07
7	245.499038	8.923077	5.242664e+07
8	190.011285	9.750000	1.024670e+08

Figura 33. Cálculo de los puntajes de RFM

Esta clasificación sigue igual lógica que los resultados obtenidos en cada clúster. Los que poseen puntaje 0 son los peores clientes, mientras que quienes poseen puntaje 8 están categorizados como los mejores clientes.

Para generar una mejor interpretación, se aplicará una lógica basada en el puntaje que categorizará a cada cliente según el valor aportado. Esta lógica está basada en los datos de la imagen anterior.

- Puntaje 0 a 2: Low Value (Clientes cuya última compra es más de un año y medio, con baja frecuencia e ingresos en menos de 1M de Euros)
- Puntaje 3 a 4: Mid Value (Clientes que al menos han comprado entre año y medio y un año, hasta 3 productos e ingresos entre 1 y 6,2M de Euros)
- Puntaje 5 o más: High Value (Clientes que han comprado hace menos de un año y medio, al menos 5 productos e ingresos sobre 6,2M de Euros)

Es posible revisar el comportamiento a nivel de gráficos, en los cuales se puede diferenciar claramente el comportamiento de cada uno de los clústeres.

- Ingresos vs Frecuencia:

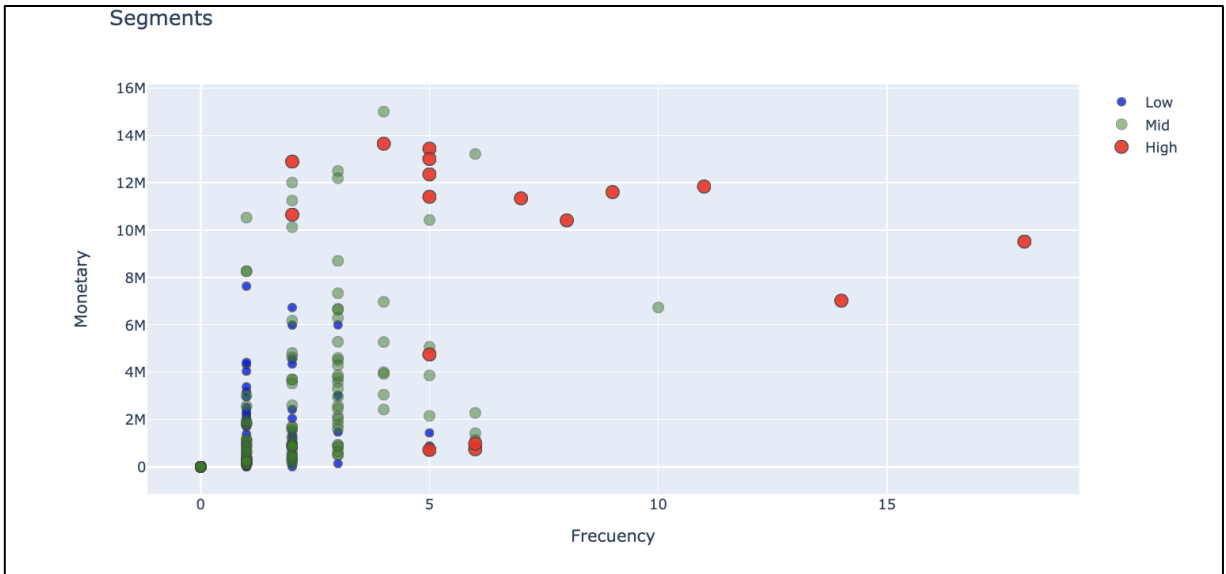


Figura 34. Análisis de comportamiento de Ingresos VS Frecuencia

- Ingresos vs Reciente

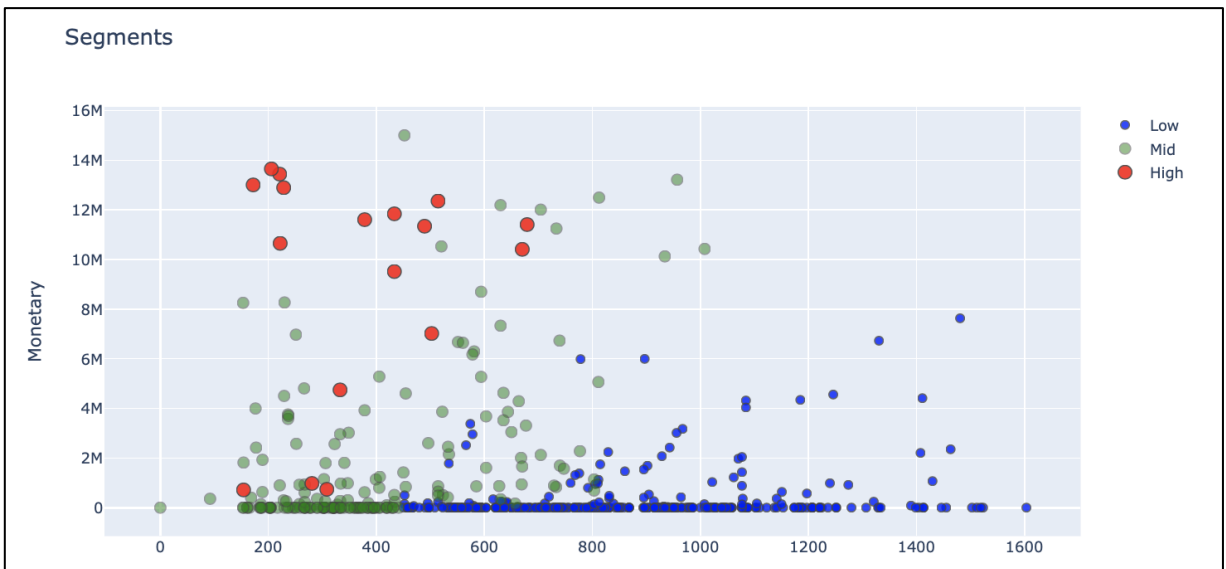


Figura 35. Análisis de comportamiento de Ingresos VS Reciente

- Frecuencia vs Reciente

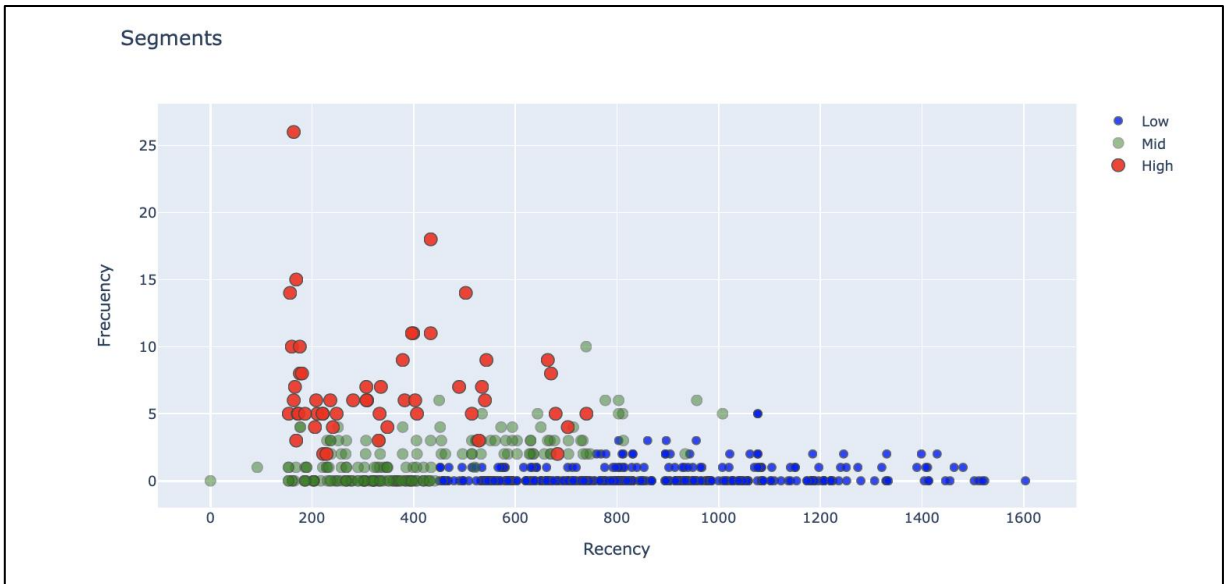


Figura 36. Análisis de comportamiento de Frecuencia VS Reciente

Los datos entregados en el análisis de RFM indican que se poseen 414 clientes categorizados como “Low value”, 180 como “Mid value” y 52 como “High value”

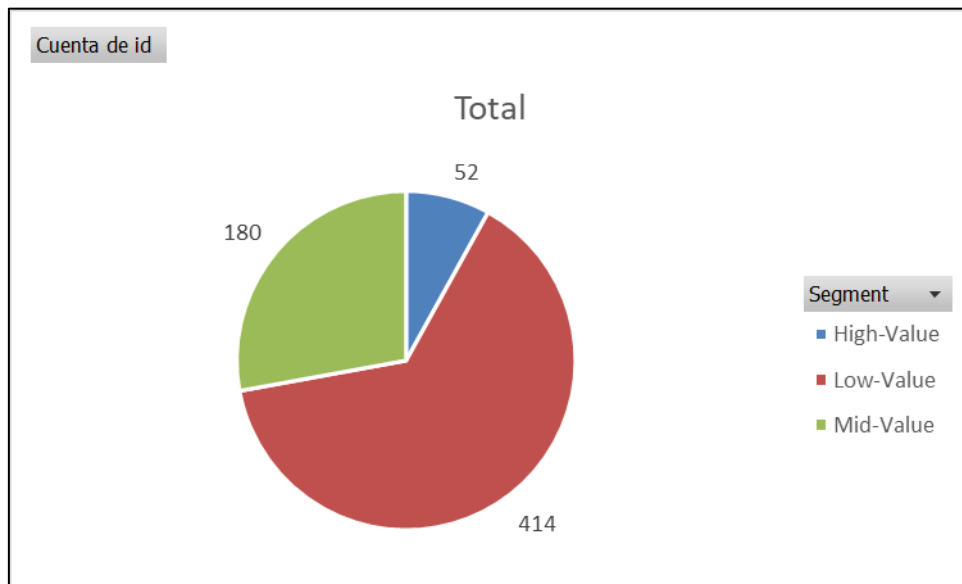


Figura 37. Gráfico de tipo de segmento

Estos datos permiten dejar planteada la siguiente parte de la implementación de este proyecto, generando acciones que se podrían aplicar y evaluar sus resultados:

- Clientes “Low Value”: Estrategia de aumentar la Frecuencia
- Clientes “Mid Value”: Estrategia enfocada en retención y aumento de frecuencia de compra

- Clientes "High Value": Estrategias de retención

CAPITULO 6. EVALUACIÓN DEL PROYECTO

6.1 Flujo de Caja

En esta sección se detalla la evaluación económica a realizar en el proyecto. Este proyecto tendrá un horizonte de 2 años debido a los continuos cambios que sufre la plataforma en el tiempo.

Para hacer el flujo de caja, se tomaron las siguientes consideraciones:

- El software será alojado en Azure, utilizando el plan S3 del producto App Services. El costo anual será de USD \$396
- Proyecto puro, sin financiamiento.
- Por plan institucional, cada cuenta de usuario cuenta con USD \$200 en productos de Azure, los cuales descuentan directamente en el plan S3.
- Moneda del flujo: USD
- Proyección: 2 años de funcionamiento, debido que el producto del que se obtendrán los datos cambia constantemente y cuenta con aplicaciones que probablemente se integren sin este desarrollo.
- La meta de venta de este año es de USD \$3,5 M. El proyecto, al buscar obtener más ingresos con la categorización de clientes, podría aumentar hasta un 35% las ventas de productos. El supuesto será que las ventas son las mismas al año 2017 (\$2.216.221).
- Los gastos de desarrollo son costo hundido debido a que sólo corresponde al sueldo del estudiante del MBE que desarrollará el proyecto.
- Si se cumple con la meta de ventas, es probable que se necesite contratar a más personas para cumplir con los proyectos (Al menos en el área de CRM). Se asume que son al menos 3 personas más en el año 1 y 2 en el año 2. Los sueldos que tendrían los recursos son:

Tabla 13: Análisis de sueldos

Contratos	Mes	Año	Dólar
Consultor	\$1.300.000	\$15.600.000	\$24.762
Jefe de proyecto	\$2.000.000	\$24.000.000	\$38.095
Desarrollador	\$1.000.000	\$12.000.000	\$19.048

- Según SII ⁵, el Software no se deprecia.

⁵ http://www.sii.cl/preguntas_frecuentes/renta/001_002_0985.htm

- De acuerdo con Forbes ⁶, Microsoft tendrá un crecimiento de un 13% en las soluciones de Dynamics 365. Como supuesto de ventas, para el año 2, aumentará lo proyectado en el año 1 en un 10%
- No se involucrarán costos en desarrollo sobre CRM porque no está permitido por órdenes de Dinamarca (El CRM ya existe y está implementado)
- La tasa de descuento es 12,28 y fue determinado con los datos de la

Tabla 14: Tasa de descuento

Tasa de descuento (r)	12,28%	Detalle
Tasa libre de riesgo (rLB)	5,18%	Según bonos Banco Central últimos 10 años
Rentabilidad del mercado	14,40%	Según IPSA últimos 25 años (anual nominal)
Riesgo sistemático	0,77	Riesgo de industria obtenido por el comportamiento bursátil de la empresa SONDA

El flujo de caja generado bajo las consideraciones señaladas corresponde al siguiente:

⁶<https://www.forbes.com/sites/greatspeculations/2017/10/27/microsofts-cloud-services-growth-continues/#4a6b315712f5>

Tabla 15: Flujo de caja

	Año 0	Año 1	Año 2
Ingresos por ventas		\$775.677	\$853.245
(Costos de ventas)			
(Gastos de administración)		-\$82.841	-\$44.746
<i>Consultor</i>		-\$24.762	-\$24.762
<i>Desarrollador</i>		-\$19.048	-\$19.048
<i>Jefe de proyecto</i>		-\$38.095	
<i>Host Azure</i>		-\$936	-\$936
(Depreciación)			
Resultado operacional (A)	\$0	\$692.836	\$808.499
Ganancia/Perdida capital (+/-)			
Ingresos financieros			
(Interés)			
(Perdidas ejercicio anterior)			
Resultado no operacional (B)	\$0	\$0	\$0
Utilidad antes de impuesto (A+B)	\$0	\$692.836	\$808.499
(Impuesto a las empresas) (27%)	\$0	-\$187.066	-\$218.295
Utilidad después de impuesto	\$0	\$505.771	\$590.205
Depreciación			
Perdidas ejercicio anterior			
Ganancia/Perdida capital (+/-)			
Flujo operacional ©	\$0	\$505.771	\$590.205
(Inversión fija)	-\$936		
Valor residual de los activos			\$0
(Capital de trabajo)			
+ Recuperación del capital de trabajo			
+ Préstamos			
(Amortizaciones)			
Flujo de capitales (D)	-\$936	\$0	\$0
Flujo de caja privado (C+D)	-\$936	\$505.771	\$590.205

Con este flujo, se generan los siguientes indicadores del proyecto:

Tabla 16: Indicadores de proyecto

VAN	\$3.049,38
------------	-------------------

TIR	54052%
-----	--------

6.2 Análisis de Sensibilidad

Debido a que es un proyecto con incertidumbre en los ingresos, es que se decide utilizar el método de análisis de sensibilidad para comprobar el flujo de caja variando los siguientes atributos:

- Ingresos por ventas: Se consideran distintos escenarios debido a que es un dato incierto y sólo se generan proyecciones de éstas. Los montos son:
 - o 10% de ventas: Obtenido del crecimiento de la empresa AlfaPeople en los últimos años
 - o 35% de ventas: De acuerdo a las proyecciones generadas por los mercados de CRM
 - o 45% de ventas: Crecimiento de mercado y aumento de ventas de AlfaPeople
 - o 58% de ventas: Lo necesario para cubrir todo el aumento de venta necesario para cumplir con cuota de ventas
- Costos: Principalmente si se utilizarán los USD \$200 dólares que tiene cada uno de los usuarios para el pago de suscripción del servicio de Azure. Con esto se podría ejecutar Azure Notebooks.

Los datos entregados, generan la siguiente tabla:

Tabla 17: Análisis de sensibilidad

Escenario	VAN	TIR
10% ventas Host completo	\$375	3022%
35% ventas Host completo	\$2.858	15124%
35% ventas 200USD Host	\$3.049	54052%
45% ventas 200USD Host	\$4.043	71335%
58% ventas 200USD Host	\$5.334	93804%

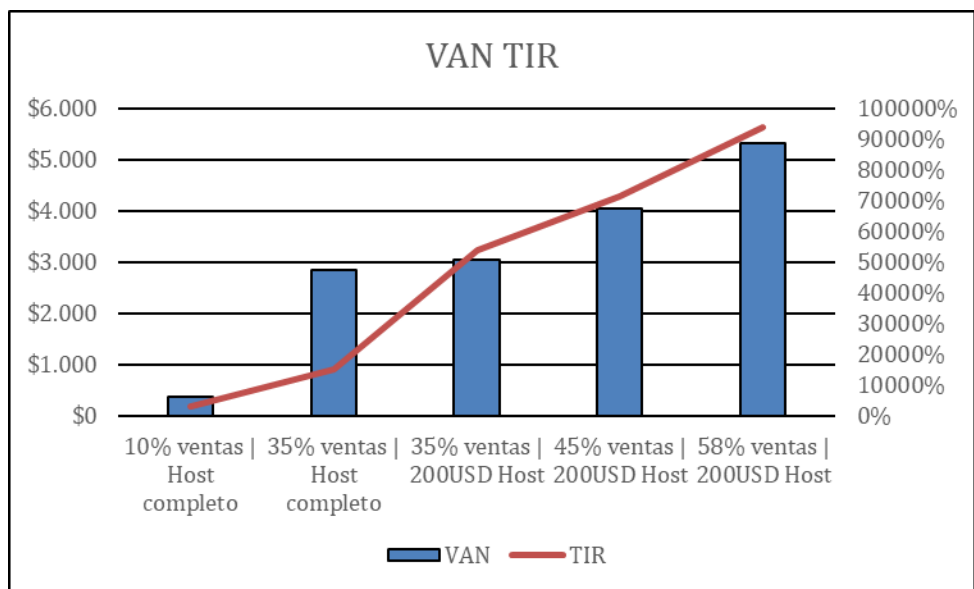


Figura 38. Análisis de VAN y TIR

En resumen del flujo de caja, este proyecto no requiere una alta inversión por parte de la empresa debido a los beneficios que posee por su categoría de partner y a los costos hundidos, despejando una de las variables complejas al momento de tomar la decisión sobre si se debe invertir o no. Incluso tomando el escenario más pesimista señalado en el análisis de sensibilidad, permite proyectar un escenario positivo con la habilitación de esta clasificación en los dos primeros años.

CAPITULO 7. CONCLUSIONES

AlfaPeople es una empresa dedicada a la venta de distintos servicios que son propuestos por distintos partners de Microsoft. Puesto a que su lógica de trabajo está basada principalmente en mantenerse como el partner preferido a bajo costo, se enfrenta frecuentemente a distintas competencias, con el objetivo de captar un buen segmento del mercado, olvidando a su base instalada. Esto genera que la salud económica de la empresa se base en la generación de negocios nuevos de altos montos más que generar flujos constantes con clientes ya cautivos.

Este trabajo propuso dar la importancia a los clientes actuales, utilizando los datos de los negocios actualmente ganados, generando segmentaciones y posibles estrategias de creación de negocios nuevos, con el objetivo de no depender sólo de los cierres en clientes nuevos, lo que podría generar una importante ventaja competitiva respecto a la competencia.

En cuanto a la revisión de los procesos, un punto importante fue identificar que, independiente no existan los diagramas respectivos, estos procesos siempre existen. El levantamiento en conjunto con el área de marketing y ventas fue lo que permitió identificar que no existía tratamiento para la base instalada. Por tanto, este proyecto abre a la posibilidad de evaluar la incorporación del diagrama definido para intentar generar mayores entradas de ingresos.

RFM es una buena técnica para comenzar en una organización que desea realizar análisis de los datos. Si bien no es un elemento que requiera utilizar técnicas complejas de data mining, genera un valor importante al momento de tener que categorizar a los clientes. Es común evaluar a un cliente sólo por los ingresos y no tomando en consideración el tiempo de compra o la frecuencia en que lo realiza, por lo que contar con un método e incorporarlo en las decisiones estratégicas puede brindar buenos resultados.

Este proyecto abre la puerta a investigaciones anexas que pueden ser generadas desde estos mismos datos, entre ellos poder definir cuáles son las mejores estrategias de contactabilidad, realizar comparaciones entre los cálculos de clúster con KMeans con los scores tradicionales de RFM o cálculo del tiempo de vida del cliente (Lifetime value) en la organización con los clústeres ya generados, pero sin olvidar la importancia de la definición de los procesos y la toma de decisiones basados en los datos existentes para el cuidado de la base instalada.

CAPITULO 8. BIBLIOGRAFIA

- Rivera, J.& Meuler, R. *Gartner Says Customer Relationship Management Software Market Grew 13.3 Percent*. Mayo 19, 2015 <<http://www.gartner.com/newsroom/id/3056118>>
- *Pulso*, enero 26, 2017 <<http://www.pulso.cl/empresas-mercados/ganancias-de-microsoft-suben-36-impulsados-por-mayor-demanda-de-servicios-cloud-en-4t16/>>
- Shaeffer, Chuck. *CRM Software Market Share Report*. <<http://www.crmsearch.com/crm-software-market-share.php>>
- Visión. *Alfapeople*. <<https://alfapeople.com/cl/acerca-de-nosotros/>>
- Misión. *Alfapeople*. <<https://alfapeople.com/cl/acerca-de-nosotros/>>
- Alfapeople. *Dynamics AX Plataforma de planificación de recursos empresariales para empresas eficientes, ágiles e innovadoras*. <<https://alfapeople.com/cl/dynamics-ax/>>
- AlfaPeople. *Dynamics CRM Solución para impulsar la productividad en ventas, servicio al cliente y eficiencia en marketing*. <<https://alfapeople.com/cl/dynamics-crm/>>
- Gestipolis. *¿Qué es el Balanced Scorecard y para qué sirve?* <<https://www.gestipolis.com/que-es-el-balanced-scorecard-y-para-que-sirve/>>
- Uriels, H. *Utiliza el Lienzo de Modelo de Negocio para potenciar tu Liderazgo*. (1/?). <https://leanstratego.wordpress.com/2013/11/22/utiliza-el-lienzo-de-modelo-de-negocio-para-potenciar-tu-liderazgo-1/>
- Guitart, I. (2020, August 14). *¿Cómo segmentar con la técnica RFM? Clase Ejecutiva UC*. <<https://www.claseejecutiva.uc.cl/blog/como-segmentar-con-la-tecnica-rfm/>>
- GeeksforGeeks. (2021, September 22). *Clustering in Machine Learning*. <<https://www.geeksforgeeks.org/clustering-in-machine-learning/>>
- Brownlee, J. (2020, August 20). *10 Clustering Algorithms With Python. Machine Learning Mastery*. <https://machinelearningmastery.com/clustering-algorithms-with-python/>
- Karaman, B. (2019, September 15). *Customer Segmentation - Towards Data Science. Medium*. <<https://towardsdatascience.com/data-driven-growth-with-python-part-2-customer-segmentation-5c019d150444>>

CAPITULO 9. ANEXOS

ANEXO A: Código de RFM

```
import pandas as pd
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
from datetime import datetime, timedelta
import pandas as pd
%matplotlib inline
import matplotlib.pyplot as plt
import numpy as np
import chart_studio.plotly as py
import plotly.offline as pyoff
import plotly.graph_objs as go
pyoff.init_notebook_mode()

###Selección de datos###
#####

opp_df = pd.read_csv("DF2020.csv", low_memory=False,error_bad_lines=False,
decimal = ',', sep =';')

###Pre procesamiento de datos###
#####

opp_df['opportunityid.psa_legalentity.psa_name'].value_counts(dropna = False)

Chile          873
Colombia       425
Central America 423
Ecuador        23
Mexico         14
Name: opportunityid.psa_legalentity.psa_name, dtype: int64

opp_df['opportunitystatecode'].value_counts(dropna = False)
2      996
1      762
Name: opportunitystatecode, dtype: int64

#Como parte de la revisión de los datos, comprobaremos los datos nulos
cols_nul = opp_df.columns[opp_df.isnull().any()]
opp_df[cols_nul].isnull().sum()
```

```

opportunityid.customerid_account.alfa_industry_entity.ap_name      163
opportunityid.customerid_account.alfa_relationshiptype              120
dtype: int64

```

```

opp_df['opportunityid.customerid_account.alfa_relationshiptype'].value_
counts(dropna = False)
2.849000e+08      754
2.849000e+08      686
NaN              120
2.849000e+08      113
2.849000e+08       21
5.698000e+08       21
5.698000e+08       12
8.547000e+08        8
5.698000e+08        7
8.547000e+08        4
1.139600e+09        2
8.547000e+08        2
1.139600e+09        2
1.424500e+09        2
2.849000e+08         1
7.782100e+08         1
1.709400e+09         1
2.279200e+09         1
Name: opportunityid.customerid_account.alfa_relationshiptype, dtype:
int64

```

Prospect: 284900000 Prospect - Former Customer : 284900001 Customer:
284900002 Supplier/Partner: 284900004 Customer & Supplier/Partner: 778210000
Competitor: 284900005 Other: 284900006

```

opp_df['opportunityid.customerid_account.alfa_relationshiptype'] = opp_d
f['opportunityid.customerid_account.alfa_relationshiptype'].replace([284
900000, 284900001, 284900002, 284900004, 778210000, 284900005, 284900006
], ['Prospect', 'Prospect - Former Customer', 'Customer', 'Supplier/Partne
r', 'Customer & Supplier/Partner', 'Competitor', 'Other'])

```

#Revisar el estado del cambio

```

opp_df['opportunityid.customerid_account.alfa_relationshiptype'].value_
counts(dropna = False)
Prospect              754
Customer              686
NaN                  120
Prospect - Former Customer  113
569800004.0           21
Other                 21
569800000.0           12
854700000.0            8
569800002.0            7
854700006.0            4
1139600000.0           2
1424500010.0           2

```

```

1139600008.0      2
854700003.0      2
2279200016.0     1
Customer & Supplier/Partner  1
1709400000.0     1
Competitor       1
Name: opportunityid.customerid_account.alfa_relationshiptype, dtype:
int64

```

```

opp_df['opportunityid.customerid_account.alfa_relationshiptype'] = opp_df[
'opportunityid.customerid_account.alfa_relationshiptype'].fillna('Unknown')

```

```

opp_df['opportunityid.customerid_account.alfa_relationshiptype'].value_counts(dropna = False)

```

```

Prospect          754
Customer          686
Unknown          120
Prospect - Former Customer  113
569800004.0       21
Other             21
569800000.0       12
854700000.0        8
569800002.0        7
854700006.0        4
1139600000.0       2
1139600008.0       2
1424500010.0       2
854700003.0        2
2279200016.0       1
Customer & Supplier/Partner  1
1709400000.0       1
Competitor        1

```

```

Name: opportunityid.customerid_account.alfa_relationshiptype, dtype:
int64

```

Al no tener los otros valores y al entender que si tenían una categoría que fue eliminada, serán categorizados como "Other". Se añaden además Competitor y Customer Supplier/Partner porque son outliers

```

opp_df['opportunityid.customerid_account.alfa_relationshiptype'] = opp_df[
'opportunityid.customerid_account.alfa_relationshiptype'].replace([569800004, 569800000, 854700000, 569800002, 854700006, 1139600008, 1139600000, 854700003, 1424500010, 1709400000, 2279200016, 'Competitor', 'Customer & Supplier/Partner'], ['Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other', 'Other'])
opp_df['opportunityid.customerid_account.alfa_relationshiptype'].value_counts(dropna = False)

```

```

Prospect          754
Customer          686
Unknown          120
Prospect - Former Customer  113

```

```
Other                                     85
Name: opportunityid.customerid_account.alfa_relationshiptype, dtype:
int64
```

```
###Transformación ###
#####
#####
df = opp_df.copy()
df['opportunityid.actualclosedate'] = pd.to_datetime(df['opportunityid.a
ctualclosedate'])
df['Q']= df['opportunityid.actualclosedate'].dt.quarter
df['QA']=pd.PeriodIndex(df['opportunityid.actualclosedate'], freq='Q')
df['QA'].value_counts(dropna = False)
2018Q4      233
2018Q2      156
2017Q4      155
2018Q1      155
2018Q3      146
2019Q3      131
2017Q3      125
2017Q2      113
2019Q4      106
2019Q1      101
2019Q2       97
2017Q1       72
2016Q4       64
2016Q3       42
2016Q2       38
2016Q1       22
2020Q1        1
2020Q2        1
Freq: Q-DEC, Name: QA, dtype: int64
```

```
df = pd.get_dummies(df, columns=['opportunityid.customerid_account.alfa_
industry_entity.ap_name', 'opportunityproductname', 'opportunityid.custom
erid_account.alfa_relationshiptype', 'opportunityid.psa_legalentity.psa_n
ame', 'QA', 'opportunitystatecode'],
                    drop_first=True, prefix=['Industria', 'Producto', 'Relacion
', 'BU', 'QA', 'estadoOp'])
```

Se eliminará Q que consistía para identificar a cuál quarter correspondía

```
df = df.drop("Q", axis= 1)
```

Se elimina el nombre de la oportunidad por no aportar

```
df = df.drop("opportunityid.name", axis= 1)
```

```
df_r = pd.DataFrame()
```

```
df_r[['opportunityid.customerid_account.name', 'opportunityid.actualclose
date']] = df[['opportunityid.customerid_account.name', 'opportunityid.act
ualclosedate']]
```

```
#Obtenemos la última fecha de cierre para calcular el recency
df_r['MasReciente'] = df_r['opportunityid.actualclosedate'].max()
```

```
#Recency = última opp cerrada - fecha de cierre de la opp
df_r['R'] = df_r['MasReciente'] - df_r['opportunityid.actualclosedate']
df_r['R'] = df_r['R']/np.timedelta64(1, 'D')
```

```
df_r= df_r.sort_values(by='opportunityid.actualclosedate', ascending=False)
df_r= df_r.drop(["opportunityid.actualclosedate", "MasReciente"], axis= 1
)
df_r= pd.DataFrame.drop_duplicates(df_r, subset= "opportunityid.customer
id_account.name")
df_r.head()
```

	opportunityid.customerid_account.name	R
0	472	303.352778
1	ABCDIN	973.116667
2	ACB Ingenieria	665.093750
3	ACHS	303.095833
4	ADVISICON	1273.901389

```
df_r =df_r.sort_values(['opportunityid.customerid_account.name'], ascend
ing=True)
```

```
df_r.reset_index(drop = True, inplace = True)
```

```
df_r.R.describe()
count      646.000000
mean       673.922853
std        317.626677
min         0.000000
25%        430.863194
50%        650.595833
75%        904.374653
max        1603.538889
Name: R, dtype: float64
```



```
#frequency: Contar todas las oportunidades ganadas. Para esto se tendrá que separar los datos
```

```
#ganadas
```

```
df_g = pd.DataFrame()  
df_g[['opportunityid.customerid_account.name', 'baseamount_base', 'estadoOp_2']] = df[['opportunityid.customerid_account.name', 'baseamount_base', 'estadoOp_2']]
```

```
df_g = df_g.drop(df_g[df_g.estadoOp_2 == 1].index)  
df_g = df_g.drop(['baseamount_base'], axis=1)  
df_g = df_g.groupby(['opportunityid.customerid_account.name']).count().reset_index()  
df_g.head()
```

	opportunityid.customerid_account.name	estadoOp_2
0	ACHS	1
1	ADVISICON	1
2	AGROSAVIA	14
3	AIEP	8
4	AKVA group Chile	1

```
df_g.rename(columns={'estadoOp_2': 'F'}, inplace=True)  
df_g.head()
```

	opportunityid.customerid_account.name	F
0	ACHS	1
1	ADVISICON	1
2	AGROSAVIA	14
3	AIEP	8
4	AKVA group Chile	1

```
df_g.F.describe()  
count    243.000000  
mean      3.135802  
std       3.119201  
min       1.000000  
25%      1.000000  
50%      2.000000
```

```
75%          4.000000
max          26.000000
Name: F, dtype: float64
```

```
#monetary= Lo mismo que F pero sumuando baseamount
```

```
df_m = pd.DataFrame()
df_m[['opportunityid.customerid_account.name', 'baseamount_base', 'estadoOp_2']] = df[['opportunityid.customerid_account.name', 'baseamount_base', 'estadoOp_2']]
df_m = df_m.drop(df_m[df_m.estadoOp_2 == 1].index)
df_m= df_m.drop(['estadoOp_2'], axis= 1)
df_m = df_m.groupby(['opportunityid.customerid_account.name']).sum().reset_index()
df_m.head()
```

	opportunityid.customerid_account.name	baseamount_base
0	ACHS	1146738
1	ADVISICON	917647
2	AGROSAVIA	52724254
3	AIEP	35403784
4	AKVA group Chile	353402

```
df_m.rename(columns={'baseamount_base': 'M'}, inplace=True)
```

```
#Agrupar con los clientes para obtener en un solo df RFM
```

```
df_rfm=pd.DataFrame()
df_rfm=df.copy()

df_rfm = df_rfm.drop(["opportunityid.actualclosedate", "baseamount_base", "priceperunit"], axis= 1)

#se borran duplicados ordenados de más reciente y así mantener R

df_rfm = df_rfm.groupby(['opportunityid.customerid_account.name']).mean().reset_index()
df_rfm= df_rfm.set_index('opportunityid.customerid_account.name').join(df_r.set_index('opportunityid.customerid_account.name'))
df_rfm= df_rfm.join(df_g.set_index('opportunityid.customerid_account.name'))
```

```

df_rfm= df_rfm.join(df_m.set_index('opportunityid.customerid_account.name'))
df_rfm.fillna(0)

#Se crea nuevo Dataframe para consultar RFM
dfrfm = pd.DataFrame()
dfrfm[['Recency', 'Frecuency', 'Monetary']] = df_rfm[['R', 'F', 'M']]
dfrfm.head()

```

Recency	Frecuency	Monetary
opportunityid.customerid_account.name		
472	303.352778	NaN
ABCDIN	973.116667	NaN
ACB Ingenieria	665.093750	NaN
ACHS	303.095833	1.0
ADVISICON	1273.901389	1.0

```

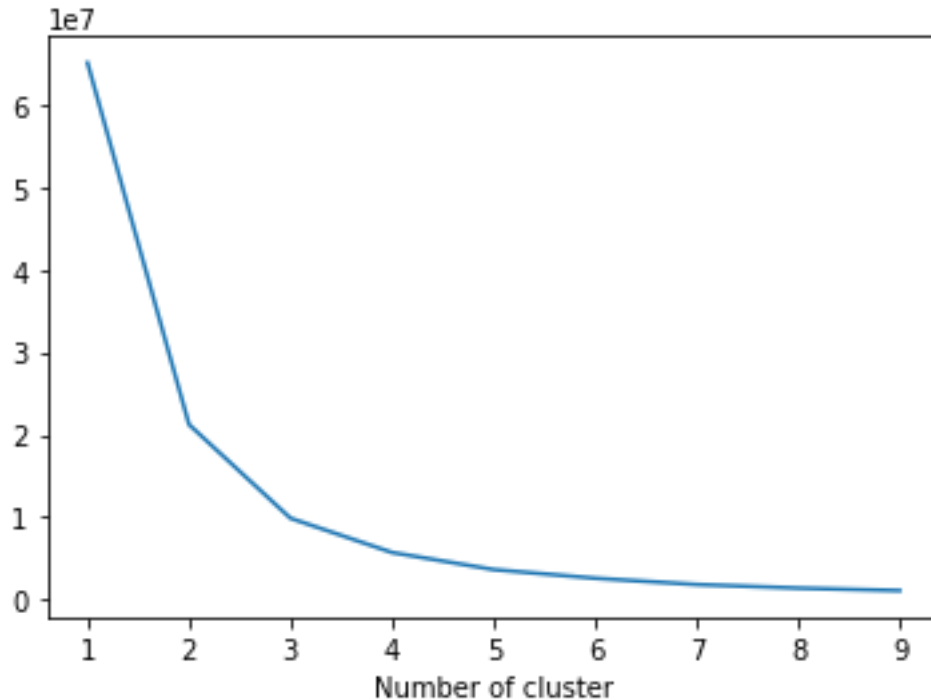
#Se llenan los nulos
dfrfm = dfrfm.fillna(0)
dfrfm.reset_index()
dfrfm.insert(0, 'id', range(1, 1+len(dfrfm)))

#guardar id y nombre para identificar cliente
dfnombre = pd.DataFrame()
dfnombre[['id', 'Cliente']] = dfrfm[['id', 'opportunityid.customerid_account.name']]

#Se inicia regla del codo para determinar K
from sklearn.cluster import KMeans

sse={}
tx_recency = dfrfm[['Recency']]
for k in range(1, 10):
    kmeans = KMeans(n_clusters=k, max_iter=1000).fit(tx_recency)
    tx_recency["clusters"] = kmeans.labels_
    sse[k] = kmeans.inertia_
plt.figure()
plt.plot(list(sse.keys()), list(sse.values()))
plt.xlabel("Number of cluster")
plt.show()

```



#4 clusters para R. Se añade a Dataframe

```
kmeans = KMeans(n_clusters=4)
kmeans.fit(dfrfm[['Recency']])
dfrfm['RecencyCluster'] = kmeans.predict(dfrfm[['Recency']])
```

#Se agrega función para ordenar por cluster number

```
def order_cluster(cluster_field_name, target_field_name, df, ascending):
    new_cluster_field_name = 'new_' + cluster_field_name
    df_new = df.groupby(cluster_field_name)[target_field_name].mean().re
set_index()
    df_new = df_new.sort_values(by=target_field_name, ascending=ascending
).reset_index(drop=True)
    df_new['index'] = df_new.index
    df_final = pd.merge(df, df_new[['cluster_field_name', 'index']], on=clus
ter_field_name)
    df_final = df_final.drop([cluster_field_name], axis=1)
    df_final = df_final.rename(columns={"index": cluster_field_name})
    return df_final
```

```
dfrfm = order_cluster('RecencyCluster', 'Recency', dfrfm, False)
```

```
dfrfm.groupby('RecencyCluster')['Recency'].describe()
```

	co unt	mean	std	min	25%	50%	75%	max
Recency Cluster								
0	78. 0	1236.98 5862	138.22 7577	1070.08 6111	1126.26 4757	1206.52 0833	1330.23 1771	1603.53 8889
1	170 .0	895.734 089	88.919 232	754.091 667	811.089 931	903.337 153	956.777 257	1061.62 8472
2	230 .0	600.934 891	84.035 041	450.068 750	536.373 958	593.030 208	664.751 389	747.910 417
3	168 .0	287.972 892	88.366 594	0.00000 0	219.838 368	300.930 556	360.454 514	441.187 500

```
#ordenar para ver cual es el mejor y el peor
def order_cluster(cluster_field_name, target_field_name, df, ascending):
    new_cluster_field_name = 'new_' + cluster_field_name
    df_new = df.groupby(cluster_field_name)[target_field_name].mean().re
set_index()
    df_new = df_new.sort_values(by=target_field_name, ascending=ascending
).reset_index(drop=True)
    df_new['index'] = df_new.index
    df_final = pd.merge(df, df_new[[cluster_field_name, 'index']], on=clus
ter_field_name)
    df_final = df_final.drop([cluster_field_name], axis=1)
    df_final = df_final.rename(columns={"index": cluster_field_name})
    return df_final

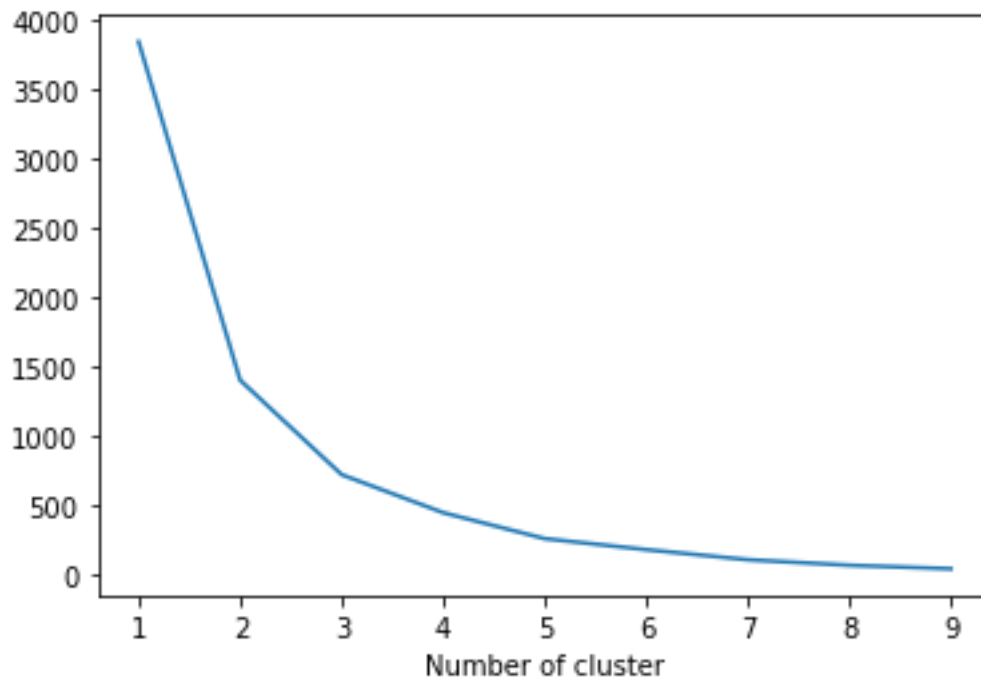
dfrfm = order_cluster('RecencyCluster', 'Recency', dfrfm, False)
dfrfm.groupby('RecencyCluster')['Recency'].describe()
```

	co unt	mean	std	min	25%	50%	75%	max
Recency Cluster								
0	78. 0	1236.98 5862	138.22 7577	1070.08 6111	1126.26 4757	1206.52 0833	1330.23 1771	1603.53 8889
1	170 .0	895.734 089	88.919 232	754.091 667	811.089 931	903.337 153	956.777 257	1061.62 8472
2	230 .0	600.934 891	84.035 041	450.068 750	536.373 958	593.030 208	664.751 389	747.910 417

	count	mean	std	min	25%	50%	75%	max
Recency Cluster								
3	168	287.972	88.366	0.00000	219.838	300.930	360.454	441.187
	.0	892	594	0	368	556	514	500

```
#Regla del codo para revisar F
from sklearn.cluster import KMeans

sse={}
tx_f = dfrfm[['Frecuency']]
for k in range(1, 10):
    kmeans = KMeans(n_clusters=k, max_iter=1000).fit(tx_f)
    tx_f["clusters"] = kmeans.labels_
    sse[k] = kmeans.inertia_
plt.figure()
plt.plot(list(sse.keys()), list(sse.values()))
plt.xlabel("Number of cluster")
plt.show()
```



```
#k-means para F
kmeans = KMeans(n_clusters=4)
kmeans.fit(dfrfm[['Frecuency']])
dfrfm['FrequencyCluster'] = kmeans.predict(dfrfm[['Frecuency']])
```

```
dfrfm = order_cluster('FrequencyCluster', 'Frequency', dfrfm, True)
#Revisar detalles de cada cluster
dfrfm.groupby('FrequencyCluster')['Frequency'].describe()
```

count	mean	std	min	25%	50%	75%	max
FrequencyCluster							
0	492.0	0.180894	0.385322	0.0	0.0	0.0	1.0
1	99.0	2.616162	0.710089	2.0	2.0	2.0	4.0
2	50.0	6.540000	1.875840	5.0	5.0	6.0	11.0
3	5.0	17.400000	5.079370	14.0	14.0	15.0	26.0

```
kmeans = KMeans(n_clusters=4)
kmeans.fit(dfrfm[['Monetary']])
dfrfm['MonetaryCluster'] = kmeans.predict(dfrfm[['Monetary']])
```

```
dfrfm = order_cluster('MonetaryCluster', 'Monetary', dfrfm, True)
#Revisar detalles de cada cluster M
dfrfm.groupby('MonetaryCluster')['Monetary'].describe()
```

count	mean	std	min	25%	50%	75%	max
Monetary Cluster							
0	586.0	5.866128e+05	1.448971e+06	0.0	0.000000e+00	0.0	225245.0 869873.8
1	370.0	1.695876e+07	7.075408e+06	9516978.0	1.160766e+07	13441900.0	19682873.0 33500571.0
2	200.0	5.330389e+07	1.480762e+07	35403784.0	4.105543e+07	50699761.0	61003475.0 79260314.0
3	3.0	1.473939e+08	3.513652e+07	114606749.0	1.288490e+08	143091349.0	163787439.5 184483530.0

```
### Interpretación y evaluación###
#####
#####
#Calcula el promedio basado en la media para generar un score promedio
dfrfm['OverallScore'] = dfrfm['RecencyCluster'] + dfrfm['FrequencyCluster'] + dfrfm['MonetaryCluster']
```

```
dfrfm.groupby('OverallScore')['Recency', 'Frecuency', 'Monetary'].mean()
```

	Recency	Frecuency	Monetary
OverallScore			
0	1240.396890	0.275362	4.550359e+05
1	916.347250	0.223684	2.317782e+05
2	630.186085	0.367021	2.357764e+05
3	393.100005	0.986755	1.204604e+06
4	438.417803	3.484848	6.218068e+06
5	444.328889	5.550000	1.709705e+07
6	356.196379	6.714286	3.407841e+07
7	236.516759	8.800000	5.016912e+07
8	190.011285	9.750000	1.024670e+08

```
#Se agrega una definición de segmentos basados en el score obtenido
dfrfm['Segment'] = 'Low-Value'
dfrfm.loc[dfrfm['OverallScore']>2, 'Segment'] = 'Mid-Value'
dfrfm.loc[dfrfm['OverallScore']>4, 'Segment'] = 'High-Value'
```

```
dfrfm.head()
```

opportunityid.cu	stomerid_accou	nt.name	id	Re	Frec	Mo	Recen	Freque	Monet	Over	Se
				cen	uen	net	cyClu	ncyClu	aryClu	allSc	gm
				cy	cy	ary	ster	ster	ster	ore	ent

0			472	1	303.352778	0.0	0.0	3	0	0	3	Mid-Value
1		ACHS		4	303.095833	1.0	1146738.0	3	0	0	3	Mid-Value

opportunityid.cu stomerid_accou nt.name	id	Re cen cy	Frec uen cy	Mo net ary	Recen cyClu ster	Freque ncyClu ster	Monet aryClu ster	Over allSc ore	Se gm ent	
2	AKR OS CIA. LTDA .	9	321. 0930 56	0.0	0.0	3	0	0	3	Mid- Val ue
3	AKVA group Chile	10	290. 3402 78	1.0	35340 2.0	3	0	0	3	Mid- Val ue
4	ASEB ANE CIO	21	257. 4534 72	1.0	14136 1.0	3	0	0	3	Mid- Val ue

```
dfrfm['Segment'].value_counts(dropna = False)
```

```
Low-Value      409
Mid-Value      184
High-Value      53
Name: Segment, dtype: int64
```

```
dfrfm['OverallScore'].value_counts(dropna = False)
```

```
2      188
1      152
3      151
0       69
4       33
5       20
7       15
6       14
8        4
Name: OverallScore, dtype: int64
```

```
### Exportar a Excel para trabajar con los resultados###
```

```
dfrfm.to_excel("RFM.xlsx")
```

```
### Análisis de gráficos###
```

```
#Revenue vs Frequency
```

```
tx_graph = dfrfm.query("Monetary < 1800000 and Frequency < 26")
```

```

plot_data = [
    go.Scatter(
        x=tx_graph.query("Segment == 'Low-Value'")['Frecuency'],
        y=tx_graph.query("Segment == 'Low-Value'")['Monetary'],
        mode='markers',
        name='Low',
        marker= dict(size= 7,
            line= dict(width=1),
            color= 'blue',
            opacity= 0.8
        )
    ),
    go.Scatter(
        x=tx_graph.query("Segment == 'Mid-Value'")['Frecuency'],
        y=tx_graph.query("Segment == 'Mid-Value'")['Monetary'],
        mode='markers',
        name='Mid',
        marker= dict(size= 9,
            line= dict(width=1),
            color= 'green',
            opacity= 0.5
        )
    ),
    go.Scatter(
        x=tx_graph.query("Segment == 'High-Value'")['Frecuency'],
        y=tx_graph.query("Segment == 'High-Value'")['Monetary'],
        mode='markers',
        name='High',
        marker= dict(size= 11,
            line= dict(width=1),
            color= 'red',
            opacity= 0.9
        )
    ),
]

plot_layout = go.Layout(
    yaxis= {'title': "Monetary"},
    xaxis= {'title': "Frecuency"},
    title='Segments'
)

fig = go.Figure(data=plot_data, layout=plot_layout)
pyoff.iplot(fig)

```

```

#Revenue vs Frequency
tx_graph = dfrfm.query("Monetary < 18000000 and Frecuency < 26")

plot_data = [
    go.Scatter(
        x=tx_graph.query("Segment == 'Low-Value'")['Frecuency'],
        y=tx_graph.query("Segment == 'Low-Value'")['Monetary'],
        mode='markers',
        name='Low',
        marker= dict(size= 7,
            line= dict(width=1),
            color= 'blue',
            opacity= 0.8
        )
    ),
    go.Scatter(
        x=tx_graph.query("Segment == 'Mid-Value'")['Frecuency'],
        y=tx_graph.query("Segment == 'Mid-Value'")['Monetary'],
        mode='markers',
        name='Mid',
        marker= dict(size= 9,
            line= dict(width=1),
            color= 'green',
            opacity= 0.5
        )
    ),
    go.Scatter(
        x=tx_graph.query("Segment == 'High-Value'")['Frecuency'],
        y=tx_graph.query("Segment == 'High-Value'")['Monetary'],
        mode='markers',
        name='High',
        marker= dict(size= 11,
            line= dict(width=1),
            color= 'red',
            opacity= 0.9
        )
    ),
]

plot_layout = go.Layout(
    yaxis= {'title': "Monetary"},
    xaxis= {'title': "Frecuency"},
    title='Segments'
)

fig = go.Figure(data=plot_data, layout=plot_layout)
pyoff.iplot(fig)

```

```
#Revenue Recency
```

```
tx_graph = dfrfm.query("Monetary < 18000000 and Frecuency < 26")
```

```
plot_data = [  
    go.Scatter(  
        x=tx_graph.query("Segment == 'Low-Value'")['Recency'],  
        y=tx_graph.query("Segment == 'Low-Value'")['Monetary'],  
        mode='markers',  
        name='Low',  
        marker= dict(size= 7,  
                    line= dict(width=1),  
                    color= 'blue',  
                    opacity= 0.8  
                )  
    ),  
    go.Scatter(  
        x=tx_graph.query("Segment == 'Mid-Value'")['Recency'],  
        y=tx_graph.query("Segment == 'Mid-Value'")['Monetary'],  
        mode='markers',  
        name='Mid',  
        marker= dict(size= 9,  
                    line= dict(width=1),  
                    color= 'green',  
                    opacity= 0.5  
                )  
    ),  
    go.Scatter(  
        x=tx_graph.query("Segment == 'High-Value'")['Recency'],  
        y=tx_graph.query("Segment == 'High-Value'")['Monetary'],  
        mode='markers',  
        name='High',  
        marker= dict(size= 11,  
                    line= dict(width=1),  
                    color= 'red',  
                    opacity= 0.9  
                )  
    ),  
]
```

```
plot_layout = go.Layout(  
    yaxis= {'title': "Monetary"},  
    xaxis= {'title': "Recency"},  
    title='Segments'  
)
```

```

fig = go.Figure(data=plot_data, layout=plot_layout)
pyoff.iplot(fig)

# Revenue vs Frequency
tx_graph = dfrfm.query("Recency < 50000 and Frecuency < 2000")

plot_data = [
    go.Scatter(
        x=tx_graph.query("Segment == 'Low-Value'")['Recency'],
        y=tx_graph.query("Segment == 'Low-Value'")['Frecuency'],
        mode='markers',
        name='Low',
        marker= dict(size= 7,
            line= dict(width=1),
            color= 'blue',
            opacity= 0.8
        )
    ),
    go.Scatter(
        x=tx_graph.query("Segment == 'Mid-Value'")['Recency'],
        y=tx_graph.query("Segment == 'Mid-Value'")['Frecuency'],
        mode='markers',
        name='Mid',
        marker= dict(size= 9,
            line= dict(width=1),
            color= 'green',
            opacity= 0.5
        )
    ),
    go.Scatter(
        x=tx_graph.query("Segment == 'High-Value'")['Recency'],
        y=tx_graph.query("Segment == 'High-Value'")['Frecuency'],
        mode='markers',
        name='High',
        marker= dict(size= 11,
            line= dict(width=1),
            color= 'red',
            opacity= 0.9
        )
    ),
]

plot_layout = go.Layout(
    yaxis= {'title': "Frecuency"},
    xaxis= {'title': "Recency"},
    title='Segments'
)

```

```
)  
fig = go.Figure(data=plot_data, layout=plot_layout)  
pyoff.iplot(fig)
```

