



**Universidad del
Rosario**

Escuela de Administración
Escuela de Ingeniería Ciencia y Tecnología

Maestría en Business Analytics

Análisis y estructuración de clientes para una empresa del sector logístico. Caso Italian
Shipping Cargo

Presentado por:

Sergio Andrés Orjuela Torres – Bianca Johanna Ordoñez Lugo

Bajo la dirección de:

Jeison Pinilla

Bogotá, D.C. 20 de mayo de 2023



Universidad del
Rosario

Escuela de Administración
Escuela de Ingeniería Ciencia y Tecnología

Análisis y estructuración de clientes para una empresa del sector logístico. Caso Italian
Shipping Cargo

Tipo de trabajo (Trabajo de Grado).

Autores.

Sergio Andrés Orjuela Torres – Bianca Johanna Ordoñez Lugo

Tutor

Jeison Pinilla

Programa

Maestría en Business Analytics

17 de junio de 2023

Bogotá, Colombia

Contenido

Declaración de originalidad y autonomía	6
Lista de tablas.....	9
Abreviaturas	10
Glosario	11
Resumen	14
Palabras clave.....	14
Abstract	15
Keywords	15
1. Introducción	16
2. Objetivos	18
2.1. Objetivo general	18
2.2. Objetivos específicos.....	18
3. Alcance.....	19
4. Metodología	20
5. Cronograma.....	22
6. Descripción de la Situación organizacional donde se realizará el proyecto (Contexto) 23	
7. Descripción de la situación estudio de caso y/o problemática empresarial y método y/o estrategia a aplicar para su solución	25
8. Comprensión del negocio.....	28
8.1. Aplicación Canvas.....	29
8.1.1. Problema	29
8.2. Solución.....	30
8.3. Bases.....	30
8.4. Hipótesis.....	30
8.5. Actores.....	30
8.6. Acciones	31
8.7. KPIs	31
8.8. Propuesta de valor	31

8.9. Riesgos	31
8.10. Impacto	32
9. Descripción de las fuentes de información	33
10. Comprensión de los datos.	34
10.1. Análisis de tipos de datos estructurados:.....	34
10.2. Preparación de los datos	36
10.3. Limpieza de datos ¿cómo es el entendimiento y el tratamiento de los datos de parte de las personas que los manipulan?.....	37
10.4. Análisis descriptivo de los datos ¿Cantidad de filas y columnas? ¿Qué número de periodos, años y meses tiene? ¿Qué métricas de análisis tiene la empresa?...37	
10.4.1. Clientes:	37
10.4.2. Operaciones:	38
10.4.3. SIICEX:	38
11. Descripción de las alternativas, estrategias y/o acciones que se toman en el análisis de la solución a la problemática.....	39
11.1. Estrategias para dar solución a las diferentes problemáticas de la empresa	39
11.2. Preparación de los datos y proceso de modelamiento a implementar ...40	
11.2.1. Análisis descriptivo no supervisado	41
11.2.2. Normalización y escalación de los datos:	44
11.2.3. Desarrollo de K-means	52
11.2.4. Análisis e interpretación de los clusters.....	58
11.2.5. Entrenamiento del modelo Kmeans	59
11.2.6. Descriptivos y características de los diferentes clusters	60
11.3. Dashboards	79
11.3.1. Dashboard Comercial	80
11.3.2. Dashboard gerencial	85
11.3.3. Funnel de ventas	87
11.4. Conclusiones de los Dashboards:	88
12. Plan y recomendaciones de implementación y aplicación	90
12.1. Recomendaciones:.....	90
12.1.1. Estructuración de datos	90

12.1.2. Procedimientos.....	90
12.1.3. Almacenamiento de la información	90
12.1.4. Capacitación al personal	91
12.2. Implementación de la alternativa:	91
12.2.1. Personal involucrado y responsabilidades:	92
12.2.2. Momento y lugar de implementación:	92
12.2.3. Resultado a obtener:.....	92
12.2.4. Medición del éxito o fracaso:.....	93
13. Conclusiones	94
Referencias bibliográficas.....	96

Declaración de originalidad y autonomía

Declaro(amos) bajo la gravedad del juramento, que he(mos) escrito el presente Proyecto Aplicado Empresarial (PAE), en la propuesta de solución a una problemática en el campo de conocimientos del programa de Maestría por mi(nuestra) propia cuenta y que, por lo tanto, su contenido es original.

Declaro(amos) que he(mos) indicado clara y precisamente todas las fuentes directas e indirectas de información y que este PAE no ha sido entregado a ninguna otra institución con fines de calificación o publicación.



BIANCA JOHANNA ORDOÑEZ



SERGIO ANDRES ORJUELA TORRES

Firmado en Bogotá, D.C. el 20 de mayo de 2023

Declaración de exoneración de responsabilidad

Declaro(amos) bajo la gravedad del juramento, que he(mos) escrito el presente Proyecto Aplicado Empresarial (PAE), en la propuesta de solución a una problemática en el campo de conocimientos del programa de Maestría por mi(nuestra) propia cuenta y que, por lo tanto, su contenido es original.

Declaro(amos) que he(mos) indicado clara y precisamente todas las fuentes directas e indirectas de información y que este PAE no ha sido entregado a ninguna otra institución con fines de calificación o publicación.



BIANCA JOHANNA ORDOÑEZ



SERGIO ANDRES ORJUELA TORRES
Firmado en Bogotá, D.C. el 20 de mayo de 2023

Lista de Graficas

Gráfica 1 Descripción bases de datos SIICEX	36
Gráfica 2 Explicación funnel de ventas	39
Gráfica 3 Modelo de analítica.....	40
Gráfica 4 Análisis de histogramas	45
Gráfica 5 distribución de datos	46
Gráfica 6 Variabilidad de los datos.....	48
Gráfica 7 Correlación scather_matrix.....	50
Gráfica 8 Mapa de calor.....	51
Gráfica 9 Método del codo	53
Gráfica 10 Índice de la silueta	54
Gráfica 11 Prueba de coeficiente de silueta.....	55
Gráfica 12 Gráfico t_SNE 2D.....	56
Gráfica 13 Gráfica 3D.....	57
Gráfica 14 Formula k-means	59
Gráfica 15 Dashboard comercial	81
Gráfica 16 Filtros dashboard comercial.....	82
Gráfica 17 Gráfico de anillos dashboard comercial.....	82
Gráfica 18 Referencias de datos	83
Gráfica 19 Grafico de barras.....	83
Gráfica 20 Tabla de clientes	84
Gráfica 21 Filtros dashboard gerencial.....	86
Gráfica 22 Gráficos de anillos dashboard gerencial.....	86
Gráfica 23 Etiquetas de datos dashboard gerencial	87
Gráfica 24 Gráficos dashboard gerencial.....	87
Gráfica 25 Funnel de ventas	88

Lista de tablas

Tabla 1 Valores de correlación	51
Tabla 3 Características clúster 1	60
Tabla 4 Características clúster 2	62
Tabla 5 Características clúster 3	65
Tabla 6 Características clúster 4	67
Tabla 7 Características clúster 5	70
Tabla 8 Características clúster 6	72
Tabla 9 Características clúster 7	74
Tabla 10 Características clúster 8	76

Abreviaturas

SIICEX Sistema de Información Arancelaria Vía Internet, el cual contiene Estadísticas de Comercio actualizadas, con datos de comercio mensual desde 2007.

FOB El Incoterm Free On Board, que se traduce como 'franco a bordo' se utiliza exclusivamente para transporte marítimo o fluvial. Con este término comercial internacional: El vendedor entrega la mercancía en el puerto de embarque y asume los costos de trámites aduaneros de exportación y licencias de exportación.

Glosario

Canvas: El Business Model Canvas, o lienzo de modelo de negocio, es una herramienta que permite describir y visualizar de manera holística el modelo de negocio de una empresa. Se representa en un lienzo dividido en diferentes bloques que representan los elementos clave del negocio, como segmentos de clientes, propuesta de valor, canales de distribución, relaciones con los clientes, fuentes de ingresos, actividades clave, recursos clave, socios clave y estructura de costos. El canvas ayuda a los emprendedores y equipos de negocio a entender y comunicar de manera clara cómo opera su negocio, identificar áreas de mejora, explorar nuevas oportunidades y diseñar estrategias para lograr el éxito empresarial.

Cluster: En el contexto de análisis de datos, un cluster se refiere a un grupo o conjunto de elementos similares que comparten características o atributos comunes. Es una técnica utilizada en minería de datos y aprendizaje automático para identificar patrones o estructuras subyacentes en un conjunto de datos. Los elementos dentro de un cluster son más similares entre sí que con los elementos de otros clústeres, lo que permite agruparlos en categorías distintas. Los algoritmos de clustering buscan maximizar la similitud interna y minimizar la similitud entre clusters diferentes, lo que permite una mejor comprensión de la estructura y relaciones dentro de los datos. El clustering puede ser utilizado para segmentar clientes, clasificar documentos, identificar grupos de productos, entre otros usos en el análisis de datos.

Dashboard: Un dashboard es una herramienta de visualización de datos que presenta información clave de manera clara y concisa a través de gráficos, tablas y otros elementos visuales. Proporciona una vista panorámica de los indicadores clave de rendimiento (KPI) y métricas relevantes para ayudar a las organizaciones a monitorear y analizar su desempeño en tiempo real. Un dashboard permite a los usuarios tomar decisiones informadas y rápidas al proporcionar una representación visual intuitiva de los datos, facilitando la identificación de tendencias, patrones y oportunidades de mejora.

Funnel: El funnel de ventas es un modelo visual que muestra las etapas por las que pasa un cliente potencial desde que muestra interés en un producto o servicio hasta que se convierte en un cliente real. Ayuda a las empresas a comprender y optimizar el proceso de ventas, identificando las etapas clave y tomando medidas para impulsar a los prospectos hacia la compra.

Gobierno de datos: Marco de políticas, procesos y controles para garantizar la calidad, integridad, seguridad y uso adecuado de los datos en una organización. Incluye la definición de roles y responsabilidades, estándares de calidad de datos y cumplimiento normativo.

K Means: K-means es un algoritmo de agrupamiento no supervisado utilizado en análisis de datos y aprendizaje automático. Su objetivo principal es agrupar un conjunto de datos en k grupos o clústeres, donde cada punto de datos pertenece al clúster más cercano en función de la distancia euclidiana. El algoritmo K-means comienza seleccionando aleatoriamente k centroides, que son puntos representativos iniciales para cada clúster. Luego, itera hasta que se logre la convergencia, reasignando los puntos de datos a los clústeres más cercanos y recalculando los centroides. El resultado final es un conjunto de clústeres donde los puntos de datos dentro de

cada clúster son similares entre sí y diferentes de los puntos en otros clústeres. El algoritmo K-means es ampliamente utilizado en la segmentación de clientes, análisis de imágenes, reconocimiento de patrones y otras aplicaciones de análisis de datos.

KPI: Un KPI (Key Performance Indicator) o indicador clave de rendimiento, es una medida cuantitativa utilizada para evaluar el desempeño de una organización, equipo o proceso en relación con los objetivos establecidos. Los KPIs son métricas seleccionadas cuidadosamente que permiten monitorear y medir el progreso hacia metas específicas.

Resumen

Se puede definir business analytics como la solución que nace por medio del análisis de datos a las problemáticas específicas de un negocio, parte de esta solución es la creación de un proceso de inteligencia empresarial el cual va a impactar positivamente las diferentes áreas de una empresa ya que establece indicadores clave para su análisis y posterior a sus conclusiones tomar mejores decisiones y estrategias de negocio.

Este proyecto tiene como fin definir una nueva estrategia de analítica en el área comercial de la empresa Italian shipping Cargo, por medio de la creación de los procesos para el análisis de datos externo (Mercado) e interno (Diagrama de flujo de datos), la definición de indicadores clave, funnel de ventas, Tarjetas de clientes con sus características. Estos hitos están enfocados en aumentar el número de operaciones (Ventas) y estructurar la gestión y logística comercial

Palabras clave

Embudo de ventas, diagrama de flujo de datos, analítica de negocios, Indicador clave de desempeño.

Abstract

Analysis and structuring of clients for a company in the logistics sector. Italian Shipping Cargo
Case

Business analytics can be defined as the solution that is born through the analysis of data to the specific problems of a business, part of this solution is the creation of a business intelligence process which will positively impact the different areas of a company since which establishes key indicators for analysis and subsequent conclusions to make better decisions and business strategies.

The purpose of this project is to define a new analytics strategy in the commercial area of the Italian shipping Cargo company, through the creation of processes for external (Market) and internal (Data Flow Diagram) data analysis, definition of key indicators, sales funnel, customer cards with their characteristics. These milestones are focused on increasing the number of operations (Sales) and structuring commercial management and logistics

Keywords

Funnel, flowchart,, business analytics, Key Performance Indicator KPI

1. Introducción

La gestión eficiente de datos se ha convertido en una necesidad estratégica para las compañías modernas, especialmente para aquellas en el sector logístico. En este sentido, la empresa Italian Shipping Cargo se ha enfrentado a diversos retos relacionados con la gestión de datos y la analítica empresarial. El presente trabajo tiene como objetivo analizar y estructurar los clientes de Italian Shipping Cargo mediante la implementación de procesos de inteligencia empresarial y el establecimiento de un gobierno de datos sólido.

El caso de estudio se enfoca en la empresa Italian Shipping Cargo, la cual ha experimentado dificultades para establecer procesos de analítica empresarial y gobierno de datos eficientes. Se busca proponer una estrategia de analítica empresarial enfocada en la gestión comercial, que permita aumentar el número de ventas y estructurar la gestión y logística comercial.

Los objetivos específicos de este proyecto incluyen la creación de procesos para el análisis de datos externo (mercado) e interno (diagrama de flujo de datos), la definición de indicadores clave, el diseño de un funnel de ventas y la creación de tarjetas de clientes con sus características. Además, se busca establecer políticas y procedimientos para la seguridad y calidad de los datos, así como brindar capacitación y formación adecuada al personal para garantizar el cumplimiento del gobierno de datos.

La importancia de este proyecto radica en que la analítica empresarial y el gobierno de datos son aspectos críticos para el éxito y la supervivencia de las empresas modernas. Además, la

implementación de esta estrategia en Italian Shipping Cargo permitirá mejorar la gestión y logística comercial, aumentar la eficiencia y eficacia en la toma de decisiones y mejorar la competitividad de la empresa.

En cuanto a los métodos utilizados, se empleará una metodología de investigación mixta que combina el análisis cuantitativo y cualitativo. Se utilizarán técnicas de minería de datos, análisis estadístico y herramientas de visualización de datos para analizar los datos de la empresa. Además, se realizarán entrevistas a expertos y personal clave de la empresa para obtener información cualitativa y garantizar la validez y fiabilidad de los resultados.

El alcance de este proyecto se limita a la gestión comercial de Italian Shipping Cargo, y se enfoca en la creación de procesos para el análisis de datos y el establecimiento de políticas y procedimientos para la seguridad y calidad de los datos.

Las limitaciones del estudio incluyen la disponibilidad y calidad de los datos, así como el tiempo y recursos limitados para la implementación de la estrategia. Se supone que el personal de la empresa está dispuesto a participar y colaborar en el proceso de implementación.

En resumen, este proyecto busca implementar una estrategia de analítica empresarial y gobierno de datos en Italian Shipping Cargo, enfocada en la gestión comercial. Se utilizará una metodología de investigación mixta, y se espera obtener beneficios significativos para la empresa, incluyendo la mejora de la gestión y logística comercial, la eficiencia y eficacia en la toma de decisiones y la competitividad en el mercado.

2. Objetivos

Los objetivos propuestos para el presente trabajo son:

2.1.Objetivo general

Definir una nueva estrategia de analítica de negocios para el área comercial de la empresa Italian Shipping para la logística de ventas y la gestión comercial incluyendo la segmentación de clientes.

2.2.Objetivos específicos

- Desarrollar un análisis de la gestión de datos estructurados de la empresa en cuestión (italian Shipping Cargo), de tal forma que contribuyan activamente al cumplimiento de los objetivos comerciales de la compañía.
- Establecer un diagrama de gestión de datos del área comercial con el fin de garantizar un debido proceso en la recolección de información, creación de indicadores clave y analítica enfocada a resultados
- Mejorar la logística de ventas con base a los datos internos de la empresa, para definir las etapas y los porcentajes de conversión internos en el proceso comercial.
- Establecer los KPIs clave en el área comercial de la logística de ventas (Perfomance) y Gestión de ventas (Resultados y clientes)
- Establecer clusters de clientes describiendo sus características

3. Alcance

El alcance de este proyecto de tesis se limita al análisis y mejora de la gestión comercial de Italian Shipping Cargo. Se centra en el desarrollo e implementación de una estrategia de analítica empresarial y gobierno de datos específicamente diseñada para el área comercial de la empresa. Los procesos, indicadores clave, funnel de ventas y tarjetas de clientes propuestos estarán enfocados en optimizar las operaciones comerciales, aumentar el número de ventas y mejorar la eficiencia y eficacia en la gestión y logística de esta área. Cabe destacar que el alcance de este proyecto no abarca otras áreas de la empresa, como la gestión de operaciones, recursos humanos o finanzas. El objetivo principal es proporcionar a Italian Shipping Cargo una estrategia analítica sólida y orientada al área comercial para mejorar su desempeño y competitividad en el mercado logístico.

4. Metodología

Para el desarrollo del proyecto se propone la siguiente metodología:

1. Revisión de literatura: Se realizará una revisión exhaustiva de la literatura académica y empresarial relacionada con la analítica empresarial, el gobierno de datos y la gestión comercial en el sector logístico. Esta revisión permitirá obtener un panorama completo de las mejores prácticas, enfoques y herramientas utilizadas en la industria, así como identificar posibles desafíos y oportunidades específicas para Italian Shipping Cargo. También se tendrán en cuenta la aplicación de los ejercicios que se desarrollaron en clase.

2. Recopilación de datos: Se recopilarán datos relevantes para el análisis de la gestión comercial de Italian Shipping Cargo. Esto incluirá datos internos de la empresa, como registros de ventas, informes de clientes y datos de operaciones comerciales, así como los datos de las bases de SIICEX Colombia, la cual es un sistema integrado de información de comercio exterior, y cuya base nos proporciona toda la información necesaria para realizar el análisis del mercado

3. Análisis cuantitativo y cualitativo: Se genera un orden para la realización del modelo no supervisado con base a la metodología utilizada

3.1. Comprensión de los datos

3.2. descripción del modelo que se va a implementar

3.3. descripción de las variables que se van a utilizar

3.4. procesamiento de los datos

3.5. Correlaciones de las variables

3.6. Desarrollo y análisis de los resultados de K-means

3.7. Análisis e interpretación de los clusters

3.8. descripción de los tipos de dashboards

4. Diseño e implementación de la estrategia de analítica empresarial: Con base en los resultados del análisis, se diseñará una estrategia de analítica empresarial específicamente dirigida al área comercial de Italian Shipping Cargo. Esto incluirá la definición de indicadores clave de rendimiento (KPIs), la aplicación de la metodología para que el área comercial logre diseñar un funnel de ventas. Se utilizarán herramientas de visualización de datos para representar de manera clara y comprensible los resultados del análisis y facilitar la toma de decisiones.

5. Establecimiento de políticas y procedimientos de gobierno de datos: Para garantizar la seguridad y calidad de los datos utilizados en el análisis y la gestión comercial, se establecerán políticas y procedimientos de gobierno de datos. Esto incluirá la definición de roles y responsabilidades, la implementación de medidas de seguridad y privacidad de los datos, y la creación de un marco para la gestión y actualización de los datos.

5. Cronograma

Se elige como escenario más conveniente la metodología híbrida entre el ágil y tradicional, ya que permite la interacción entre el diseño de la logística y gestión comercial, el desarrollo iterativo de las tareas de analítica y pruebas del producto, asegurando la calidad en los datos y resultados. Es riesgoso elegir el escenario tradicional únicamente ya que sin terminar la etapa de diseño de las tareas de analítica no es posible iniciar el desarrollo para realizar pruebas, teniendo como consecuencia pérdida de tiempo si se presentaran ajustes significativos en de diseño en los procesos de la estrategia a medida del tiempo por lo cual utilizamos también la metodología scrum para iterar en los procesos analíticos y de gestión comercial.

El sistema híbrido tiene en cuenta aspectos de gran importancia a la hora de la gestión de proyectos predictivos y adaptativos con base a la salida de nuevos procesos y resultados.

En la metodología ágil se escogerá la Scrum y en el modelo tradicional diagrama de Gantt.

6. Descripción de la Situación organizacional donde se realizará el proyecto (Contexto)

La empresa Italian Shipping Cargo se caracteriza por ser una empresa familiar que opera en el sector de la logística internacional. A lo largo de su trayectoria, ha experimentado un crecimiento constante y ha logrado establecerse en el mercado. Sin embargo, la falta de inversión en el control y estandarización de la información ha generado diversos desafíos en cuanto a la gestión de datos y la toma de decisiones informada.

En el contexto actual, la empresa enfrenta dificultades para acceder a información confiable, actualizada y precisa sobre sus operaciones comerciales. La falta de procesos establecidos para el análisis de datos y la carencia de un gobierno de datos sólido han llevado a la existencia de información dispersa, inconsistente y poco estructurada en diferentes áreas y departamentos de la organización. Pero para el desarrollo del presente proyecto nos concentraremos únicamente en el área comercial de la organización.

Además, Italian Shipping Cargo no ha desarrollado programas de seguridad de la información, lo que implica una vulnerabilidad en términos de confidencialidad, integridad y disponibilidad de los datos. La falta de políticas y procedimientos de seguridad de la información ha dejado expuesta a la empresa a posibles amenazas y riesgos relacionados con la privacidad y la protección de datos sensibles.

El hecho de ser una empresa familiar también ha influido en la gestión de la información. En muchas ocasiones, las decisiones se han basado en la intuición y la experiencia, sin contar con análisis y datos respaldados. La falta de un enfoque analítico ha limitado la

capacidad de la organización para identificar oportunidades de mejora, optimizar procesos y aumentar su competitividad en el mercado.

En resumen, Italian Shipping Cargo se encuentra en una situación donde la gestión de la información y el análisis de datos se han vuelto críticos para su crecimiento y éxito continuo. La falta de inversión en el control y estandarización de la información, así como la ausencia de programas de seguridad de la información, han creado una necesidad urgente de implementar una estrategia de analítica empresarial y gobierno de datos en el área comercial. El objetivo es mejorar la toma de decisiones, aumentar la eficiencia en la gestión y logística comercial, y establecer medidas de seguridad adecuadas para proteger la información sensible de la empresa.

7. Descripción de la situación estudio de caso y/o problemática empresarial y método y/o estrategia a aplicar para su solución

El problema principal en Italian Shipping Cargo es la falta de gestión eficiente de datos y la ausencia de un gobierno de datos sólido en el área comercial. Esto ha llevado a la dispersión, inconsistencia y falta de estandarización de la información, lo que dificulta la toma de decisiones fundamentadas y limita el potencial de crecimiento y competitividad de la empresa.

A raíz del problema anterior se derivan los siguientes problemas: Falta de procesos establecidos para el análisis de datos externos (mercado) e internos (diagrama de flujo de datos).

Ausencia de indicadores clave y herramientas de análisis para medir el desempeño y evaluar la eficacia de las estrategias comerciales, por lo que es inexistente el funnel de ventas que permita visualizar y optimizar el flujo de clientes potenciales a lo largo del proceso de ventas.

Falta de tarjetas de clientes con sus características, lo que dificulta la comprensión y segmentación adecuada de la base de clientes.

Ausencia de políticas y procedimientos para la seguridad y calidad de los datos, como también del gobierno de datos, lo que pone en riesgo la confidencialidad y la integridad de la información.

Las evidencias que respaldan estos problemas se basan en hechos y observaciones concretas dentro de la empresa. Estas incluyen:

- Datos dispersos y desorganizados en diferentes áreas y departamentos.
- Información inconsistente y no estandarizada en los reportes y sistemas utilizados.
- Falta de indicadores y métricas para evaluar el desempeño comercial.
- Dificultades para realizar un seguimiento efectivo de los clientes y su ciclo de vida.
- Falta de medidas de seguridad y control para proteger los datos de la empresa.

Los síntomas que sugieren que estos son los problemas reales en el caso incluyen dificultades para tomar decisiones informadas debido a la falta de información confiable y actualizada, pérdida de oportunidades de negocio debido a la falta de análisis y seguimiento adecuados de clientes potenciales, lo que también se ve afectado por no responder rápidamente a los clientes potenciales. Por último, encontramos riesgos potenciales de seguridad de la información debido a la falta de políticas y procedimientos establecidos.

Estos problemas están interrelacionados y se refuerzan mutuamente. La falta de gestión eficiente de datos y el gobierno de datos deficiente son la causa raíz de los problemas secundarios identificados. La falta de procesos establecidos y la ausencia de indicadores y métricas claras son el resultado directo de la falta de gestión de datos. Además, la falta de políticas y procedimientos para la seguridad de la información agrava aún más la situación.

A corto plazo, las ramificaciones de estos problemas incluyen:

- Toma de decisiones ineficaz y basada en intuición en lugar de datos.
- Pérdida de oportunidades de negocio y clientes potenciales debido a la falta de seguimiento y análisis adecuados.
- Ineficiencia y falta de coherencia en la gestión comercial y logística.

A largo plazo, las ramificaciones de estos problemas pueden incluir:

- Pérdida de competitividad en el mercado debido a la falta de análisis de datos y toma de decisiones estratégicas.
- Riesgos de seguridad y protección de datos sensibles.

- Estancamiento en el crecimiento y desarrollo de la empresa debido a la falta de una cultura de análisis y gobierno de datos. En cuanto a este punto se pudo observar que en los últimos años el crecimiento de la compañía no ha sido notorio en el mercado, por lo que se busca impulsar este crecimiento.

En cuanto al método y/o estrategia a aplicar para su solución, se propone implementar una estrategia de analítica empresarial y gobierno de datos enfocada en la gestión comercial. Esto implica establecer procesos para el análisis de datos externos e internos, definir indicadores clave, diseñar un funnel de ventas y crear tarjetas de clientes con sus características. Además, se deberán establecer políticas y procedimientos para la seguridad y calidad de los datos, y se brindará capacitación y formación adecuada al personal para garantizar el cumplimiento del gobierno de datos.

8. Comprensión del negocio.

La etapa de "Comprensión del negocio" en la metodología CRISP para el caso de Italian Shipping Cargo implica una revisión exhaustiva de la compañía, sus procesos de gobierno de datos y el análisis de datos de los clientes. Esta revisión permitirá identificar fortalezas, áreas de mejora y alinear el proyecto con los objetivos estratégicos de la compañía. Sin embargo, hasta el momento, la empresa ha enfrentado desafíos en la recolección y aprovechamiento adecuado de datos en el área comercial.

A pesar de contar con más de 30 años de experiencia en el mercado de exportaciones e importaciones, Italian Shipping Cargo ha carecido de un proceso establecido para la recopilación y seguimiento de datos comerciales. Esto ha resultado en la falta de trazabilidad y en la limitada utilización de recursos valiosos, como la plataforma SIICEX, que proporciona información detallada sobre el mercado, incluyendo datos sobre las empresas importadoras, tipos de productos, cantidades, vías de transporte y otras variables relevantes.

La falta de aprovechamiento de estos recursos ha generado una brecha significativa en la analítica de datos de la empresa, lo cual ha retrasado su crecimiento y ha limitado la toma de decisiones comerciales estratégicas. La problemática identificada se convierte en una espina de pescado con múltiples aspectos que requieren atención y corrección.

En particular, se evidencia la falta de buenas prácticas en el manejo de datos en el área comercial, donde actualmente se utilizan bases de datos que generan informes planos mediante

tablas dinámicas. Esto limita la capacidad de análisis y la generación de información accionable para impulsar el aumento de ingresos y la captación de nuevos clientes potenciales.

Para superar esta problemática y mejorar la comprensión del negocio, se propone implementar una estrategia de analítica empresarial y gobierno de datos en el área comercial de Italian Shipping Cargo, enfocada al área comercial. Esto implica establecer procesos adecuados para la recolección, almacenamiento y análisis de datos, aprovechar al máximo los recursos como la plataforma SIICEX, y adoptar herramientas y técnicas avanzadas de análisis de datos para obtener información relevante y accionable.

Al abordar esta situación y fortalecer la comprensión del negocio a través de la analítica de datos, Italian Shipping Cargo podrá identificar nuevas oportunidades de crecimiento, mejorar la toma de decisiones comerciales y maximizar su potencial en el mercado de exportaciones e importaciones.

8.1. Aplicación Canvas

La aplicación del Business Model Canvas en Italian Shipping Cargo se justifica especialmente en el punto de conocimiento del negocio. Dado las dificultades en la gestión de datos y la falta de un enfoque estratégico en la toma de decisiones comerciales, el Canvas nos brinda una estructura visual y analítica para comprender y evaluar el conocimiento profundo del negocio.

8.1.1. Problema

- No hay trazabilidad en la recolección y seguimiento de datos del área comercial con respecto a los clientes y operaciones

- No se utilizan los recursos disponibles para análisis de datos para toma de decisiones

8.2.Solución

- Creación del funnel de ventas para establecer bases de datos estructuradas en la gestión interna de la empresa
- Implementar un sistema de gobernanza y analítica de datos
- Clusterización de clientes por medio del método K-means con el fin de aumentar el porcentaje de conversión de cotizaciones a ventas

8.3. Bases

- Bases internas: clientes y operaciones
- Bases externas: SIICEX (DIAN)

8.4.Hipótesis

¿Si implementamos una estructura de recolección y seguimiento de datos del área comercial se podrá evaluar con mayor facilidad el performance del área?

¿Utilizando los recursos disponibles y aplicando la clusterización de clientes mediante el método K-means, aumentaremos el % de conversión de cotizaciones a operaciones?

8.5. Actores

- Patrocinador: Gerencia y junta directiva
- Consumidor final del producto: Gerencia y área comercial
- Impacto del producto: Clientes y Gerencia comercial

8.6. Acciones

- Gestión y control del funnel de ventas mediante las bases de datos estructuradas en el área comercial.
- Actualización del modelo kmeans con las nuevas operaciones de los clientes
- Actualización diaria de los dashboards para evidenciar el resultado de los KPIs

8.7.KPIs

- Lograr una mayor satisfacción del cliente y una mayor retención de estos.
- Lograr una mayor fidelización y generación de nuevas oportunidades de negocio
- Obtener una posición de liderazgo en el mercado nacional e internacional de operadores logísticos
- generar una segmentación de datos de los clientes y optimizar la toma de decisiones en el proceso de importación

8.8.Propuesta de valor

- El problema es de alto nivel en el área comercial ya que no se tiene control de resultados
- Control del performance del área comercial
- Aumentar el número de operaciones

8.9.Riesgos

- No gestionar debidamente el proceso de analítica en el área comercial y generar información de mala calidad
- Falta de adopción por parte de los empleados

- Falta de recursos y capacidades

8.10. Impacto

- Control de los resultados del performance comercial
- Medir número de operaciones por cada asesor y cuáles son los clientes que más retiene
- Medición del comportamiento de ventas antes y después de la implementación del producto

9. Descripción de las fuentes de información

1. ARCHIVO: Celdas de clientes potenciales y prospectos.
 - A. Clientes de Tecnología – es un nicho de mercado que durante la pandemia se incrementó el volumen de carga.
 - B. Con esta información se realiza un filtro por medio de Google, para buscar la empresa y realizar el contacto – con la persona de Importaciones
 - C. En aplicación de Rúes Registro mercantil – se coloca el nombre de la empresa y se verifica su cámara de comercio si está vigente
 - D. Control del contacto con el cliente y que tipo de producto o materia importa
 - E. Registro de clientes con tarifa especial de acuerdo con el volumen de carga se trabaja con ZIX CORP – es el FORWARDERS – Aéreo / Marítimo de la empresa
 - F. Proceso de Italian Shipping de visitas y control de llamadas de clientes para ingresar a la CIA
 - G. Preliquidación esto es para hacer una proyección de cuanto es el costo de la operación
2. ARCHIVO: Proceso para cotizaciones ya formulado para realizar dicha operación
3. ARCHIVO: Información de cotizaciones
 - H. Seguimiento de la cotización si es aprobada o no – control estadístico
 - I. Volúmenes de carga y orígenes por cada cliente

10. Comprensión de los datos.

La etapa inicial de "Comprensión de los datos" en el de acuerdo a la metodología CRISP se centra en la recopilación y análisis exhaustivo de los datos pertinentes para el proyecto de Italian Shipping Cargo. Este proceso implica la exploración y evaluación de datos internos, como registros de ventas y datos de operaciones comerciales, así como datos externos provenientes de fuentes como SIICEX Colombia. A través de este análisis detallado, se busca obtener una comprensión profunda de la calidad, disponibilidad y relevancia de los datos, sentando así las bases para las siguientes etapas de modelado y evaluación.

Las técnicas que se utilizaron para realizar un análisis exhaustivo sobre el manejo y el estado de los datos internos y externos en la empresa son:

10.1. Análisis de tipos de datos estructurados:

- Datos estructurados

Base de datos de clientes de diferentes meses y años

Fuente: Bases de datos de la compañía, elaboración empleados internos.

Variables:

Clientes: Registros de los clientes de los últimos 5 años sin embargo no se tiene trazabilidad de todas las operaciones

Contacto: Es la información del representante legal o el asesor de la empresa que genera la operación con Italian shipping cargo

Telefono: Es el numero de celular o telefono de la empresa que necesita el servicio de importacion, no se tiene bien establecida la nomenclatura

Correo: Correo electronico de la persona que contacta a italian shipping cargo para el servicio de importacion sin embargo no se tiene bien clasificado si es el correo de un asesor o del representante legal

Base de datos de operaciones (Ventas)

Fuente: Bases de datos de la compañía, elaboración empleados internos.

Variables:

Correo: Correo electronico de la persona que contacta a italian shipping cargo para el servicio de importacion sin embargo no se tiene bien clasificado si es el correo de un asesor o del representante legal

Tipo de producto: Según sicex se registra el Des.Cuote el cual describe el tipo de sector de producto que se importa

Region de destino: Se especifica cual es la region de destino que especifica el cliente en la operación

Total Cif: Valor total de la operación por cada cliente en el rango de tiempo establecido

Gráfica 1 Descripción bases de datos SIICEX

1	REPRESENTANTE LEGAL	TOTAL VALOR CIF (US\$)	VIA	REGION, DESTINO	DESC. CUODE
2	RINCON PALLARES LUIS FERNANDO	2,306,78	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
3	RINCON PALLARES LUIS FERNANDO	517,83	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
4	RINCON PALLARES LUIS FERNANDO	909,50	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
5	RINCON PALLARES LUIS FERNANDO	187,10	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
6	VILLEGAS RAMIREZ FABIO	1,219,69	AEREA	CARIBE	MATERIAS PRIMAS Y PDTOS INTERMEDIOS PARA LA IND PDTOS QCOS Y FARMACEUTICOS ELABORADOS
7	VILLEGAS RAMIREZ FABIO	749,07	AEREA	CARIBE	BIENES DE CAPITAL PARA LA INDUSTRIA OTRO EQUIPO FIJO
8	RINCON PALLARES LUIS FERNANDO	954,42	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
9	RINCON PALLARES LUIS FERNANDO	461,86	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
10	RINCON PALLARES LUIS FERNANDO	8,571,35	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
11	RINCON PALLARES LUIS FERNANDO	342,80	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
12	RINCON PALLARES LUIS FERNANDO	62,92	AEREA	CARIBE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
13	RINCON PALLARES LUIS FERNANDO	2,636,25	AEREA	CENTRAL	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
14	RINCON PALLARES LUIS FERNANDO	1,117,98	AEREA	OCCIDENTE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
15	RINCON PALLARES LUIS FERNANDO	155,57	AEREA	OCCIDENTE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
16	RINCON PALLARES LUIS FERNANDO	117,62	AEREA	OCCIDENTE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
17	RINCON PALLARES LUIS FERNANDO	129,81	AEREA	OCCIDENTE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
18	RINCON PALLARES LUIS FERNANDO	878,83	AEREA	OCCIDENTE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
19	RINCON PALLARES LUIS FERNANDO	161,84	AEREA	OCCIDENTE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
20	RINCON PALLARES LUIS FERNANDO	233,89	AEREA	OCCIDENTE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES
21	RINCON PALLARES LUIS FERNANDO	97,79	AEREA	OCCIDENTE	BIENES DE CONSUMO NO DURADERO - VESTUARIO Y OTRAS CONFECCIONES DE TEXTILES

Fuente: Bases de datos de la DIAN, 2022, SICEX

Se muestra una evidencia de las variables que se van a trabajar, son bases de datos de sicex

Variables:

Representante legal: Nombre del cliente el cual general la operación en la empresa

Total valor Cif (US): Valor total de la operación por cada cliente en el rango de tiempo establecido

Via: Tipo de via (Terrestre, Aerea, Maritima) que se utiliza para llevar el tipo de importación a la region de destino

Desc.Cuode: Es el tipo de de sector de producto de importación que realiza el representante legal el cual es el cliente.

10.2. Preparación de los datos

En qué estado están los datos estructurados:

-Clientes: Son las bases de datos más estructuradas en la empresa sin embargo se necesita hacer una limpieza y normalización en estas bases de datos ya que hay métricas y formatos diferentes (DATOS CRUDOS)

-Operaciones: Hace falta recopilar estructuradamente mas datos de las operaciones de la empresa, ya que se encuentran en diferentes cuentas de los comerciales por ende no esta consolidada la informacion en un mismo archivo (DATOS CRUDOS)

- SIICEX: Son las bases de datos principales con las que se va a realizar el proyecto estan limpias y normalizadas sin embargo se debe hacer un analisis de valor en los datos (DATOS ENRIQUECIDOS)

10.3. Limpieza de datos ¿cómo es el entendimiento y el tratamiento de los datos de parte de las personas que los manipulan?

Como se describe en las problemáticas del área comercial, el personal no tiene unanimidad en el momento de almacenar y digitar los datos, se generó un focus group con algunos de los empleados y se evidencia desalineación en el entendimiento de estos, en la forma de consolidarlos y en el valor que se busca encontrar en ellos.

Se realizaron preguntas como:

- ¿Cuáles son las variables de medición más significativas de un cliente? R/ Para algunos es correo electrónico o número de contacto y para otro número de operaciones

- ¿Se tiene una nomenclatura para digitar los datos en la empresa? R/ No en ningún departamento

10.4. Análisis descriptivo de los datos ¿Cantidad de filas y columnas? ¿Qué número de periodos, años y meses tiene? ¿Qué métricas de análisis tiene la empresa?

10.4.1. Clientes:

-Entre 4 y 7 columnas y 480 filas

-Se tienen registro de clientes desde hace 10 años y de ciertos los meses, sin embargo, la información no ha sido aprovechada ya que no está organizada, unas bases están en el RP recurso poco utilizado y otras en las carpetas de la empresa entando en su totalidad desorganizadas

-No tienen métricas establecidas para el análisis de estas bases de datos, el enfoque está en hacer seguimiento a los clientes que efectúan operaciones

10.4.2. Operaciones:

-Entre 8 y 10 columnas y 350 filas

-Se tienen registros aproximadamente de hace 5 años sin embargo no se tiene normalizada la clasificación de los años en algunas bases internas de la empresa

-No se tienen métricas establecidas para el análisis de estas bases de datos, el enfoque está en auditar únicamente en el momento de la operación la trazabilidad del proceso

10.4.3. SIICEX:

-Entre 96 y 150 columnas y más de 2 millones de datos

-Aproximadamente desde 1950 hasta el momento en todos los meses

-Las principales variables que son importantes y se van a tener en cuenta para realizar la segmentación de clientes son: Tipo de vía en la operación, Valor CIF (Venta), Representante legal, Región de destino, Empresa y Desc. Cuode (Descripción del tipo de producto)

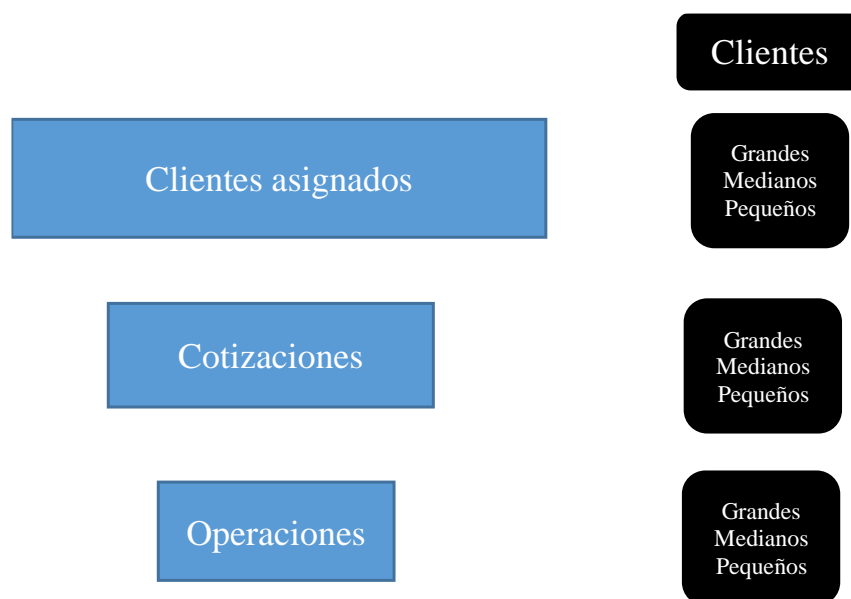
-Las métricas están estandarizadas en darle trazabilidad a las operaciones a nivel mundial

11. Descripción de las alternativas, estrategias y/o acciones que se toman en el análisis de la solución a la problemática

11.1. Estrategias para dar solución a las diferentes problemáticas de la empresa

Solución: Con respecto a la problemática de la mala gestión y análisis de datos internamente en el area comercial se plantea como primera solución un funnel de ventas para la creación de bases de datos que expliquen la trazabilidad y resultados mes a mes del proceso comercial.

Gráfica 2 Explicación funnel de ventas



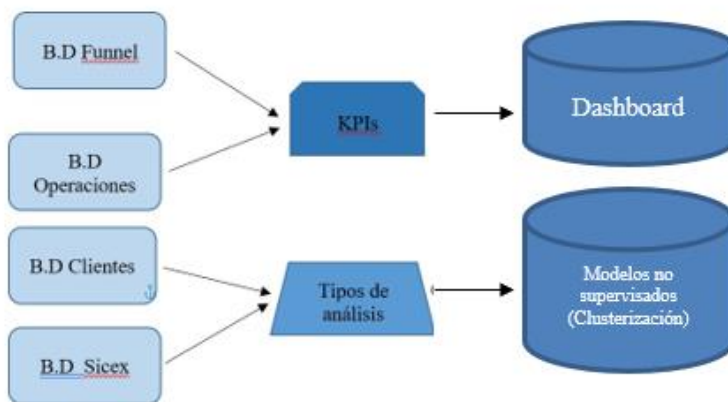
Fuente: Elaboración propia

Este funnel de ventas se gestionará en una base de datos en Excel las columnas son Año mes clientes (Grandes, medianos, pequeños) Clientes asignados, cotizaciones y operaciones

Solución: Al determinar que no se tienen alineados los procesos y objetivos de la parte comercial en la empresa para toma de decisiones se propone como solución la creación de un

dashboard con los KPIs de las bases de datos de clientes, operaciones y funnel para establecer el proceso de manejo de datos mediante un diagrama inteligente

Gráfica 3 Modelo de analítica



Fuente: Elaboración propia

Ante el poco uso de los recursos de datos que tiene la empresa para la toma de decisiones enfocadas al crecimiento en el área comercial, se implementara dentro del alcance establecido un tipo de analítica de negocio sin embargo se propone una segunda, puesto que si es posible realizarla complementaria en buena forma el tipo de analítica implementada dentro del alcance

11.2. Preparación de los datos y proceso de modelamiento a implementar

En la etapa de "Preparación de los datos" dentro de la metodología CRISP, se llevará a cabo un proceso exhaustivo de limpieza, transformación y selección de los datos recopilados. Esto implica identificar y tratar valores atípicos, manejar datos faltantes, normalizar formatos y realizar otras técnicas de preparación de datos necesarias para garantizar la calidad y consistencia de los datos. Además, se realizarán selecciones y filtrados de variables relevantes para el análisis. Esta preparación de los datos sentará las bases para las siguientes etapas de modelado y evaluación, asegurando que los datos estén listos para ser utilizados en los análisis posteriores.

Para este modelamiento se va a utilizar K-means para la clusterización de las bases de datos de SIICEX utilizando únicamente las operaciones que registra la empresa en esta plataforma de la DIAN, se utiliza este tipo de modelamiento para que los asesores comerciales puedan identificar con mayor facilidad las características de cada uno de estos clientes y así puedan abordarlos generando mayor conversión de ventas

11.2.1. Análisis descriptivo no supervisado

Para el análisis descriptivo no supervisado y en concordancia a la aplicación de la metodología CRISP a continuación se inicia con el modelado de los datos para dar solución al problema.

Tarea: Clusterización

Descripción del modelo K-means

11.2.1.1. Introducción

Definición de K-means: K-means es un tipo de análisis no supervisado de aprendizaje automático que cluzteriza o agrupa los datos significativos de los clientes, el objetivo de utilizar kmeans es dividir en cierto conjunto de datos los clientes que tienen similitudes en las variables que se utilizan en el análisis. Además K.-means se enfoca en minimizar las distancias al cuadrado entre cada uno de los datos y los centroides de los clústeres a los cuales pertenece. Este apartado teniendo en cuenta la metodología Crisp, corresponde al modelamiento.

“Los enfoques de clusterización por medio de K- -means son más efectivos. Los clientes tienen diferentes tipos de características y tienen diferente importancia para la empresa.

La clave para una segmentación significativa es definir sus propias variables y atributos de clientes relevantes para el negocio” (Kaur & Singh, 2023. p.3)

11.2.1.2. Necesidad de utilizar K-means

Segmentación de clientes: la clusterización utilizando K-means permite agrupar a los diferentes tipos de clientes que registran operaciones en la empresa en clústeres con respecto al comportamiento de las variables a analizar las cuales son tipos de productos, vías y regiones de importación. Esto facilita a entender a los asesores comerciales cuales son las características principales de los diferentes clientes, permitiendo personalizar las estrategias de ventas y marketing a cada uno de los clientes que se encuentran en los diferentes clusters.

11.2.1.3. Descripción de las variables que se van a utilizar:

Se van a utilizar tres tipos de variables con sus diferentes categorías:

Tipos de productos de importación: La empresa tiene nueve tipos de sectores de productos los cuales son relevantes para clasificar, para identificar cuales son los tipos de clientes que importan estos diferentes productos para anticipar la gestión de compra y logística de la importación de estos y así puedan llegar al destino que el cliente desea en el menor tiempo posible, estos productos son

- BCIMCH: Bienes de capital para la agricultura, maquinas, herramientas y materia de transporte
- BCNDAF: Bienes de consumo no duradero productos alimenticios, farmacéuticos, textiles y bebidas
- CESP: Combustibles elaborados semielaborados y primarios
- ETPR: Equipo de transporte rodante, partes fijas y accesorios

- MPAA: Materias primas y productos intermedios para la agricultura elaborados, semielaborados y primarios
- MPANS: Materias primas y productos intermedios agropecuarios no alimenticios elaborados, primarios y semielaborados.
- MPIA: Materias primas y productos intermedios para la industria productos alimenticios primarios
- MPM: Materia primas y productos intermedios para la industria de productos mineros elaborados, semielaborados, primarios y desechos
- MQF: Materia primas y productos químicos y farmacéuticos

Tipos de vías: Es de gran importancia el análisis de cuáles son los tipos de vías que más utilizan los clientes ya que este proceso en ciertas ocasiones por desconocimiento del comportamiento de los clientes afecta a la empresa ya que no se aprovecha todo el espacio disponible en cada uno de los tipos de vía de importación, generando sobrecostos. Además, ayuda a identificar cuales son los clientes que más utilizan los diferentes tipos de vías para así generar eficiencia en la logística de importación, las vías son:

- AEREA
- TERRESTRE
- MARITIMA

Regiones de destino: La empresa a nivel de logística tiene que tener en cuenta cuales son los diferentes clientes que importan a cierta parte del país, ya que el proceso requiere para mayor eficiencia, anticipación para que los productos le lleguen a los clientes en el menor tiempo posible, si se sabe cuál es la tendencia de los clientes con respecto a las regiones de importación,

la empresa podrá anticipar las peticiones de los clientes y así mejorar la logística para que lleguen los tipos de bienes en el menor tiempo posibles, estas regiones son:

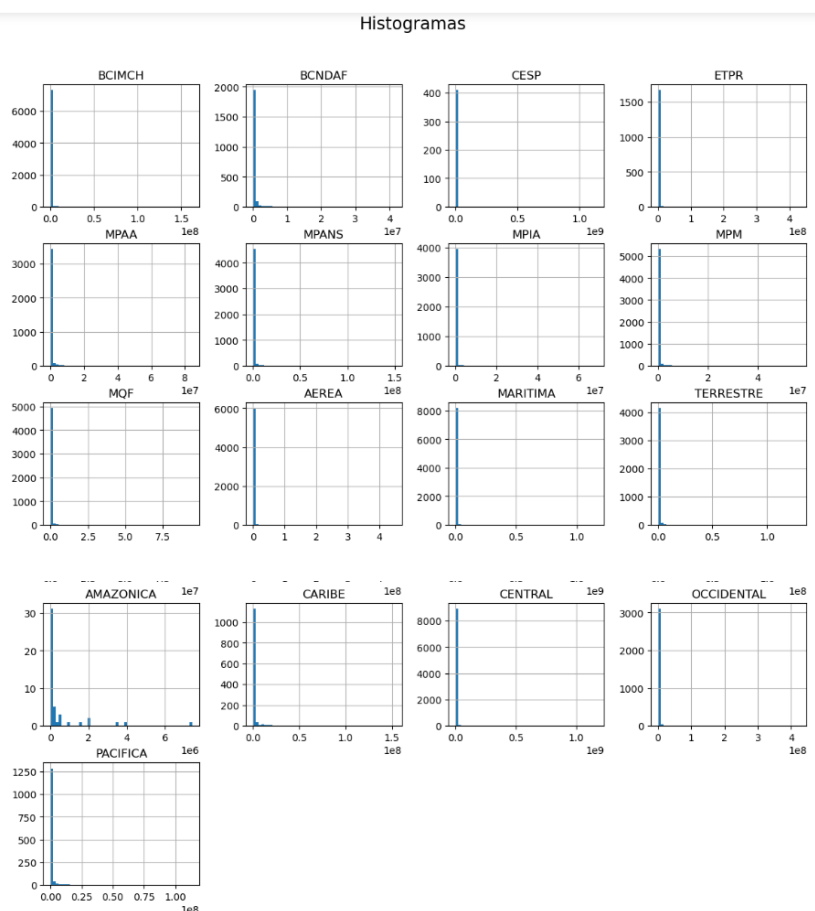
- AMAZONICA
- CENTRAL
- OCCIDENTE
- PACIFICA
- CARIBE

La clave de tener en cuenta estas variables es la personalización del servicio de la empresa ante los clientes de parte de la gestión del área comercial, además enfoca y disminuye el ruido con respecto a cuáles son las variables más importantes que la empresa debe recopilar y generar seguimiento día a día. Identificando los diferentes clusters y sus características la empresa Italia Shipping cargo podrá adaptar y ser más eficiente con respecto a los servicios que ofrece.

11.2.2. Normalización y escalación de los datos:

Inicialmente se identifica por medio de histogramas el comportamiento de la distribución de cada una de las variables para evidenciar si hay sesgos.

Gráfica 4 Análisis de histogramas



Fuente: elaboración propia

Los histogramas de las diferentes columnas evidencian sesgos en la distribución de los datos. Porque tienden a estar a la izquierda y después de los bigotes de la derecha hay bastantes outliers. Además, muestran la frecuencia de la ocurrencia de los datos, en diferentes intervalos, esto afecta patrón de distribución.

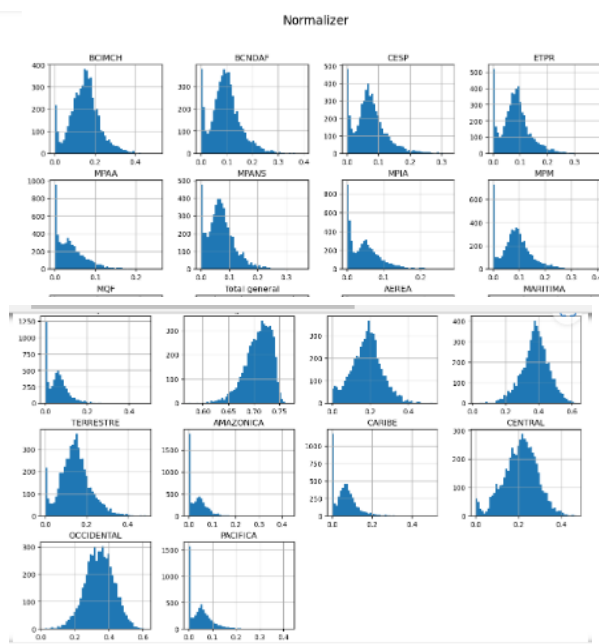
Concluimos con respecto a la distribución de los datos que hay columnas sesgadas a la izquierda porque su distribución únicamente tiende hacia a la izquierda, tiene una curtosis mayor a 3.

Vamos a utilizar algunos métodos de imputación y normalización de los datos para métodos de escalación en un rango específico para así tener distribución normal. Además, esto nos ayuda a eliminar los outliers los cuales son datos que van a afectar negativamente al modelo K-means

Se van a utilizar las técnicas de normalizer y Power Transformer para transformar las variables y generar un rango específico con respecto a la distribución, esto proporciona una influencia equilibrada en el algoritmo de aprendizaje de kmeans, además evita que las diferencias en la magnitud de los valores afecten al modelo en su rendimiento.

Normalized: Se evidencian el cambio de los resultados en los histogramas de las diferentes variables ya que normalizer se enfoca en la escala de los datos en cuanto a su enfoque objetivo en ajustar los datos en una escala específica

Gráfica 5 distribución de datos



Fuente: elaboración propia

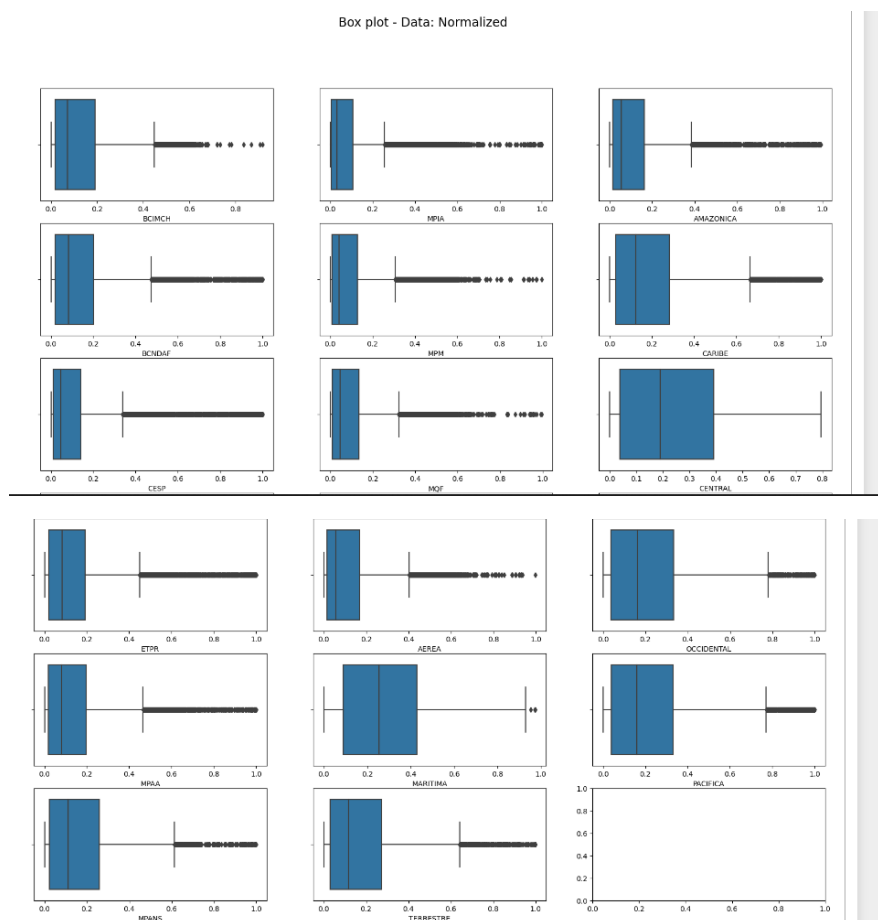
Podemos evidenciar el cambio de las distribuciones al utilizar normalizer, el sesgo que tenían los datos hacia la izquierda se soluciona y esto nos ayuda a que los datos tengan una distribución más gaussiana esta representación genera un conjunto de datos los cuales se reparten en valores bajos, medios y altos, esto genera en el gráfico de histogramas un gráfico con forma de campana y simetría, el punto máximo de los datos representa el promedio

Se puede observar que la distribución es simétrica y gaussiana a diferencia de los primeros resultados los cuales mostraban una distribución sesgada hacia la izquierda

Posteriormente realizamos nuevamente los gráficos de boxplot con power transformer y normalized para visualizar el nuevo comportamiento de la distribución de los datos t sus valores atípicos de cada una de las variables

Boxplot con normalized: Inicialmente vemos que los datos están distribuidos en rangos más normales esta información nos ayuda a empezar a entender las relaciones de las variables, podemos evidenciar que la mediana la cual es la línea dentro del rectángulo en las variables de los tipos de productos (BCIMCH, BCNDAF, CESP, ETPR, MPAA, MPANS, MPIA, MPM y MQF) tienda a estar más a la izquierda, esto significa que sigue habiendo un sesgo ya que la base que se ha utilizado no ha estado estructurada correctamente por parte del área comercial de la empresa, además podemos evidenciar con respecto a los cuartiles la longitud del rectángulo de los tipos de productos que no hay tanta dispersión ya que están entre un rango de 0 y 0,2 a diferencia de MPANS que está en un rango mayor de 0,2 indicando que es un tipo de producto con más dispersión.

Gráfica 6 Variabilidad de los datos



Fuente: elaboración propia

Con respecto a las regiones de destino (CENTRAL, CARIBE, OCCIDENTE, AMAZONICO Y PACIFICA) podemos evidenciar en el box plot que tienen datos con más dispersión ya que están en un rango entre 0 y 0,4 a diferencia de AMAZONICA que es una región con pocas operaciones

Con base a los resultados podemos evidenciar que los tipos de vía (AEREA,

MARITIMA Y TERRESTRE) tienen comportamientos diferentes entre cada una de las variables, marítima tiene mayor dispersión de los 3 tipos de vía ya que está en un rango mayor a 0,4 concluyendo que tiene mayor cantidad y dispersión en los datos, además es la variable de los tres tipos de vía que tiene menos outliers como se evidencia en los puntos a la derecha de los bigotes. La vía terrestre está en un rango entre 0 y 0,3 es el segundo tipo de vía con mayor dispersión y cantidad de datos sin embargo muestra bastantes outliers en comparación a la vía marítima. Por último, evidenciamos que los resultados de la vía aérea tienen una dispersión reducida ya que el rango de la longitud del rectángulo es menor que 0,2 evidenciando que es una variable con poca cantidad y dispersión en los datos, además en comparación a los otros tipos de vía tiene bastantes outliers.

En general con respecto a los bigotes los cuales representan la variabilidad de los datos se puede observar que hay datos de un valor alto que son atípicos ya que son mayores a 0,2 y 0,4

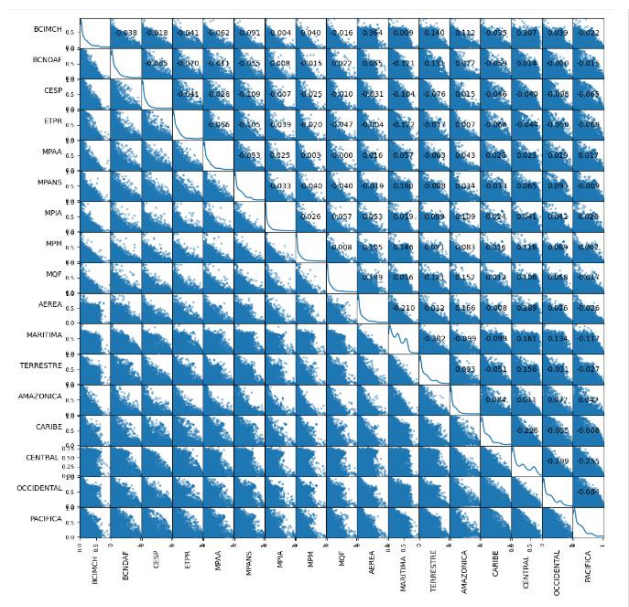
Fuente: Elaboración propia

11.2.2.1. Correlaciones:

El análisis de las correlaciones de las diferentes variables que se van a utilizar en la clusterización con kmeans genera información de alto valor sobre la relación entre las variables, esto ayuda a comprender la estructura que tienen los datos.

“La correlación alta entre las variables numéricas puede afectar la estabilidad de los centroides en el algoritmo de K-Medias ya que conocer lo que pasa con una permite saber lo que pasa con la otra siendo un problema de colinealidad” (Perez & Galvis, 2021)

Gráfica 7 Correlación scatter_matrix



Se utiliza la función `scatter_matrix` para la creación de una matriz de gráficos de dispersión y agregar las anotaciones de correlaciones entre las diferentes variables

Fuente: elaboración propia

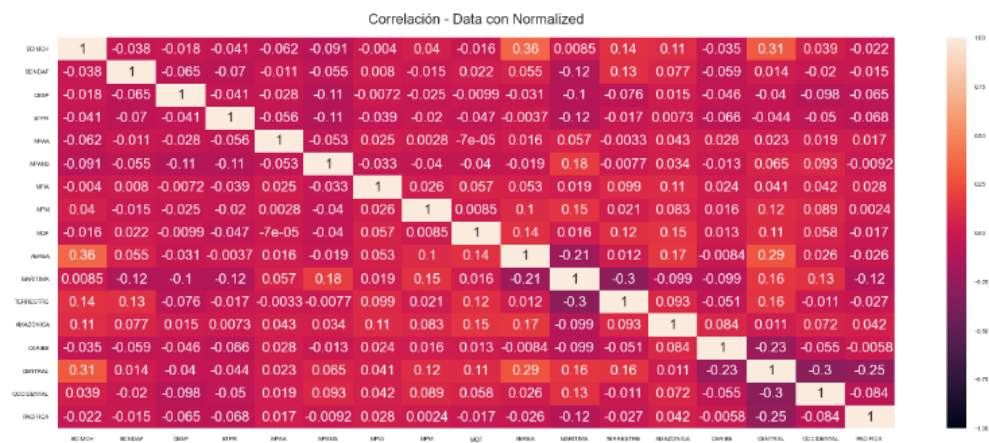
Podemos evidenciar con la siguiente imagen que el patrón de los datos con respecto a su distribución no es netamente positivo ya que no se distribuyen de manera lineal, además con las anotaciones de correlación podemos identificar que la mayoría de las variables no están cercanas de uno por lo cual no hay una correlación fuerte. Sin embargo, generaremos un mapa de calor de las correlaciones entre las variables para evaluar con mejor precisión los patrones de correlación.

Podemos evidenciar que los valores más altos de correlación están entre las siguientes variables:

- BCIMCH y AEREA con 0,36

- BCIMCH y CENTRAL con 0,31
- AEREA y CENTRAL

Gráfica 8 Mapa de calor



Fuente: elaboración propia

La siguiente imagen se toma de un informe científico del uso adecuado del coeficiente de correlación. Esta imagen explica con respecto al coeficiente que se generó en el mapa de calor, cual es la debida interpretación de las correlaciones entre las variables con respecto a su coeficiente.

Tabla 1 Valores de correlación

Rango de valores de r_{XY}	Interpretación
$0.00 \leq r_{XY} < 0.10$	Correlación nula
$0.10 \leq r_{XY} < 0.30$	Correlación débil
$0.30 \leq r_{XY} < 0.50$	Correlación moderada
$0.50 \leq r_{XY} < 1.00$	Correlación fuerte

Fuente: Hernández Lalinde, Espinosa Castro, Rodríguez, & Chacón Rangel, 2018

Se demuestra que las variables en promedio tienen una correlación débil y nula, a diferencia de las variables con mayor correlación antes expuestas que entrarían en una

correlación moderada, esto significa que las correlaciones entre las variables no van a afectar el resultado de kmeans ya que en su mayoría el coeficiente de correlación está por debajo de 0,30.

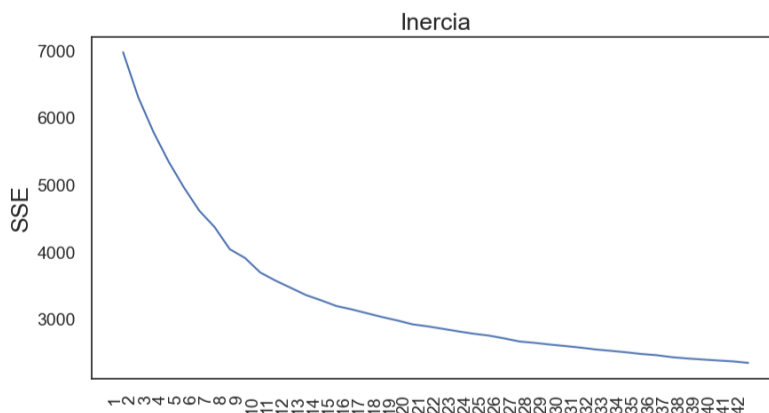
11.2.3. Desarrollo de K-means

Para determinar la cantidad de clusters que se desprenden del método K-means se utilizarán los criterios del método del codo y el de la silueta para respaldar cual va a ser el número de clusters ideales y cual base en definitiva se tomará para escoger los mejores resultados de K-means, la metodología será igual para los dos tipos de bases con sus respectivos análisis teniendo en cuenta que esta en la etapa de evaluación del modelo con respecto a la metodología Crisp.

11.2.3.1. Método del codo:

Para un modelo no supervisado inicialmente se debe encontrar cual es el número óptimo de clusters o grupos en los datos para segmentar. Este método determina el valor ideal de K por medio del quiebre del codo. “La razón de esto es que encontrar un centroide en medio de los dos grupos reduce la distancia relativa entre los puntos de datos. Por lo tanto, requerimos una fórmula de distancia para calcular el número eficiente de grupos para la segmentación” (Kaur & Singh, 2023)

Con la cita anterior del informe de modelo de segmentación de clientes usando K-means para comercio electrónico respaldamos el porqué de utilizar el método del codo para determinar cuál es el número ideal de clusters según el quiebre de la línea.

Gráfica 9 Método del codo

Fuente: elaboración propia

Podemos ver en la anterior imagen el resultado de cuáles serían los número de clusters ideales con respecto a quiebre de la línea, determinamos que 6, 8 y 11 con respecto a la imagen es el número ideal de segmentos sin embargo utilizaremos otro método para corroborar cual es el número ideal de clusters con la base normalized

11.2.3.2. *Método de la silueta:*

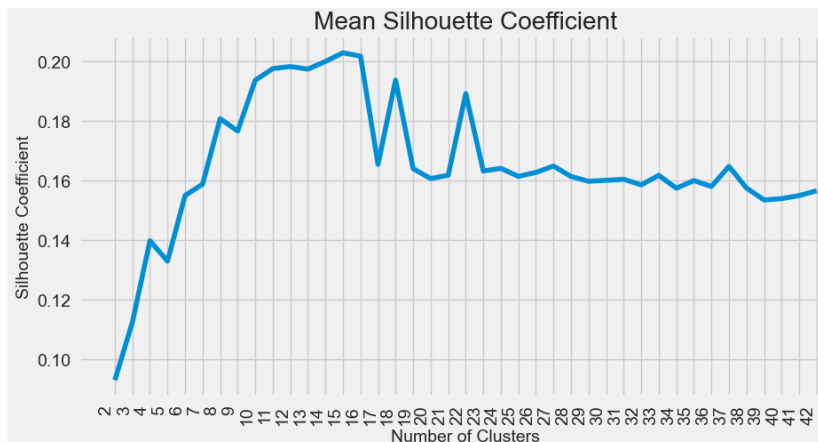
Este método establece para cada uno de los conjuntos de datos una medida de similitud entre cada dato individual y los demás datos que están dentro del cluster para así generar una comparación con el comportamiento de otros datos en clusters diferentes

Índice de silueta: Se utiliza para indicar la calidad de agrupación “Un valor de índice de silueta cerca a cero denota que la observación es cercana a un cluster vecino, lo que indica que no es confiable el número de clusters, un índice cercano o mayor a 1 significa que está bien integrado el cluster y está lejos del resto de grupos, es confiable el número de clusters, y un valor próximo a -1 indica que los clusters no están asignados correctamente” (Perucha & Carlos, 2022)

Con la cita anterior del proyecto de la explicación del método k medias de la

universidad de Valladolid identificamos como entender los resultados del índice de silueta para tomar la decisión más óptima en el número de clusters.

Gráfica 10 Índice de la silueta



Fuente: elaboración propia

Podemos identificar que después de 7 clusters, el índice de silueta es bueno ya que es más alto que 0,1, sin embargo, para generar relación con los resultados del código el cual fueron 6, 8 y 11 clusters ideales podemos concluir que en este resultado de silueta los clusters que cumplen los parámetros ideales son 8 y 11.

11.2.3.3. *Coefficiente de silueta:*

Mide cual es la calidad de la asignación de los elementos y datos de los clusters, se comparan las distancias que hay entre los diferentes datos de un mismo cluster para posteriormente comparar con datos de otro cluster.

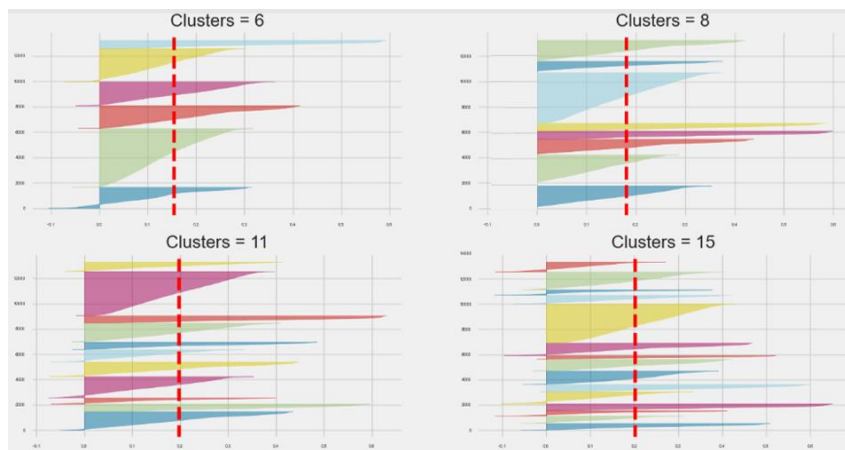
“El coeficiente de silueta del elemento i se denota $s(i)$.

- Si $s(i) = -1$, el dato i está mal agrupado
- Si $s(i) > -1$ y $(i) < 0$, el dato i está sesgado y sus distancias no son ideales
- Si $s(i) \approx 0$, el dato i está entre dos clusters

- Si $s(i) > 0$ y $(i) < 1$, si el dato i es mayor a dos cluster, es ideal la agrupación
- Si $s(i) \approx 1$, el dato i en la totalidad de clusters está bien agrupado “

(Ochoa, Paredes, & Esquicha, 2017)

Gráfica 11 Prueba de coeficiente de silueta



Fuente: elaboración propia

Como podemos observar en la anterior imagen se hizo la prueba del coeficiente de silueta para 6, 8, 11 y 15 clusters, estos muestran sus distribuciones y el comportamiento de los datos entre -1, 0 y 1, en el eje X evidenciamos los resultados del coeficiente de silueta para cada cluster. Además, generamos una línea vertical para comprobar que el volumen de datos de los cluster esté más allá de esta línea, que indica que tienen el mínimo coeficiente ideal, cada uno de los clusters para kmeans. Concluimos que el cluster 8 tiene el mejor coeficiente de silueta ya que esta entre 0,3 y 0,6 al igual que el cluster 15 entre 0,2 y 0,7, sin embargo, el 15 tiene ciertos datos de los diferentes clusters que están entre -1 y 0 como se evidencia en la anterior imagen.

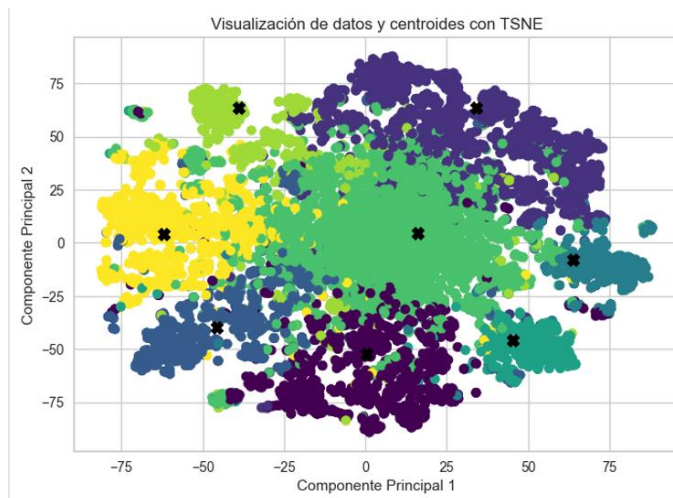
Podemos concluir que para la base normalized teniendo los dos tipos de análisis de

codigo y silueta con sus índices y coeficientes el cluster que mejor se comporta es el cluster numero 8 debido a su similitud de resultados en los dos tipos de análisis anteriormente expuestos, por ende, realizaremos graficas para visualizar los datos después de realizar el Kmean para esta base.

11.2.3.4. Gráfico *t_SNE 2D*:

Este tipo de grafico se utiliza para reducir la dimensionalidad de los datos ya que son bastantes en un rango extenso, y así poder ver su comportamiento mediante etiquetas de colores y los centroides de cada cluster.

Gráfica 12 Gráfico *t_SNE 2D*



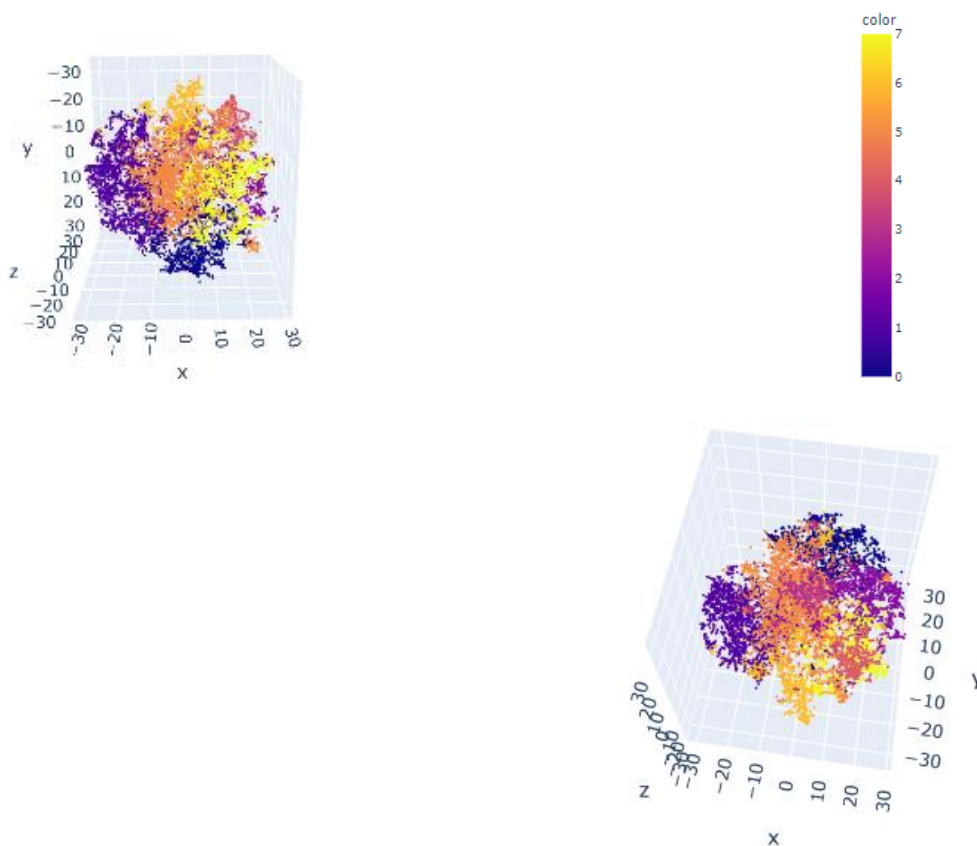
Fuente: elaboración propia

Escogiendo dos componentes de la totalidad de variables que se utilizaron para Kmeans, podemos evidenciar que hay un comportamiento cercano en cada uno de los datos

etiquetados con colores diferentes, además los centroides que son las x muestran que están bien alineados con base al cumulo total de datos de cada cluster

Gráfico 3D: Se utiliza nuevamente t-SNE para reducir la dimensionalidad de los datos en un espacio de tres dimensiones, se busca evidenciar si el comportamiento es similar al grafico 2D.

Gráfica 13 Gráfica 3D



Fuente: elaboración propia

11.2.3.5. *Conclusión*

La interpretación de los resultados de las imágenes anteriores muestra coherencia con

el gráfico generado en 3D, el cual en este caso es con 3 variables principales de la totalidad que se utilizaron en el modelo, podemos ver patrones de datos gaussianos ya que muestra una forma circular, esférica y condensada en cada cluster por medio de las etiquetas de colores de cada uno.

Podemos concluir que posteriormente a los métodos utilizados de evaluación y los resultados que se evidencian en las gráficas se puede entrenar el modelo K-means

“Una vez asignamos puntos a un centro de cluster en particular, dado que el criterio elegido ha sido aquel de minimizar la distancia euclídea al cuadrado y sabemos que los subconjuntos de la forma $\{x \in \mathbb{R}^d / \|x - C\|^2 \leq R^2\}$ son bolas en el espacio, la apariencia de los cluster de datos que forma k medias debe ser de apariencia circulares o esférica” (Perucha & Carlos, 2022, p. 76)

Con la anterior cita corroboramos que el comportamiento y resultado de los datos con la base `normalized` es ideal para K-means. Esto nos aproxima al resultado que queremos generar y es categorizar idealmente a los clientes con sus principales características para que los comerciales puedan abordarlos de una manera más fácil y se puedan generar más ventas.

11.2.4. Análisis e interpretación de los clusters

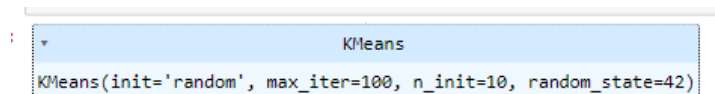
Con el fin de desarrollar evaluación de los resultados dentro de la metodología CRISP, se analizarán y evaluarán los resultados obtenidos a través del modelado de los datos. Se compararán los resultados con los objetivos y criterios establecidos durante la etapa de comprensión del negocio para determinar su validez y utilidad.

Con respecto a los resultados de los cluster definimos que el numero ideal de cluster es $K = 8$ utilizando la base normalized, en este apartado entrenaremos K-means con $K = 8$ y describiremos cuales son las características de cada cluster con los descriptivos.

Los descriptivos son el valor agregado que tendrá el área comercial para adecuar fácilmente su hoja de ruta de ventas a cada uno de los clientes de los diferentes clusters, identificando cuales son los puntos sensibles que busca y son necesidad para los clientes para así suplirlos por medio del servicio de la empresa importadora.

11.2.5. Entrenamiento del modelo Kmeans

Gráfica 14 Formula k-means



```
KMeans
KMeans(init='random', max_iter=100, n_init=10, random_state=42)
```

Fuente: elaboración propia

- Utilizamos la biblioteca **scikit-learn** para utilizar Kmeans configurando los parámetros:
- **I = 8** (número de cluster)
- **Init** para especificar los centroides con la función random
- **N_init** lo utilizamos para establecer que el numero 10 serán las veces que se ejecutara el algoritmo con sus centroides iniciales
- Establecemos un máximo de 100 iteraciones
- Utilizamos **fit** para calcular los centroides y que asigne una muestra de los clusters basados en las distancias de los diferentes centroides

11.2.6. Descriptivos y características de los diferentes clusters

Se aclara inicialmente que en Python se toma de 0 a 7 la cuenta de los 8 clusters

Cluster 1

Tabla 2 Características clúster 1

```

LABEL = 0, cantidad de clientes = 1765
=====
count    BCIMCH    BCNDAF    CESP     ETPR     MPAA     MPANS \
mean    439178.7  292619.4  270580.1  766446.1  607365.7  799392.2
std     3209772.0  936996.5  860831.3  4088947.9  2986272.5  2104800.9
min      0.1        0.4        62.4      8.1        6.8        4.0
25%     5667.1    5504.6    871.4     3433.0    2824.5     7129.3
50%     21983.1   23138.5   7804.4    18463.7   28419.4    71586.8
75%     120261.0  110333.5  118606.8  180227.8  200030.1   484866.7
max     85632474.4  7708090.0  4820975.6  46964481.8  49946043.5  17887089.7

count    MPIA      MPM      MQF      AAREA    MARITIMA  TERRESTRE \
mean    526385.8  477523.0  326783.8  157465.7  2370654.7  333542.4
std     3784373.7  2033092.9  1747609.9  778509.0  8519710.3  1266892.5
min      1.4        0.1        0.1      180.0     719.1      1.1
25%     1199.3    4357.0    3438.4    5414.2    83923.5    9636.1
50%     7611.1    23692.0  19991.1   15242.7   346176.0   39658.3
75%     53684.5   134128.5  105277.6  62893.5   1321569.8  167226.2
max     56654511.9  26628069.2  36993858.1  13975395.5  142130615.0  23464216.8

count    AMAZONICA  CARIBE    CENTRAL  OCCIDENTAL  PACIFICA  labels
mean    728941.6  234614.0  197684.8  2823777.1  343335.0  0.0
std     1282661.9  671312.1  1939420.8  8873184.7  972001.7  0.0
min      543.0      1.2        1.5      402.4      2.8      0.0
25%     12356.2   7285.0    5738.0    213416.8  6899.6  0.0
50%     34152.3   27342.3  14752.3   552176.2  33554.9  0.0
75%     856017.0  135478.3  51012.3   1759445.8  178120.8  0.0
max     3983473.5  5468165.6  55995664.0  123528369.5  8300811.1  0.0

```

Fuente: Elaboración propia

Características:

- Tiene una cantidad de 1,765 clientes los cuales pueden ser abordados individualmente por cada uno de los comerciales
- Se identifican que los tipos de producto que más importan y a los cuales debería el comercial enfocarse en ofrecer operación para estos clientes son:
 - MPANS Con un promedio de importación total de 799.392 (US), El 50% de sus clientes tienen en promedio de importaciones de 71.586 (US) y el 75% de clientes tiene un promedio de importación de 484.866 (US)
 - ETPR con un promedio de importación total de 766.446 (US), el 50% de sus

clientes tienen un promedio de importación menor o igual a 18.463 (US) y el 75% de los clientes importa para este tipo de producto un promedio de 180.227 (US)

- MPAA con un promedio de importación total de 607.365 (US), el 50% de sus clientes importan menor o igual a 28.419 (US) y el 75% de los clientes importa para este tipo de producto menor o igual a 200.030 (US)

Se identifican que los tipos de productos que menos importa este cluster y a los cuales debería evitar el comercial recomendar al cliente son:

- CESP Con un promedio de importación total de 270.580 (US),
- BCNDAF con un promedio de importación total de 292.619 (US)
- Se identifica que la vía que más utilizan estos clientes es la MARITIMA con un promedio de venta de 2.370.654 (US), el 50% de los clientes de este cluster importan menor o igual a 346.176 (US) y el 75% de los clientes importa por esta vía menor o igual a 1.321.569 (US)
- Se identifica que la vía que menos utilizan los clientes en este cluster es la vía AEREA con un promedio de importación de 157.465 (US)
- Las regiones de destino que más utilizan los clientes de este cluster son:
 - OCCIDENTAL con un promedio de importación total de 2.823.777 (US), el 50% de sus clientes importan menor o igual a 552.176 (US) y el 75% de los clientes importa para este tipo de producto menor o igual a 1.759.445 (US)
 - CENTRAL con un promedio de importación total de 197.684 (US) es un promedio menor al de la región caribe sin embargo en esta región hay 966 clientes que importan a la región central, a diferencia de 110 que importan a la región

caribe, el 50% de sus clientes importan menor o igual a 14.752 (US) y el 75% de los clientes importa para este tipo de producto menor o igual a 51.012 (US)

- Las regiones que no son significativas para este cluster son Amazónica con tan solo 12 clientes y Caribe con 110 clientes

Conclusión

El enfoque de los asesores comerciales para este cluster debe estar en los clientes que importan los tipos de bienes MPANS, ETPR y MPAA por medio de la vía MARITIMA hacia las regiones OCCIDENTAL y CENTRAL, para mayor eficiencia es clave que el comercial en el orden como se describió anteriormente establezca los filtros en el dashboard comercial para así identificar cuales son los clientes que cumplen con estas variables las cuales generan mayores ventas.

Cluster 2

Tabla 3 Características clúster 2

```

=====
LABEL = 1, cantidad de clientes = 2443
=====

```

	BCIMCH	BCNDAF	CESP	ETPR	MPAA	MPANS \
count	1736.0	683.0	180.0	518.0	1117.0	1377.0
mean	1148427.8	948837.4	6620966.8	836500.1	580478.3	880292.2
std	7462211.4	3871807.7	85136274.3	4554438.7	3351276.5	3434327.8
min	0.5	2.2	4.3	7.7	0.5	0.2
25%	15515.2	4594.3	832.2	5327.2	2961.1	4842.1
50%	93020.7	39235.2	6964.7	37521.5	27598.2	69425.2
75%	412704.7	283220.4	63304.2	207513.1	168450.0	431331.4
max	162485377.2	41463984.6	1142333164.4	61752483.0	77428451.3	83753047.6

	MPAA	MPM	MQF	AEREA	MARITIMA	TERRESTRE
count	1321.0	1553.0	1508.0	1526.0	2003.0	1238.0
mean	372443.2	466358.8	765748.2	1032119.4	2851823.4	978310.8
std	3024277.3	2429842.6	4780515.9	6124758.5	28267587.9	5077394.2
min	0.1	0.6	0.5	2.3	0.1	0.2
25%	1611.1	6465.1	4933.3	12528.5	96387.3	21285.5
50%	11035.1	36226.6	32876.5	71018.0	331329.6	105194.4
75%	75338.4	175689.3	172315.0	366069.1	1136466.4	412523.7
max	68961195.1	56559694.6	94399477.9	128961598.6	1166359811.2	129714444.3

	AMAZONICA	CARIBE	CENTRAL	OCCIDENTAL	PACIFICA	labels
count	11.0	230.0	2133.0	401.0	174.0	2443.0
mean	1009488.9	918464.3	3702712.6	496282.1	1029420.0	1.0
std	2180963.4	6627062.6	27787904.0	6434449.8	8699216.9	0.0
min	149.3	0.1	14593.3	0.1	0.1	1.0
25%	32124.7	1778.6	235531.5	4356.7	6699.7	1.0
50%	235257.3	22031.0	615642.2	25678.4	35871.1	1.0
75%	775503.0	123313.7	1857771.1	110891.5	136354.1	1.0
max	7413498.0	96123388.2	1173449013.0	128005647.7	113515352.0	1.0

Fuente; elaboración propia

Características:

- Tiene una cantidad de 2443 clientes los cuales pueden ser abordados individualmente por cada uno de los comerciales
- Se identifican que los tipos de producto que más importan los clientes de este cluster y los cuales debería ofrecer el comercial para este nicho son:
 - BCIMCH Con un promedio de importación total de 1.148.427 (US), el 50% de sus clientes tienen en promedio importaciones de 93.020 (US), el 75% de clientes tiene un promedio de importación de 412704 (US)
 - MQF con un promedio de importación total de 765748 (US), el 50% de sus clientes tienen un promedio de importación menor o igual a 32876 (US) y el 75% de los clientes importa para este tipo de producto un promedio de 366069 (US) o menos
 - MPANS con un promedio de importación total de 880.292 (US), el 50% de sus clientes importan menor o igual a 69.425 (US) y el 75% de los clientes importa para este tipo de producto menor o igual a 431.331 (US)
- Se identifican que los tipos de productos que menos importa este cluster y los cuales el comercial no debe ofrecer son:
 - CESP en este caso el promedio de importación no es relevante ya que de la totalidad de clientes de este cluster, únicamente generan operación 180.
 - ETPR en este caso el promedio de importación no es relevante ya que de la totalidad de clientes únicamente generan operación 518
- Se identifica que la vía que más utilizan los clientes de este cluster es la MARITIMA con

un promedio de venta de 2.851.823 (US), el 50% de los clientes de este cluster importan menor o igual a 331.329 (US) y el 75% de los clientes importa por esta vía menor o igual a 1.136.466 (US)

- Se identifica que la vía que menos utilizan los clientes en este cluster es la vía TERRESTRE con un promedio de importación de 978.310 (US)

- Las regiones de destino que más utilizan los clientes de este cluster son:
 - CENTRAL con un promedio de importación total de 3.702.712 (US), el 50% de sus clientes importan menor o igual a 615.642 (US) y el 75% de los clientes importa para este tipo de producto menor o igual a 1.857.771 (US), generan operación 2133 clientes de la totalidad del cluster
 - OCCIDENTAL con un promedio de importación total de 496.282 (US) es un promedio menor al de la región caribe sin embargo en esta región hay 401 clientes que importan a la región occidental a diferencia de 230 que importan a la región caribe, el 50% de sus clientes importan menor o igual a 25.678 (US) y el 75% de los clientes importa para este tipo de producto menor o igual a 110.891 (US)

- Las regiones que no son significativas para este cluster son Amazónica con tan solo 11 clientes y Pacífica con 174 clientes

Conclusión

El enfoque de los asesores comerciales para este cluster debe estar en los clientes que importan los tipos de bienes BCIMCH, MQF y MPANS por medio de la vía MARITIMA hacia las regiones CENTRAL y OCCIDENTAL, para mayor eficiencia es clave que el comercial en el

orden como se describió anteriormente establezca los filtros en el dashboard comercial para así identificar cuáles son los clientes que cumplen con estas variables las cuales generan mayores ventas.

Cluster 3

Tabla 4 Características clúster 3

```

-----
LABEL = 2, cantidad de clientes = 1225
-----

```

	BCIMCH	BCNDAF	CESP	ETPR	MPAA	MPANS \
count	702.0	166.0	41.0	164.0	250.0	348.0
mean	256121.3	268551.7	1939450.7	402773.2	1662551.9	1642134.1
std	1487703.2	1088536.2	6450498.6	1673668.0	8215352.3	10912622.1
min	6.8	0.9	28.8	31.3	11.1	2.2
25%	4857.3	5559.5	1678.4	3537.2	5881.9	3846.3
50%	15189.3	22002.9	19605.8	15120.4	31159.8	20603.9
75%	54852.7	120641.3	155670.8	66871.2	199826.6	195482.1
max	22061237.6	12481550.8	33383049.7	13831179.2	84315375.6	149682734.9

	MPIA	MPH	MPF	AEREA	MARITIMA	TERRESTRE \
count	260.0	417.0	372.0	643.0	467.0	454.0
mean	515163.0	482731.6	285471.4	119346.8	3313574.3	383936.7
std	2865920.3	2484595.1	2075258.7	639965.3	12570571.4	2950674.0
min	0.2	6.6	0.1	2.0	450.0	1.5
25%	1798.5	3140.4	2290.1	5257.6	24453.3	7471.3
50%	10219.6	13501.8	13758.9	15316.8	146602.5	26406.2
75%	38556.6	69014.8	64406.5	40133.8	793187.7	79358.9
max	30573639.3	25158375.4	35372857.0	10542111.2	145359947.4	59823974.3

	AMAZONICA	CARIBE	CENTRAL	OCCIDENTAL	PACIFICA	labels
count	5.0	264.0	818.0	238.0	85.0	1225.0
mean	26517.6	6314256.7	102405.8	138750.2	170990.5	2.0
std	27049.9	17057630.6	580049.6	748981.7	468244.8	0.0
min	1362.0	23421.6	1.1	0.6	1.7	2.0
25%	5652.6	298778.8	6572.4	5555.4	7675.9	2.0
50%	16805.3	759881.2	18344.6	18555.1	30024.3	2.0
75%	44107.2	3517619.9	47137.4	52161.8	82372.3	2.0
max	64661.0	152977794.4	10869284.6	10512370.9	3273006.7	2.0

```

-----

```

Fuente; elaboración propia

Características:

- Tiene una cantidad de 1,225 clientes los cuales pueden ser abordados individualmente por cada uno de los comerciales
- Se identifican que los tipos de producto que más importan y a los cuales debería el comercial enfocarse en ofrecer operación para estos clientes son:
 - BCIMCH Con un promedio de importación total de 256.121 (US) El 50% de sus clientes tienen en promedio importaciones de 71.586 (US) o menos y el 75% de clientes un promedio de importación de 484.866 (US) o menos
 - MPANS con un promedio de importación total de 1.642.134 (US), el 50% de sus clientes tienen un promedio de importación menor o igual a 20.603 (US) y el 75%

de los clientes importa para este tipo de producto un promedio de 195.482 (US)

- MPAA con un promedio de importación total de 1.662.551 (US), el 50% de sus clientes importan menor o igual a 31.159 (US) y el 75% de los clientes importa para este tipo de producto menor o igual a 199.826 (US)
- Se identifican que los tipos de productos que menos importa este clúster y a los cuales no debería abordar el comercial son:
 - BCNDAF ya que únicamente generan operación 166 clientes de los 1225
 - ETPR únicamente generan operación 164 clientes de los 1225
 - Se identifica que hay dos vías que más utilizan estos clientes. AEREA ya que la utilizan 643 clientes y tienen un promedio de venta menor o igual a 119.346 (US), MARITIMA con 467 clientes y un promedio de importación o menor o igual a 793.187 (US)
 - Se identifica que la vía que menos utilizan los clientes en este clúster es la vía TERRESTRE únicamente 454 clientes generan operación
 - Las regiones de destino que generan más valor a la empresa y a las cuales debería abordar los comerciales son:
 - CENTRAL con un promedio de importación total de 102.485 (US) este valor es menor que el de la región caribe, sin embargo, en la región central hay 818 clientes y en la región caribe 264
 - CARIBE con un promedio de importación total de 6.314.256 (US) el 50% de sus clientes importan menor o igual a 759.881 (US) y el 75% de los clientes importa para este tipo de producto menor o igual a 3.517.619 (US)

- Las regiones que no son significativas para este clúster son Amazónica con tan solo 5 clientes y Pacífica con 85 clientes

Conclusión

El enfoque de los asesores comerciales para este cluster debe estar en los clientes que importan los tipos de bienes BCIMCH, MPANS y MPAA por medio de las vías AEREA y MARITIMA hacia las regiones CENTRAL y CARIBE, para mayor eficiencia es clave que el comercial en el orden como se describió anteriormente establezca los filtros en el dashboard comercial para así identificar cuáles son los clientes que cumplen con estas variables las cuales generan mayores ventas.

Cluster 4

Tabla 5 Características clúster 4

```

=====
LABEL = 3, cantidad de clientes = 630
=====
count    BCIMCH  BCNDAF  CESP    ETPR    MPAA    MPANS    MPAA  \
mean    108710.5  69661.7  24769.5  29560722.0  65867.3  90997.8  25101.6
std     321760.0  169579.3  43329.3  85349908.7  198441.8  194141.7  105167.6
min      13.4     7.2     9.0     137425.8    12.0     0.6     1.3
25%     7750.5   899.1   1394.9  1165582.6   798.9   1340.0   266.2
50%    29826.2  12350.9  2734.7  3978903.4   4772.4  14361.5  1386.3
75%    107500.0  36221.0  28235.3  9809150.9  33521.4  64721.3  11804.1
max    6133109.7  1104130.9  136121.4  430452104.6  1902711.2  1163786.6  1191264.3

count     MPM     MQF     AEREA  MARITIMA  TERRESTRE  AMAZONICA  \
mean    38599.7  43810.5  1560064.4  206937.2  1779968.5  NaN
std    317233.4  151450.4  23193124.1  359842.7  9482219.4  NaN
min      1.3     1.2     9.8     97.4     0.8     NaN
25%     653.0   787.6   8655.8  26423.7   8820.9  NaN
50%    1219.9  4023.4  29581.2  76814.8  47722.9  NaN
75%    7761.4  24999.8  92536.5  224463.0  156771.2  NaN
max   6040665.1  2121366.4  447689600.0  3302296.6  80231476.5  NaN

count    CARIBE  CENTRAL  OCCIDENTAL  PACIFICA  labels
mean    76740.1  726484.8  4479585.6  423270.5  3.0
std    152588.1  4875598.6  39571083.8  1421726.4  0.0
min      0.8     295.3     0.4     1270.3  3.0
25%    3564.7  18244.3  11473.4  11566.0  3.0
50%    19845.8  72756.7  37255.0  52303.9  3.0
75%    58119.9  252347.5  117414.3  218932.9  3.0
max    810715.5  71648500.8  424696382.4  10401197.0  3.0

```

Fuente: elaboración propia

Características:

- Tiene una cantidad de 630 clientes los cuales pueden ser abordados individualmente por cada uno de los comerciales
- Se identifican que los tipos de producto que más importan los clientes de este cluster y los que debería ofrecer el comercial son:
 - BCIMCH con 479 clientes que generan operación, tiene un promedio de importación de 108.710 (US)
 - MPM con 479 clientes que generan operación este tipo de producto tiene un promedio de importación de 38.599 (US),
 - MQF con 303 clientes que generan operación este tipo de producto tiene un promedio de importación de 43.810 (US),
 - Los clientes que generan operación con ETPR a pesar de ser 26 tienen un valor de importación alto ya que el promedio de importación es de 29.560.722 (US) el comercial debe abordarlos ya que a pesar de que no sean tantos clientes que importen, el valor de importación es significativa para la empresa
- Se identifican que los tipos de productos que menos importa este clúster y los que no debería ofrecer el comercial son:
 - CESP Con un promedio de importación total de 24.769 (US) con únicamente 17 clientes que generan operación
- Se identifica que la vía que más utilizan estos clientes es la AEREA con un promedio de importación de 1.560.064 (US) y 381

- Se identifica que la vía que menos utilizan los clientes en este clúster es la vía TERRESTRE con un promedio de importación de 157.465 (US) y únicamente 130 clientes generan operación
- Las regiones de destino que más utilizan los clientes de este clúster son:
 - CENTRAL con un promedio de importación total de 726.484 (US) y 443 clientes que generan operación
 - OCCIDENTAL con un promedio de importación total de 726.484 (US) y 121 clientes que generan operación,
- Las regiones que no son significativas para este clúster son Amazónica con 0 clientes y caribe con 47 clientes

Conclusión

El enfoque de los asesores comerciales para este cluster debe estar en los clientes que importan los tipos de bienes BCIMCH, MPM, MQF y CESP por medio de la vía AEREA hacia las regiones CENTRAL y OCCIDENTAL, para mayor eficiencia es clave que el comercial en el orden como se describió anteriormente establezca los filtros en el dashboard comercial, para así identificar cuáles son los clientes que cumplen con estas variables las cuales generan mayores ventas.

Clúster 5

Tabla 6 Características clúster 5

```

=====
LABEL = 4, cantidad de clientes = 609
=====
count  BCIMCH  BCINDAF  CESP  ETPR  MPAA  MPANS  MPIA  \
mean  48993.8  17259.4  114824576.3  29227.5  8266.5  20783.4  128297.4
std   77907.0  35361.7  NaN  40605.0  25136.8  80641.7  1362753.7
min   30.0    15.4    114824576.3  22.7    3.0    0.1    9.3
25%   5632.0  806.7   114824576.3  2550.1  351.7  564.5  927.8
50%   17970.0  3249.0  114824576.3  14360.3  1251.1  2316.0  4020.4
75%   68799.8  10218.5  114824576.3  31671.4  3005.9  13563.0  25992.1
max   737457.0  175082.5  114824576.3  178824.5  148606.8  750117.5  18989059.3

count  MPM  MQF  AEREA  MARITIMA  TERRESTRE  AMAZONICA  \
mean  14315.8  19098.3  56695.5  2874687.1  355229.3  NaN
std   31358.4  29679.7  111384.8  7318695.4  5083754.6  NaN
min   2.5    1.1    2.0    6892.0    44.7    NaN
25%   560.9  1395.9  8269.3  8198.8  12750.8  NaN
50%   2815.2  7948.2  31596.5  8696.0  41604.0  NaN
75%   11608.4  22046.0  78560.8  73282.5  104006.1  NaN
max   287487.0  174855.4  1813946.4  22737808.6  92265364.6  NaN

count  CARIBE  CENTRAL  OCCIDENTAL  PACIFICA  labels
mean  3857302.7  119413.0  69230.7  64884.8  4.0
std   20992122.5  904573.2  69885.2  87405.6  0.0
min   2.0    128.0    1.7  1716.0  4.0
25%   7076.7  14352.4  13434.6  10483.7  4.0
50%   14994.0  39172.2  51646.1  25882.3  4.0
75%   34771.6  115974.6  91742.7  89746.5  4.0
max   115003173.2  18989059.3  358682.7  441982.8  4.0
=====

```

Fuente: elaboración propia

Características:

- Tiene una cantidad de 609 clientes los cuales pueden ser abordados individualmente por cada uno de los comerciales
- Se identifican que los tipos de producto que más importan los clientes de este clúster ya los cuales debería ofrecer el comercial son:
 - BCIMCH Con un promedio de importación total de 48.993 (US) y 428 clientes que generan operación
 - MPM con un promedio de importación de 14.315 (US), 310 clientes generan operación
 - MQF con un promedio de importación de 19.098 (US) y 249 clientes que generan operación
- Se identifican que los tipos de productos que menos importa este clúster y los cuales

generaran menores ingresos son:

- CESP solo un cliente genera operación
- ETPR con un promedio de importación total de 29.227 (US) y únicamente generan operación 78 clientes
- Se identifica que la vía que más utilizan estos clientes es la TERRESTRE con un promedio de venta de 355.229 (US) 329 clientes generan operación por medio de esta vía
- Se identifica que la vía que menos utilizan los clientes en este clúster es la vía MARITIMA con únicamente 14 clientes que generan operación
- Las regiones de destino que más utilizan los clientes de este clúster son:
 - CENTRAL con un promedio de importación total de 119.413 (US) y 443 clientes que generan operación,
 - OCCIDENTAL con un promedio de importación total de 69.230 (US) y 110 clientes que generan operación
- Las regiones que no son significativas para este clúster son Amazónica con 0 clientes y Pacífica con 47 clientes

Conclusión

El enfoque de los asesores comerciales para este cluster debe estar en los clientes que importan los tipos de bienes BCIMCH, MPM y MQF por medio de la vía TERRESTRE hacia las regiones CENTRAL y OCCIDENTAL, para mayor eficiencia es clave que el comercial en el orden como se describió anteriormente establezca los filtros en el dashboard comercial, para así

identificar cuáles son los clientes que cumplen con estas variables las cuales generan mayores ventas.

Cluster 6

Tabla 7 Características clúster 6

```

-----
LABEL = 5, cantidad de clientes = 3983
-----
count      BCIMCH      BCNDAF      CESP      ETPR      MPAA      MPANS \
mean      88276.5    145655.0    205699.5    89954.8    258043.8    214375.0
std       280426.6    543958.8    665148.2    284852.1    930379.0    635655.3
min        2.9        1.8         53.1        5.3        2.8         2.0
25%       9956.1     8350.8     2440.6     4438.9     13134.6     21612.9
50%       29622.0    27705.1    23719.9    20401.5    50004.7     63336.4
75%       74723.8    84157.0    61450.9    70495.2    176480.0    163304.6
max       6336530.3  6681611.9  3735875.8  3932118.7  19629909.0  14896695.3

count      MPIA        MPM        MQF        AEREA    MARITIMA    TERRESTRE \
mean      161612.1    129030.6    77707.0    110517.3    243105.1    190366.1
std       902673.8    556418.2    231100.3    499347.0    780517.5    609233.3
min        0.1         2.1         5.2         4.2         1.3         1.8
25%       2720.0     5301.0     5190.2     10292.6    30532.1    14750.9
50%       16511.1    22497.4    21708.1    26774.2    69044.0    50907.9
75%       61256.0    83209.8    62660.6    71384.9    179463.7    141383.7
max       15126130.9  12834638.9  4043655.3  8570323.1  19629909.0  11522050.8

count      AMAZONICA    CARIBE    CENTRAL    OCCIDENTAL    PACIFICA    labels
mean      166602.5    284194.5    236857.5    212913.8    256885.4    5.0
std       172367.4    867932.2    853459.8    452817.4    538532.0    0.0
min       20000.0     18.2      1.5        0.5         0.3         5.0
25%       32226.8     26873.1    29688.4    31868.1    27600.8    5.0
50%       58126.7     73079.2    66411.8    74413.8    72654.0    5.0
75%       245916.9    174949.6    165464.4    173778.3    214291.3    5.0
max       525313.0    11522050.8  19629909.0  5924001.8  4039051.1    5.0

```

Fuente: elaboración propia

Características:

- Tiene una cantidad de 3983 clientes los cuales pueden ser abordados individualmente por cada uno de los comerciales
- Se identifican que los tipos de producto que más importan los clientes de este clúster y los cuales generarían más ingresos son:
 - BCIMCH, 1510 clientes generan operación, tienen un promedio de importación

total de 88.276 (US),

- MPANS, 1177 clientes generan operación, tienen un promedio de importación total de 214.375 (US),
 - MPM, 1069 clientes generan operación, tienen un promedio de importación total de 129.030 (US),
- Se identifican que los tipos de productos que menos importa este clúster y los que no generan altas ventas son:
 - CESP Con un promedio de importación total de 205.699 (US) es alto este promedio, pero únicamente generan operación 61 clientes a comparación de 318 de ETPR, el cual es el segundo tipo de producto que menos clientes tiene y genera un promedio de importación de 89.954 (US)
 - Se identifica que la vía que más utilizan estos clientes es la MARITIMA con un promedio de venta de 243.105 (US), generan operación 3167 clientes
 - Se identifica que la vía que menos utilizan los clientes en este clúster es la vía TERRESTRE con un promedio de importación de 190.366 (US) y únicamente importan 717
 - Las regiones de destino que más utilizan los clientes de este cluster son:
 - CENTRAL con un promedio de importación total de 236.857 (US) y 2662 clientes que generan operación
 - OCCIDENTAL con un promedio de importación total de 212.913 (US) y 847 clientes que generan operación.

- Las regiones que no son significativas para este clúster son Amazónica con tan solo 11 clientes y caribe con 315 clientes

Conclusión

El enfoque de los asesores comerciales para este cluster debe estar en los clientes que importan los tipos de bienes BCIMCH, MPANS y MPM, por medio de la vía MARITIMA hacia las regiones CENTRAL y OCCIDENTAL, para mayor eficiencia es clave que el comercial en el orden como se describió anteriormente establezca los filtros en el dashboard comercial, para así identificar cuáles son los clientes que cumplen con estas variables las cuales generan mayores ventas.

Cluster 7

Tabla 8 Características clúster 7

	BCIMCH	BCNDAF	CESP	ETPR	MPAA	MPANS	MPIA
count	500.0	104.0	13.0	102.0	193.0	291.0	277.0
mean	38321.8	2048108.0	4648.4	19466.4	33834.3	45088.6	71770.0
std	116779.0	4398438.8	6831.4	27566.4	101043.7	99583.0	424451.1
min	1.4	206.1	93.3	23.5	5.2	1.6	0.8
25%	3691.3	159124.3	340.6	3239.9	1830.6	3166.9	194.0
50%	11543.9	455534.2	1451.2	9029.7	7310.6	12304.6	2622.2
75%	31546.1	1239617.3	4653.6	27961.8	29575.9	50868.0	9933.2
max	2098249.7	24361739.2	21061.5	200453.7	1031417.3	1143632.8	5311487.2

	MPM	MQF	AEREA	MARITIMA	TERRESTRE	AMAZONICA
count	337.0	347.0	338.0	600.0	169.0	3.0
mean	28841.5	42439.0	178829.9	112807.8	1006260.3	11789.4
std	85528.0	151858.8	1146828.3	450778.9	2821032.3	13879.3
min	0.4	3.4	0.8	0.6	414.8	3208.0
25%	1732.1	1538.3	4312.2	12208.6	37602.3	3783.1
50%	6994.4	7785.9	10730.7	32865.6	170681.0	4358.2
75%	24686.8	31544.2	36509.7	78961.6	632274.9	16080.2
max	1061350.9	2186053.3	14832082.4	6022325.7	21480090.8	27802.1

	CARIBE	CENTRAL	OCCIDENTAL	PACIFICA	labels
count	115.0	354.0	314.0	137.0	881.0
mean	68838.0	607041.0	147581.7	211695.1	6.0
std	173085.2	2459715.8	755702.2	1721809.4	0.0
min	6.0	1.6	667.8	428.5	6.0
25%	10870.6	14020.0	11127.6	9174.0	6.0
50%	26000.0	55429.3	36161.2	25307.8	6.0
75%	52587.9	189988.3	91853.7	67466.6	6.0
max	1320371.8	26630552.3	10516559.1	20178505.2	6.0

Fuente: elaboración propia

Características:

- Tiene una cantidad de 881 clientes los cuales pueden ser abordados individualmente por cada uno de los comerciales
- Se identifican que los tipos de producto que más importan los clientes de este clúster y pueden generar altas ventas son:
 - BCIMCH, 500 clientes generan operación, tienen un promedio de importación total de 38.321 (US),
 - MQF, 347 clientes generan operación, tienen un promedio de importación total de 42.439 (US),
 - MPM, 337 clientes generan operación, tienen un promedio de importación total de 28.841 (US).
- Se identifican que los tipos de productos que menos importa este cluster son:
 - CESP con un promedio de importación total de 4.648 (US), y únicamente generan operación 13 clientes
 - ETPR con un promedio de importación total de 9.029 (US), y únicamente generan operación 102 clientes
- Se identifica que la vía que más utilizan estos clientes es la MARITIMA con un promedio de venta de 112.807 (US) y 600 clientes generan operación,
- Se identifica que la vía que menos utilizan los clientes en este cluster es la vía TERRESTRE con un promedio de venta de 1.006.260 (US) y 169 clientes generan

operación

- Las regiones de destino que más utilizan los clientes de este cluster son:
 - CENTRAL con un promedio de importación total de 607.041 (US) y 354 clientes que generan operación
 - OCCIDENTAL con un promedio de importación total de 147.581 (US) y 314 clientes que generan operación
- Las regiones que no son significativas para este clúster son Amazónica con tan solo 3 clientes y caribe con 115 clientes

Conclusión

El enfoque de los asesores comerciales para este cluster debe estar en los clientes que importan los tipos de bienes BCIMCH, MQF y MPM por medio de la vía MARITIMA hacia las regiones CENTRAL y OCCIDENTAL, para mayor eficiencia es clave que el comercial en el orden como se describió anteriormente establezca los filtros en el dashboard comercial, para así identificar cuáles son los clientes que cumplen con estas variables las cuales generan mayores ventas.

Clúster 8

Tabla 9 Características clúster 8

```

=====
LABEL = 7, cantidad de clientes = 1654
=====
count      BCIMCH      BCNDAF      CESP      ETPR      MPAA      MPANS \
mean 158671.7  447093.1  466463.5  202450.7  596481.4  491384.1
std  826375.1  1890718.5  2018197.5  927122.0  2413442.7  1323945.7
min    13.1     17.6       11.3      4.4       1.4       0.9
25%   5596.7   5879.7    1160.9    3051.6    2687.6    5400.7
50%  21568.9  24086.0   7441.4    13089.8   17290.2   29936.5
75%  59931.5  85606.0   45379.1   56422.2   119093.7  243032.8
max 17386996.8 17193541.3 12254458.4 8927043.9 24689994.1 10190876.2

count      MPIA      MPM      MQF      AEREA      MARITIMA      TERRESTRE \
mean 544378.7  289119.6  185312.0  110446.7  1472481.9  295069.2
std  3557539.0  2254225.6  1516213.0  657716.1  5527176.6  1380021.9
min    0.3       1.0       0.6      216.9      5.3       1.6
25%   547.7    2812.8    769.2    5590.8    31554.6    12686.8
50%  4007.7   10962.5   6690.8   13875.7  104445.3   32661.8
75%  32011.7  41840.2  46717.0  40998.4  484093.0   92400.1
max 36196598.3 49035233.1 31350295.6 13518164.2 62500663.1 26961070.5

count      AMAZONICA      CARIBE      CENTRAL      OCCIDENTAL      PACIFICA      labels
mean 688162.2  275620.9  85012.0  184358.0  3389615.4  7.0
std  1522684.3  1026497.4  353712.0  1419631.8  8043180.5  0.0 |
min  1827.5    1.9       0.3      2.9      33307.8    7.0
25%  6131.1    7923.4    5888.0    8976.1    238567.1    7.0
50%  9242.6    27760.4  20330.4  27984.0  667562.7    7.0
75%  11595.0   87929.4  51093.8  78057.4  2256017.9    7.0
max 3412014.8 8980480.7 7656520.6 22455373.8 61753154.3 7.0

```

Fuente: elaboración propia

Características:

- Tiene una cantidad de 1654 clientes los cuales pueden ser abordados individualmente por cada uno de los comerciales
- Se identifican que los tipos de producto que más importan los clientes de este cluster son:
 - BCIMCH, 955 clientes generan operación, tienen un promedio de importación total de 158.671 (US)
 - MPM, 658 clientes generan operación, tienen un promedio de importación total de 289.119 (US).
 - MQF, 604 clientes generan operación, un promedio de importación total de 185.312 (US).
- Se identifican que los tipos de productos que menos importa este cluster y los cuales no

generan altas ventas son:

- CESP con un promedio de importación total de 466.463 (US), y únicamente generan operación 39 clientes
- ETPR con un promedio de importación total de 202.450 (US), y únicamente generan operación 222 clientes
- Se identifica que las dos vías que más utilizan estos clientes es la AEREA y MARITIMA la diferencia está en que MARITIMA tiene más promedio de venta de 1.472.481 (US) y 708 clientes que generan operación y AEREA con un promedio de venta de 110.446 (US) y 905 clientes que generan operación.
- Se identifica que la vía que menos utilizan los clientes en este cluster es la vía TERRESTRE con un promedio de importación de 295.069 (US) Y 551 clientes que generan operación
- Las regiones de destino que más utilizan los clientes de este cluster son:
 - CENTRAL con un promedio de importación total de 85.012 (US) y 1129 clientes que generan operación
 - PACIFICA con un promedio de importación total de 3.389.615 (US) y 332 clientes que generan operación.
- Las regiones que no son significativas para este cluster son Amazónica con tan solo 5 clientes y caribe con 106 clientes

Conclusion

El enfoque de los asesores comerciales para este cluster debe estar en los clientes que importan los tipos de bienes BCIMCH, MQF y MPM por medio de las vías AEREA y

MARITIMA hacia las regiones CENTRAL y PACIFICA, para mayor eficiencia es clave que el comercial en el orden como se describió anteriormente establezca los filtros en el dashboard comercial, para así identificar cuáles son los clientes que cumplen con estas variables las cuales generan mayores ventas.

Conclusion general

Los asesores comerciales por medio del siguiente dashboard comercial que se explicara a continuación, deben aprovechar las características principales evidenciadas en cada cluster para abordar a los clientes, estas características ayudaran a disminuir el tiempo de recompra de los clientes y disminuirá igualmente el tiempo de importación del país que produce el producto hasta el lugar del destino, para así mejorar la logística comercial. También ayuda a la gerencia a identificar cuales son los clientes que los comerciales convierten mas de cotizaciones a ventas para así enfocarlos en estos y generar mayor trazabilidad en el área comercial y aumento en las ventas.

En conclusión las la priorización de productos, vías y regiones maximizara las oportunidades de crecimiento de la empresa y aumentara la satisfacción del cliente

11.3. Dashboards

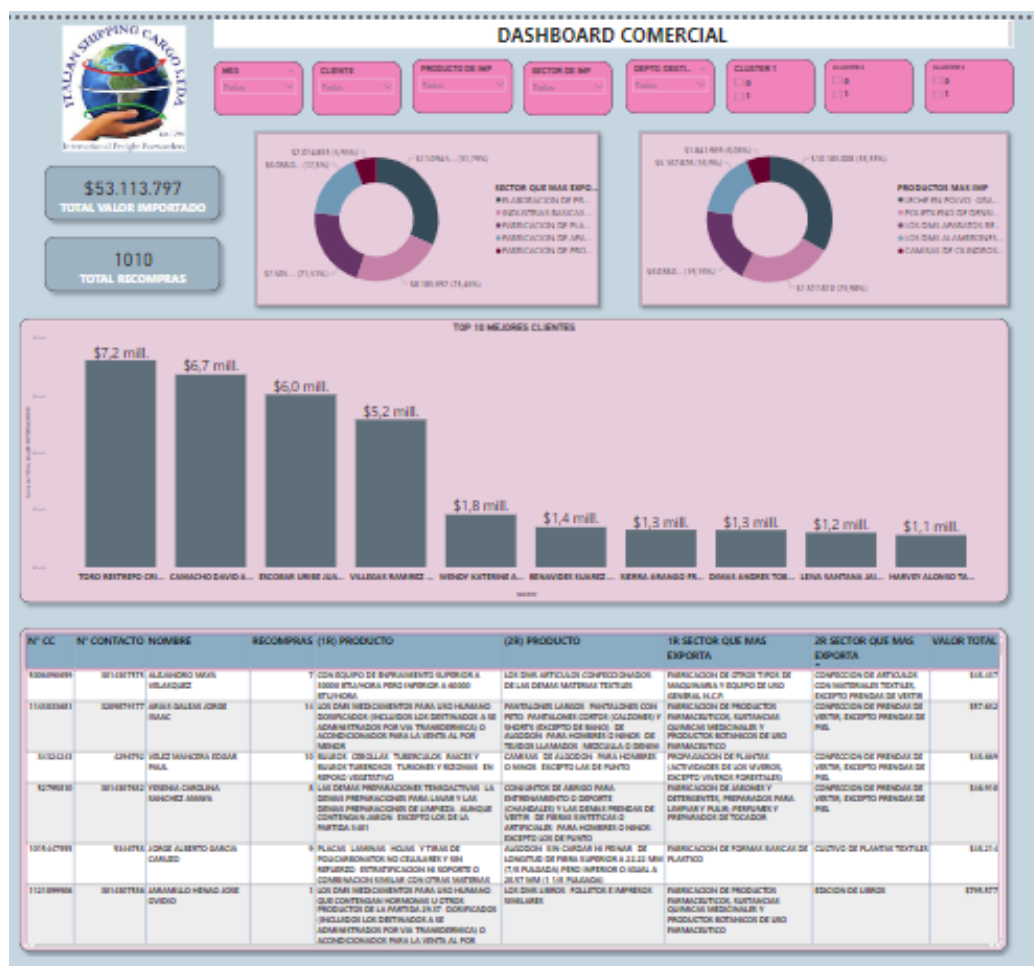
Finalmente, con el fin de desarrollar el despliegue del proyecto, y teniendo en cuenta la metodología CRISP se crearon dos dashboards en Power BI para que los datos del resultado de las características de los cluster sean fáciles de digerir y entender tanto para el área comercial

como para la Gerencia. Además, esta es una herramienta clave para la toma de decisiones informadas en la empresa.

11.3.1. Dashboard Comercial

El propósito de este dashboard está enfocado principalmente en el resultado de la clusterización con kmeans de los clientes, la idea es proporcionar esta información en tiempo real y visualmente atractiva sobre el estado actual de las características de los clientes y las oportunidades comerciales. Este dashboard está enfocado en los asesores comerciales, para que puedan evidenciar claramente cuáles son las características de los clientes en los 8 tipos de clusters y así aumentar el porcentaje de conversión en el funnel de ventas. Además, ayuda a los asesores a anticipar las necesidades de los clientes para si reducir el tiempo de venta. Por último, ayuda al director comercial a ver cómo puede enfocar a cada asesor con base en los resultados de venta con las especificaciones o características de esos clientes

Gráfica 15 Dashboard comercial



Fuente: elaboración propia

11.3.1.1. Filtros específicos

Utilizamos los filtros para:

- Mes
- Cliente
- Producto de importación
- Sector o región de importación
- Los 8 clúster del resultado de K-means (1 significa que pertenece y 0 que no)

pertenece)

Estos filtros interactivos permiten a los comerciales acceder rápidamente a datos relevantes sobre clientes lo que ayuda a tomar decisiones mas fundamentadas en su trabajo y en como abordar al cliente

Gráfica 16 Filtros dashboard comercial

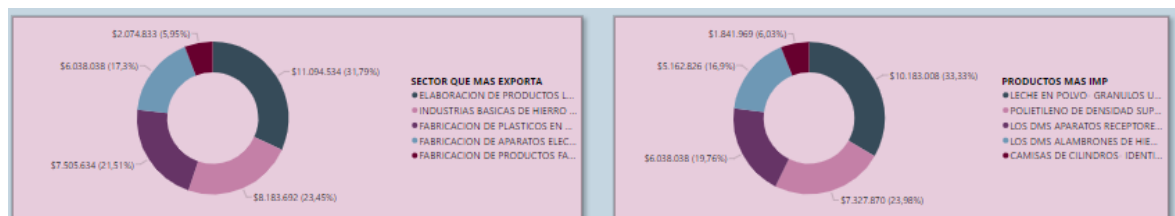


Fuente: elaboración propia

Gráfico de anillos de tipos de productos y sectores de importación

Con estos tipos de tableros el director comercial y los asesores pueden identificar cuáles son los sectores de productos que más están generando ventas en los diferentes meses y clusters, además permite explorar de forma interactiva los porcentajes que cada producto y sector de importación tiene.

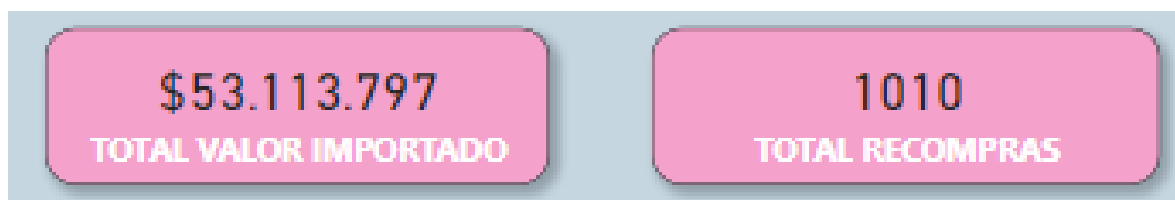
Gráfica 17 Gráfico de anillos dashboard comercial



Fuente: elaboración propia

Etiquetas de datos

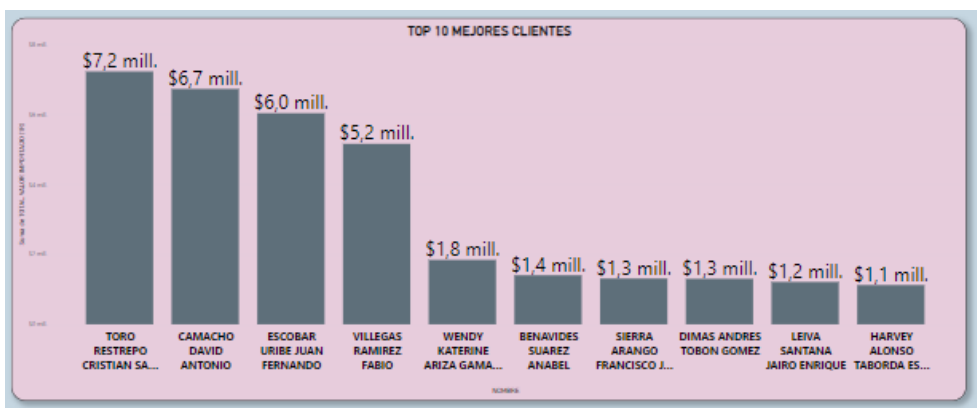
En las etiquetas el área comercial podrá visualizar fácilmente el valor del total importado y el número de recompras de cada uno de los clientes a partir de su primera compra, esto ayudara al aérea comercial a mapear de una manera más fácil cuales clientes están llevando utilidad a la empresa

Gráfica 18 Referencias de datos

Fuente: elaboración propia

11.3.1.2. *Gráfico de barras de los 10 mejores clientes*

En coherencia con las anteriores partes del dashboard y para que la historia de los datos sea más fácil de digerir, se identifica cuáles son los 10 primeros clientes y se especifica el monto de importación con base a la interacción con cada uno de los filtros para que los asesores comerciales los tengan en cuenta a la hora de hacer activación de compra.

Gráfica 19 Grafico de barras

Fuente: elaboración propia

11.3.1.3. *Tabla de clientes*

En esta tabla estará la información principal de los clientes con respecto a cuáles son los dos sectores de tipos de productos que más importa al igual que la información de los dos productos que más importa, el número de contacto, número de recompras y total importado. Es

clave el manejo de los filtros por parte los comerciales de Italian shipping cargo para generar mayor afinidad en el momento de abordar a los clientes teniendo en cuenta en las características de los resultados de kmeans ya que la información va a estar en vivo para la hoja de ruta comercial con el cliente

Gráfica 20 Tabla de clientes

N° CC	N° CONTACTO	NOMBRE	RECOMPRAS (1R) PRODUCTO	(2R) PRODUCTO	1R SECTOR QUE MAS EXPORTA	2R SECTOR QUE MAS EXPORTA	VALOR TOTAL
5006090659	3014307575	ALEJANDRO MAYA VELASQUEZ	7 CON EQUIPO DE ENFRIAMIENTO SUPERIOR A 30000 BTU/HORA PERO INFERIOR A 60000 BTU/HORA	LOS DMS ARTICULOS CONFECCIONADOS- DE LAS DEMAS MATERIAS TEXTILES	FABRICACION DE OTROS TIPOS DE MAQUINARIA Y EQUIPO DE USO GENERAL N.E.P.	CONFECCION DE ARTICULOS CON MATERIALES TEXTILES, EXCEPTO PRENDAS DE VESTIR	\$48.437
1143833681	3203873177	ARIAS GALEAS JORGE ISAAC	14 LOS DMS MEDICAMENTOS PARA USO HUMANO- DOSIFICADOS (INCLUIDOS LOS DESTINADOS A SE ADMINISTRADOS POR VIA TRANSDERMICA) O ACONDICIONADOS PARA LA VENTA AL POR MENOR	PANTALONES LARGOS- PANTALONES CON PETO- PANTALONES CORTOS (CALZONES) Y SHORTS (EXCEPTO DE BAÑO)- DE ALGODON- PARA HOMBRES O NIÑOS- DE TEJIDOS LLAMADOS -MEZCLILLA O DENIM-	FABRICACION DE PRODUCTOS FARMACEUTICOS, SUSTANCIAS QUIMICAS MEDICINALES Y PRODUCTOS BOTANICOS DE USO FARMACEUTICO	CONFECCION DE PRENDAS DE VESTIR, EXCEPTO PRENDAS DE PIEL	\$57.632
34324243	4296732	VELEZ MANCERA EDGAR PAUL	10 BULBOS- CEBOLLAS- TUBERCULOS- RAICES Y BULBOS TUBEROSOS- TURIONES Y RIZOMAS- EN REPOSO VEGETATIVO	CAMISAS- DE ALGODON- PARA HOMBRES O NIÑOS- EXCEPTO LAS DE PUNTO	PROPAGACION DE PLANTAS (ACTIVIDADES DE LOS VIVEROS, EXCEPTO VIVEROS FORESTALES)	CONFECCION DE PRENDAS DE VESTIR, EXCEPTO PRENDAS DE PIEL	\$48.669
52795810	3014307632	YESENIA CAROLINA SANCHEZ AMAYA	8 LAS DEMAS PREPARACIONES TENSOACTIVAS- LA DEMAS PREPARACIONES PARA LAVAR Y LAS DEMAS PREPARACIONES DE LIMPIEZA- AUNQUE CONTENGAN JABON- EXCEPTO LOS DE LA PARTIDA 3401	CONJUNTOS DE ABRIGO PARA ENTRENAMIENTO O DEPORTE (CHANDALES) Y LAS DEMAS PRENDAS DE VESTIR- DE FIBRAS SINTETICAS O ARTIFICIALES- PARA HOMBRES O NIÑOS- EXCEPTO LOS DE PUNTO	FABRICACION DE JABONES Y DETERGENTES, PREPARADOS PARA LIMPIAR Y PULIR; PERFUMES Y PREPARADOS DE TOCADOR	CONFECCION DE PRENDAS DE VESTIR, EXCEPTO PRENDAS DE PIEL	\$46.910
1015447555	5346758	JORGE ALBERTO GARCIA CARUZO	9 PLACAS- LAMINAS- HOJAS- Y TIRAS DE POLICARBONATOS NO CELULARES Y SIN REFUERZO- ESTRATIFICACION NI SOPORTE O COMBINACION SIMILAR CON OTRAS MATERIAS	ALGODON- SIN CARDAR NI FEINAR- DE LONGITUD DE FIBRA SUPERIOR A 22.22 MM (7/8 PULGADA) PERO INFERIOR O IGUAL A 28.57 MM (1 1/8 PULGADA)	FABRICACION DE FORMAS BASICAS DE PLASTICO	CULTIVO DE PLANTAS TEXTILES	\$48.214
1121899906	3014307536	JARAMILLO HENAO JOSE OVIDIO	1 LOS DMS MEDICAMENTOS PARA USO HUMANO- QUE CONTENGAN HORMONAS U OTROS PRODUCTOS DE LA PARTIDA 29.37- DOSIFICADOS (INCLUIDOS LOS DESTINADOS A SE ADMINISTRADOS POR VIA TRANSDERMICA) O ACONDICIONADOS PARA LA VENTA AL POR	LOS DMS LIBROS- FOLLETOS E IMPRESOS SIMILARES	FABRICACION DE PRODUCTOS FARMACEUTICOS, SUSTANCIAS QUIMICAS MEDICINALES Y PRODUCTOS BOTANICOS DE USO FARMACEUTICO	EDICION DE LIBROS	\$795.577

Fuente: elaboración propia

Conclusión

En el dashboard comercial se encuentran gráficos intuitivos los cuales muestran las importaciones totales de los productos. Por ejemplo se destacan tres tipos de productos los cuales son relevantes en los clusters, MPANS; ETPR y MPAA, estos productos generan una oportunidad comercial significativa. Además este dashboard permite a los comerciales explorar cuales son las vías que mas utilizan los clientes de los clusters, se destaca con respecto a los cluster que la via marítima es la preferida por los clientes, también ofrece información sobre las regiones de destino de importación que son mas relevantes para los clusters, se identifica que las

regiones a las que mas importan los clientes son occidental y central, al igual el dashboard permitira explorar a medida del tiempo cuales son las tendencias de importación.

11.3.2. Dashboard gerencial

Este dashboard es una poderosa herramienta la cual ofrecerá a los gerentes de Italian shipping cargo una visión panorámica y de gran valor con respecto al rendimiento comercial de la empresa, mas al detalle está enfocado principalmente en el resultado del performance de los diferentes asesores, para evidenciar claramente cuáles son los clientes que más han sido convertidos en operaciones por los diferentes asesores mediante el funnel de ventas y así poder tomar decisiones en cuanto como debe la gerencia, enfocar a los asesores con respecto a las características de los clientes y las necesidades del mercado

11.3.2.1. Filtros específicos

Utilizamos los filtros para:

- Mes
- Cliente
- Producto de importación
- Sector o región de importación
- Los 8 clúster del resultado de K-means (1 significa que pertenece y 0 que no pertenece)

Los filtros interactivos permiten a los gerentes acceder rápidamente a datos relevantes

sobre clientes y la gestión de los diferentes asesores, esto ayudara a tomar decisiones mas fundamentadas con respecto al área comercial

Gráfica 21 Filtros dashboard gerencial



Fuente: elaboración propia

11.3.2.2. Gráfico de anillos de tipos de productos y sectores de importación

Con este tipos de tablero la Gerencia puede identificar cuáles son los sectores que más utilidad y operaciones le están generando en los diferentes meses a la empresa, para así tomar decisiones de valor internas con respecto a los tipos de productos, vías y regiones de importación que se ofrecen al cliente, para aumentar el engagement y el número de operaciones (ventas)

Gráfica 22 Gráficos de anillos dashboard gerencial



Fuente: elaboración propia

11.3.2.3. Etiquetas de datos

En las etiquetas de datos la gerencia podrá visualizar fácilmente el valor del total importado y el número de recompras de cada uno de los clientes desde su primera compra, esto ayudara a la gerencia a mapear de una manera más fácil cuales clientes están llevando utilidad a la empresa y cuáles son sus características con respecto a tipos de productos, vías y regiones de importación.

Gráfica 23 Etiquetas de datos dashboard gerencial



Fuente: elaboración propia

11.3.2.4. *Gráfico de barras para los resultados del aérea comercial y mapa de ventas:*

Estos dos tipos de tableros nos ayudan a identificar como es el resultado de cada uno de los asesores comerciales de la empresa, ver su cumplimiento con respecto a la cuota individual y evidenciar en el mapa cuales son las regiones y tipos de productos que más generan operación por cada asesor, para así poder enfocarlos en donde mejor resultados.

Gráfica 24 Gráficos dashboard gerencial

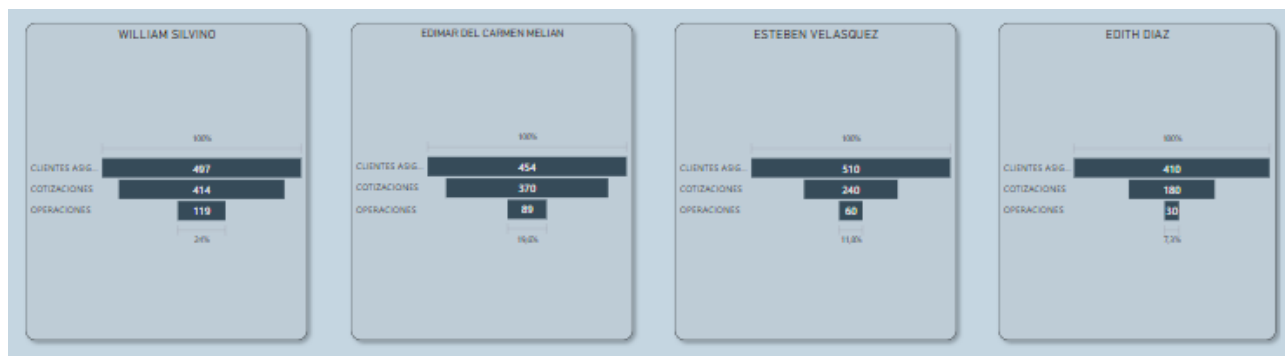


Fuente: elaboración propia

11.3.3. **Funnel de ventas**

En este funnel de ventas podemos identificar por cada uno de los resultados de los asesores en la operación, cuáles son los clientes que hay en cada etapa del funnel, en que etapa generan más conversión en cual se está limitando el proceso, para así evaluar desempeño y poder tomar decisiones más precisas con la operatividad de los asesores.

Gráfica 25 Funnel de ventas



Fuente. elaboración propia

Esta poderosa herramienta ofrecera a los gerentes una visión panorámica y de gran valor del rendimiento general de la empresa en el área comercial. Al ingresar al dashboard, los gerentes pueden acceder a gráficos y tablas que muestran los indicadores clave de rendimiento y las tendencias comerciales más relevantes.

El dashboard gerencial permite a los gerentes monitorear el desempeño de las operaciones (ventas), el crecimiento y decrecimiento de las importaciones que genera cada uno de los comerciales y la eficacia de las estrategias comerciales implementadas. Esto les proporciona información valiosa para evaluar el éxito de las acciones tomadas y realizar ajustes estratégicos con base a los resultados de los datos de la operación en el transcurso de cada mes. Además, el dashboard permite una comparación entre períodos, lo que facilita el seguimiento del progreso a lo largo del tiempo.

11.4. Conclusiones de los Dashboards:

Con base al cumplimiento de la metodología propuesta, concluimos que generar la trazabilidad para la implementación de un producto de analítica de datos en la empresa, ha

mostrado detalles de gran importancia en el aérea comercial para generar mucho más valor en la toma de decisiones, además se evidencian en los resultados, como se puede tener una mejor gestión para el área con base a el manual de gobierno de datos, el cual traza la hoja de ruta para entender los resultados finales que se muestran en el dashboard. Encontramos cuales son los tipos de productos más relevantes, las vías más utilizadas y las regiones de destino más influyentes descritos en las características de cada uno de los clusters las cuales fueron tratadas mediante el dashboard para que sea más digerible la información para los dos actores más importantes, la gerencia y el área comercial.

12. Plan y recomendaciones de implementación y aplicación

12.1. Recomendaciones:

12.1.1. Estructuración de datos

Lo primero que vamos a recomendar a la organización es que a partir del análisis que se realizó, y por el cual se determinó que no se tienen procesos para manipular y darle valor a los datos estructurados, se deben generar procedimientos claros en el área comercial para que la compañía cuente con la información necesaria para generar análisis del movimiento de sus propios clientes y así mismo poder generar análisis hacia el mercado

12.1.2. Procedimientos

Se plantea a la organización la necesidad de generar procedimientos que incluyan responsabilidades y políticas claros para que, desde la creación de los datos como datos maestros tanto en el ERP como en todos, de tal forma que se cuente con los datos necesarios y con las herramientas necesarias para que se pueda hacer una clara asociación de esta información. Se recomienda que se cree en especial un proceso para el gobierno de datos.

En caso de que la compañía no implemente este tipo de procedimientos podría continuar con la ausencia de un análisis de mercado lo que no ayudaría con el crecimiento en ventas. Al conocer que los competidores realizan este tipo de análisis nos encontramos en desventaja para procurar el crecimiento de la compañía.

12.1.3. Almacenamiento de la información

Al contar con la información totalmente descentralizada, se propone a la compañía que valide las diferentes opciones disponibles para que la información sensible se encuentre administrada y centralizada.

Lo anterior para evitar que, por rotación de personal, daño en dispositivos de empleados o algún posible robo de la información, esta se pueda perder, o que la incorrecta manipulación de esta ponga en riesgo que la organización no cuente con lo necesario para toma de decisiones o pierda información valiosa del negocio.

12.1.4. Capacitación al personal

Se le plante a la organización la necesidad de generar un plan de capacitación que vaya acorde a los procedimientos que se plantean, y que incluya como una de las finalidades del mismo la mejora continua y el encaminar los datos de la organización en pro de tener datos confiables para el análisis y toma de decisiones.

12.2. Implementación de la alternativa:

Para implementar las recomendaciones propuestas se diseñó un manual de gobierno de datos, para la implementación del mismo se deben seguir los siguientes pasos:

- Establecer un equipo de trabajo encargado de la implementación de las acciones propuestas.
- Definir un plan detallado que incluya los plazos, recursos necesarios y las actividades específicas para cada recomendación.
- Asignar responsabilidades claras a cada miembro del equipo para garantizar una ejecución eficiente.
- Realizar un seguimiento continuo del progreso y ajustar el plan según sea necesario.

12.2.1. Personal involucrado y responsabilidades:

El equipo de implementación estará conformado por personal clave de diferentes áreas de la empresa, incluyendo el área comercial, tecnología de la información y gestión de datos. Las responsabilidades serán las siguientes:

El área comercial será responsable de proporcionar los datos necesarios y participar en la definición de los procedimientos y políticas relacionadas con la gestión de datos.

El equipo de tecnología de la información será responsable de establecer los procesos técnicos necesarios para la estructuración y almacenamiento de datos, así como garantizar la seguridad de la información.

Se propone la creación del área de gestión o de un cargo que haga las veces de la misma, el cual se encargará de supervisar y asegurar el cumplimiento de las políticas y procedimientos establecidos, así como de brindar capacitación al personal.

12.2.2. Momento y lugar de implementación:

La implementación de las recomendaciones debe llevarse a cabo de manera progresiva y planificada. Se deben establecer hitos y plazos específicos para cada acción propuesta. La implementación puede realizarse en las instalaciones de la empresa o de manera remota, dependiendo de las circunstancias y los recursos disponibles.

12.2.3. Resultado a obtener:

El objetivo principal es mejorar la gestión de datos y la toma de decisiones comerciales en Italian Shipping Cargo. Se espera lograr los siguientes resultados:

- Establecimiento de procesos claros para la estructuración y análisis de datos.

- Mejora en la calidad y confiabilidad de los datos utilizados en el área comercial.
- Implementación de políticas y procedimientos para el gobierno de datos.
- Centralización y seguridad de la información sensible.
- Capacitación del personal en el manejo adecuado de datos y análisis.

12.2.4. Medición del éxito o fracaso:

El éxito de la implementación se medirá mediante indicadores clave de rendimiento asociados a la espera de la obtención de los resultados del punto anterior.

13. Conclusiones

En el presente proyecto empresarial, se ha abordado la problemática de la gestión de datos y la falta de análisis en el área comercial de la empresa Italian Shipping Cargo. A través del análisis de la situación organizacional, la comprensión del negocio y la identificación de los problemas principales y secundarios, se han propuesto recomendaciones y estrategias para abordar dichos problemas y mejorar la eficiencia en la toma de decisiones comerciales.

Una de las principales conclusiones obtenidas es que la falta de una adecuada estructuración de datos y la ausencia de procesos establecidos para el análisis y seguimiento comercial han limitado el potencial de crecimiento de la empresa. Asimismo, se evidenció que la falta de aprovechamiento de recursos intangibles, como la plataforma SIICEX, ha dejado a la empresa en desventaja frente a la competencia y ha limitado su capacidad para tomar decisiones basadas en datos precisos y actualizados.

Para solucionar estas problemáticas, se propone la implementación de un manual de gobierno de datos que establezca los procedimientos y políticas necesarios para una gestión eficiente de la información en el área comercial. Este manual incluirá directrices para la estructuración de datos, la centralización y seguridad de la información sensible, y la capacitación del personal en el manejo adecuado de datos y análisis.

La implementación de este manual de gobierno de datos permitirá a Italian Shipping Cargo contar con una base sólida para la gestión de datos, mejorando la calidad y confiabilidad de la información utilizada en la toma de decisiones comerciales. Además, facilitará la

generación de análisis y reportes más precisos y actualizados, lo cual contribuirá a un mayor entendimiento del mercado y a la identificación de oportunidades de crecimiento y mejora en la gestión y logística comercial.

Como conclusión adicional se evidencio la necesidad de la adopción del funnel de ventas en Italian Shipping Cargo lo cual llevaría una serie de beneficios significativos para la empresa. En primer lugar, proporcionará una visión clara y estructurada del proceso de ventas, desde la adquisición de clientes potenciales hasta el cierre de la venta. Esto permitirá identificar y comprender mejor los pasos clave en el proceso de ventas, así como los posibles obstáculos o cuellos de botella que puedan surgir. Además, el funnel de ventas facilitará la segmentación y clasificación de los clientes potenciales, lo que permitirá a la empresa enfocar sus esfuerzos en aquellos más propensos a convertirse en clientes reales. A su vez, esto generará un aumento en la eficiencia y efectividad de las actividades comerciales, al dirigir los recursos y esfuerzos hacia los prospectos con mayor probabilidad de cerrar una venta exitosa. Por último, el funnel de ventas proporcionará métricas y datos cuantitativos que permitirán medir el desempeño y la eficacia de las estrategias comerciales, lo que facilitará la toma de decisiones basada en datos y la implementación de mejoras continuas en el proceso de ventas.

Referencias bibliográficas

- Baker, R. (2019). Marketing Funnel: La Guía Completa Para Entender La Psicología Del Cliente, Crear Un Embudo De Ventas y Aumentar Los Beneficios. Cómo Configurar Google Analytics y Optimizar El Tipo De Conversión. Independently Published
- Schwaber, K., & Sutherland, J. (2020). “La Guía Definitiva de Scrum: Las Reglas del Juego” Scrum Guides. <https://scrumguides.org/docs/scrumguide/v2020/2020-Scrum-Guide-Spanish-Latin-South-American.pdf>
- Sharda, R. (2017). Business Intelligence, Analytics and Data Science. Pearson. ISBN: 1-292-22054-6.
- DAMA International. (2010). DAMA - GUIA DE FUNDAMENTOS PARA LA GESTION DE DATOS. DAMA International.: Technics Publications, LLC.
- Hernández Lalinde, J. D., Espinosa Castro, F., Rodríguez, J. E., & Chacón Rangel, J. (2018). Sobre el uso adecuado del coeficiente. Venezuela: Avft.
- Kaur, P., & Singh, M. (2023). Modelo de segmentación de clientes usando K-means. India: (ICSCDS).
- Ochoa, I., Paredes, K., & Esquicha, J. (2017). Estudio Comparativo de Técnicas no Supervisadas de Minería de Datos para Segmentación de Alumnos. Boca Raton: Global Partnerships for Development and Engineering Education.
- Perez, D., & Galvis, E. (2021). Análisis de la estabilidad de los centroides en K-Medias en presencia de correlación. Bogotá: Universidad Nacional de Colombia.
- Perucha, C., & Carlos, M. (2022). EL METODO DE. Valladolid: Universidad de Valladolid.

Raquel Arganza Salcedo & Miriam Arroyo López. (2019). Big data: Aplicaciones de la gestión del dato en las distintas etapas del funnel de conversión. *Revista de Marketing y Publicidad*, 1(1).

Directory of Open Access Journals. <https://doi.org/10.51302/marketing.2019.684>

Hernández Lalinde, J. D., Espinosa Castro, J. F., Peñalosa Tarazona, M. E., Fernández González, J. E., Chacón Rangel, J. G., Toloza Sierra, C. A., Arenas Torrado, M. K., Carrillo Sierra, S. M., & Bermúdez Pirela, V. J. (2018). *Sobre El Uso Adecuado Del Coeficiente De Correlación De Pearson: Definición, Propiedades Y Suposiciones*

(edsair.dedup.wf.001..45a92c2cd9f65f99a26cb492d4c18d16). OpenAIRE.

https://explore.openaire.eu/search/publication?articleId=dedup_wf_001::45a92c2cd9f65f99a26cb492d4c18d16

Gonzalez Jaramillo, V. H., Zanzzi Diaz, P. F., Altamirano Zambrano, R., & Ganchozo-Ortega, G.

(s. f.). Proceedings of the 15th LACCEI International Multi-Conference for Engineering, Education, and Technology: “Global Partnership for Development and Engineering Education”. En *Proceedings of the 15th LACCEI International Multi-Conference for Engineering: Vol. Education* (p. 20170821).

A guide to the Scrum Body of knowledge (SBOK Guide) (2013 edition). (2013). SCRUM study, A brand of VMEdU, Inc.

Earley, S., Henderson, D., & Data Management Association (Eds.). (2017). *DAMA-DMBOK: Data management body of knowledge* (2nd edition). Technics Publications.