

UNIVERSIDAD CATÓLICA SANTO TORIBIO DE MOGROVEJO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN



Sistema de inteligencia de negocios haciendo uso de tecnologías cloud para apoyar en la toma de decisiones comerciales de un minimarket en Chiclayo

**TESIS PARA OPTAR EL TÍTULO DE
INGENIERO DE SISTEMAS Y COMPUTACIÓN**

AUTOR

Tatiana Leonor Flores Navarrete

ASESOR

Segundo Jose Castillo Zumaran

<https://orcid.org/0000-0001-5613-5519>

Chiclayo, 2023

Sistema de inteligencia de negocios haciendo uso de tecnologías cloud para apoyar en la toma de decisiones comerciales de un minimarket en Chiclayo

PRESENTADA POR:

Tatiana Leonor Flores Navarrete

A la Facultad de Ingeniería de la
Universidad Católica Santo Toribio de Mogrovejo
para optar el título de

INGENIERO DE SISTEMAS Y COMPUTACIÓN

APROBADA POR:

Consuelo Ivonne Del Castillo Castro

PRESIDENTE

Hector Miguel Zelada Valdiviezo

SECRETARIO

Segundo Jose Castillo Zumaran

VOCAL

DEDICATORIA

A Dios por darme la fortaleza necesaria para culminar mi carrera profesional.

A mis padres y hermanos por su apoyo incondicional y por impulsarme a cumplir con mis metas personales y profesionales.

AGRADECIMIENTOS

A mi asesor de tesis por su gran apoyo en el desarrollo y culminación de mi trabajo de investigación

A la empresa por brindarme las facilidades de desarrollar y poner en practica mis conocimientos.

Tesis V2.0

INFORME DE ORIGINALIDAD

25% INDICE DE SIMILITUD	25% FUENTES DE INTERNET	2% PUBLICACIONES	6% TRABAJOS DEL ESTUDIANTE
-----------------------------------	-----------------------------------	----------------------------	--------------------------------------

FUENTES PRIMARIAS

1	hdl.handle.net Fuente de Internet	10%
2	repositorio.unprg.edu.pe:8080 Fuente de Internet	4%
3	tesis.usat.edu.pe Fuente de Internet	2%
4	repositorio.udl.edu.pe Fuente de Internet	1%
5	repositorio.upao.edu.pe Fuente de Internet	1%
6	repositorio.ucv.edu.pe Fuente de Internet	1%
7	Submitted to Universidad Autónoma de Bucaramanga, UNAB Trabajo del estudiante	<1%
8	www.captio.net Fuente de Internet	<1%

Índice

RESUMEN.....	9
ABSTRACT.....	10
I. INTRODUCCIÓN	11
II. REVISIÓN DE LA LITERATURA.....	13
2.1. ANTECEDENTES.....	13
2.1.1. ANTECEDENTES INTERNACIONALES	13
2.1.2. ANTECEDENTES NACIONALES	15
2.1.3. ANTECEDENTES LOCALES	16
2.2. BASES TEÓRICO CIENTÍFICAS	18
2.2.1. BUSINESS INTELLIGENCE	18
2.2.1.1. Data Mart	18
2.2.1.2. Data Flow.....	19
2.2.2. CLOUD COMPUTING	19
2.2.2.1. Azure.....	19
2.2.2.2. Azure Database for MySQL.....	20
2.2.2.3. Azure Data Factory.....	20
2.2.3. METODOLOGÍA DE RALPH KIMBALL	21
III. MATERIALES Y MÉTODOS.....	22
3.1. TIPO DE INVESTIGACIÓN	22
3.2. MÉTODOS DE INVESTIGACIÓN	22
3.3. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS	22
3.4. PROCEDIMIENTOS	22
3.4.1. METODOLOGÍA DE DESARROLLO.....	22
3.4.2. PRODUCTO ACREDITABLE	24
3.4.3. MANUAL DE USUARIO	24
3.5. MATRIZ DE CONSISTENCIA	25
3.6. CONSIDERACIONES ÉTICAS.....	26
IV. RESULTADOS Y DISCUSIÓN.....	27

4.1. EN BASE A LA METODOLOGÍA UTILIZADA	27
4.1.1. FASE #1: DEFINICIÓN DE REQUERIMIENTOS DEL NEGOCIO	27
4.1.2. FASE #2: DISEÑO DE LA ARQUITECTURA TÉCNICA	28
4.1.3. FASE #3: SELECCIÓN DE PRODUCTOS E IMPLEMENTACIÓN	28
4.1.4. FASE #4: MODELADO DIMENSIONAL.....	29
4.1.5. FASE #5: DISEÑO FÍSICO	29
4.1.6. FASE #6: DISEÑO E IMPLEMENTACIÓN DEL SUBSISTEMA DE ETL	33
4.1.7. FASE #7: ESPECIFICACIÓN DE APLICACIONES DE BI.....	36
4.1.8. FASE #8: DESARROLLO DE APLICACIONES DE BI	37
4.1.9. FASE #9: ADMINISTRACIÓN DEL PROYECTO DE DWBI.....	40
4.2. IMPACTOS ESPERADOS	41
4.2.1. IMPACTOS ECONÓMICOS	41
4.2.2. IMPACTOS SOCIALES.....	41
4.2.3. IMPACTOS EN TECNOLOGÍA	41
V. DISCUSIÓN	42
VI. CONCLUSIONES.....	44
VII. RECOMENDACIONES.....	45
VIII. REFERENCIAS.....	46
IX. ANEXOS.....	48
ANEXO N° 01. INSTRUMENTOS DE RECOLECCIÓN DE DATOS.....	48
ANEXO N° 02. RESPUESTAS JUICIO DE EXPERTOS	51
ANEXO N° 03. MANUAL DE USUARIO	54
A. REQUERIMIENTOS DE HARDWARE.....	54
SUGERIDO DE MICROPROCESADOR	54
SUGERIDO DE MEMORIA RAM	54
B. REQUERIMIENTO DE SOFTWARE	54
SISTEMA OPERATIVO	54
NAVEGADOR WEB	54
C. RUTA DE ACCESO	54

Lista de tablas

TABLA I.....	22
TABLA II.....	22
TABLA III	25
TABLA IV	53

Lista de figuras

FIG. 1 FASES RALPH KIMBALL.....	23
FIG. 2 ARQUITECTURA LÓGICA DEL SISTEMA DE INTELIGENCIA DE NEGOCIO.....	28
FIG. 3 DIAGRAMA STARNET.....	29
FIG. 4 DIAGRAMA DE BASE DE DATOS DEL SISTEMA COMERCIAL.....	30
FIG. 5 BASE DE DATOS DIMENSIONAL.....	31
FIG. 6 PIPELINE DIM_PRODUCTO.....	33
FIG. 7 PIPELINE DIM_PERSONA.....	34
FIG. 8 PIPELINE DIM_CLIENTE.....	34
FIG. 9 PIPELINE DIM_SUCURSAL.....	34
FIG. 10 PIPELINE DIM_TIEMPO.....	35
FIG. 11 PIPELINE DIM_TIPO_DOCUMENTO.....	35
FIG. 12 PIPELINE DIM_TIPO_VENTA.....	35
FIG. 13 PIPELINE HECHO_INGRESOS.....	36
FIG. 14 DASHBOARD DE ANÁLISIS GENERAL.....	37
FIG. 15 DASHBOARD DE ANÁLISIS POR EL CRITERIO DE PRODUCTO.....	38
FIG. 16 DASHBOARD DE RANKING DE INGRESOS.....	39
FIG. 17 DASHBOARD DE TENDENCIA DE INGRESOS.....	40

Resumen

La presente tesis aborda la situación problemática de una incorrecta toma de decisiones comerciales de un minimarket en Chiclayo, ante lo cual se tuvo como objetivo general, la implementación de un sistema de inteligencia de negocio que haga uso de tecnología cloud y que pueda brindar información descriptiva como apoyo. Para llevar a cabo esta investigación se tuvieron como objetivos específicos el determinar los datos relevantes para comprender el comportamiento de las ventas en las diferentes sucursales del minimarket, analizar y limpiar la data resultante del sistema de ventas, construir reportes analíticos para dar soporte en la toma de decisiones comerciales del jefe de ventas y validar los reportes analíticos mediante el juicio de expertos. La investigación fue de tipo aplicada e hizo uso del método analítico, deductivo e implementación. Finalmente se obtuvo como resultado, la implementación de un datamart compuesto por 7 dimensiones y 1 hecho que alimentaron de información a los 4 dashboards que se presentaron en base a los indicadores definidos durante la primera fase de la metodología de Ralph Kimball. Cabe resaltar que las herramientas empleadas, Data Factory y Power BI, pertenecen a la plataforma de Microsoft Azure, pero sería recomendable poder evaluar otras tecnologías cloud.

Palabras clave: inteligencia de negocios, minería de datos, toma de decisiones, ventas, Microsoft Azure.

Abstract

This thesis addresses the problematic situation of incorrect commercial decision making of a minimarket in Chiclayo, in this case the general objective was the implementation of a business intelligence system that makes use of cloud technology and that can provide descriptive information as support. To carry out this research, the specific objectives were to determine the relevant data to understand the behavior of sales in the different branches of the minimarket, analyze and clean the data resulting from the sales system, build analytical reports to support the commercial decision making of the sales manager and validate the analytical reports through the judgment of experts. The research was of an applied type and made use of the analytical, deductive and implementation method. Finally, the result was the implementation of a datamart composed of 7 dimensions and 1 fact that fed information to the 4 dashboards that were presented based on the indicators defined during the first phase of Ralph Kimball's methodology. It should be noted that the tools used, Data Factory and Power BI, belong to the Microsoft Azure platform, but it would be recommendable to be able to evaluate other cloud technologies.

Keywords: business intelligence, data mining, decision making, sales, Microsoft Azure.

I. Introducción

En la actualidad, es de conocimiento mundial que, el mayor número de información se encuentra digitalizada y el volumen de la misma aumenta de forma exponencial a gran velocidad; el 90% de los datos mundiales se han creado en los últimos dos años, logrando así que sea de fácil acceso para cada vez más personas. Dicho crecimiento de información es beneficioso si se usa de forma correcta, pero para muchas empresas termina siendo un problema por no saber gestionar una gran cantidad de datos [1]. Dichos datos por sí solos no suelen tener valor, pero si se contextualizan de forma correcta y se analizan, es posible que se revelen ciertos patrones que permitirían comprender su comportamiento [2]. Es por ello, que aquellas empresas que implementan soluciones de TI avanzadas, pueden analizar los datos y generar ventajas competitivas. Dicho tipo de implementaciones son mejor conocidas como Inteligencia de negocio (BI) y dan soporte al proceso de toma de decisiones [3].

El Business Intelligence es un proceso que aprovecha los datos que se tienen y poder así implementar mejoras en las empresas. La BI prioriza un análisis descriptivo que permite visualizar un resumen de los datos, tanto históricos como actuales, y determinar qué es lo que ha estado sucediendo o sucede actualmente. Gracias a la Inteligencia de negocios, se logra dar respuestas a preguntas del tipo “qué” y “cómo” con la finalidad de poder replicar lo que ha sido útil y evitar lo que no [4]. Según [5], sin importar el tamaño de la empresa, incluso las PyMES obtendrían grandes beneficios con solo poder contar con una solución de reporting, pero la realidad es que solo un 21% de las empresas latinoamericanas usan sus datos para acelerar su transformación digital [6].

En el Perú, según [7], la creación de empresas ha aumentado en un 11.9% durante el 2021, de las cuales, el mayor número de altas se dio en aquellas que se dedican al comercio al por menor (29.5%), seguidas por aquellas que se dedican al comercio al por mayor (15.4%) y pertenecen al rubro de PyMES, las cuales representan un 99.5% de las empresas que existen en el país [8]. Estas cifras demuestran lo que nos dice [9], el cual sostiene que las PyMES son la base del

crecimiento de nuestro país y que éstas durante la pandemia del coronavirus, han tenido que hacer cambios de forma brusca para poder mantenerse. La principal modificación ha sido el tener que pasar por un proceso de digitalización, adoptando nuevas tecnologías acorde a sus necesidades y que les permitieran enfrentar así los cambios de una forma más versátil y rápida.

El minimarket ubicado en la ciudad de Chiclayo, es un claro ejemplo de lo antes mencionado dado que es una mediana empresa que tuvo que adoptar un sistema comercial que le permitiera llevar un mejor control sobre sus ventas y productos desde los inicios de la pandemia, pero a pesar de que éste sistema le apoyó en poder tener sus ingresos y stock correctamente ordenado, su crecimiento ha sido lento al no poder determinar con rapidez sobre las mejores estrategias comerciales para poder brindar un mejor servicio a sus clientes. Según expresó el jefe de ventas del minimarket, a pesar de contar con datos, estos no eran analizados con detenimiento debido a que eran demasiados y resultaba una tarea complicada de llevar a cabo, así que cuando decidían usarlos para gestionar sus ofertas, no siempre eran éstas exitosas. Mientras que cuando escogían el producto correcto, sus ganancias lograban aumentar hasta en un 20%; cuando no lo hacían, las pérdidas podían llegar a superar hasta un 50% de la ganancia diaria debido al sobre stock del producto que ofrecían en promoción. Además, de que no contaban con otros indicadores que le permitieran tomar mejores decisiones para acelerar su crecimiento.

Ante dicha situación, surgió la siguiente pregunta ¿Cómo se podría apoyar en la toma de decisiones comerciales de un minimarket de Chiclayo? y para dar respuesta a esta problemática se determinó que con la implementación de un Sistema de inteligencia de negocios que haga uso de tecnologías cloud se podría brindar la solución deseada dado que con dicho sistema alimentado por la data histórica generada por su sistema transaccional, el jefe de ventas lograría visualizar información relevante sobre sus ventas que le permitieran analizarlas a detalle y lograr determinar por ende los productos más vendidos, los días en los que realiza más ventas, la sucursal que más ingresos genera, entre otros indicadores relevantes.

El objetivo principal de la investigación fue implementar un sistema de inteligencia de negocios que apoye en la toma de decisiones comerciales de un minimarket de Chiclayo. Se plantearon como objetivos específicos el determinar los datos relevantes para comprender el comportamiento de las ventas en las diferentes sucursales del minimarket, analizar y limpiar la data resultante del sistema de ventas, construir reportes analíticos para dar soporte en la toma de decisiones comerciales del jefe de ventas y validar los reportes analíticos mediante el juicio de expertos.

La investigación fue de tipo aplicada y se justificó a nivel económico porque al poder contar con información procesada, resultante de los datos generados por su sistema transaccional, el jefe de ventas pudo tomar decisiones acertadas no solo al momento de definir las ofertas sino además al momento de invertir en sus sucursales, lo cual generó un ahorro significativo. A nivel social, la investigación se justificó puesto que al ofertar los productos que realmente son más consumidos, las necesidades reales de los clientes, fueron atendidas y por ende su satisfacción aumentó y; finalmente, se justificó a nivel tecnológico dado que se emplearon tecnologías cloud para el desarrollo del sistema de inteligencia de negocios y gracias a ello se pudo brindar acceso inmediato a la información desde cualquier dispositivo y lugar.

II. Revisión de la literatura

2.1. Antecedentes

A continuación se detallan los antecedentes considerados para la investigación:

2.1.1. Antecedentes internacionales

Zapata [10], narra los problemas existentes en la ferretería “del ahorro” a nivel de sus procesos de facturación, para lo cual propone desarrollar un aplicativo para celulares mediante el desarrollo de un modelo de inteligencia de negocios usando la herramienta Microsoft Power BI. El valor agregado de esta investigación es que empleó el entorno móvil de Power BI y además hizo una comparación con el

entorno de escritorio. Determinando que, en su entorno móvil se pueden evidenciar ciertas deficiencias menores, que no afectarían las medidas grandes de toma de decisiones, a excepción de que en algunas oportunidades el tiempo de respuesta es bastante extenso y por ende se requieren celulares de gama media/alta para remediar el hecho de que requieren tomar otras medidas, lo que complica la forma en la que se visualiza la información. Finalmente, concluyó que es recomendable emplear Power BI a la par que se utilizan aplicaciones de Microsoft tales como, servicios en la nube, bases de datos y servicios de plataforma. Se hizo uso de esta tesis ya que llevó a cabo la integración con Microsoft Azure como medio de propagación del producto final.

Según [11], en su investigación menciona que retraso mayor en los proyectos de Inteligencia de Negocios se produce cuando se llevan a cabo los métodos de Extracción, Transformación y Carga de datos (ETL) y es por ello que propone una metodología en la cual se fortalezca la fase del proceso ETL durante los proyectos de BI. El valor diferencial de esta tesis fue la metodología moderna propuesta con la que finalmente concluye que de esta manera lograría obtener que periódicamente se cuente con una formalidad mayor y una mayor estructuración cuando se ejecute y dé seguimiento de forma apropiada a los mismos. Esta investigación es valiosa para la propia dado que plantea estrategias que permiten disminuir el tiempo invertido en el desarrollo de los proyectos de BI mediante el uso de una metodología que se oriente a la toma de decisiones.

Según [12], el exceso de información almacenada por el Sistema de Contratación Pública del Ecuador hacía que el análisis de los datos y la regulación de los procesos de contrataciones realizadas sea difícil. Para ello, propuso un sistema de inteligencia de negocios que permitiese analizar los datos de los proyectos de software ejecutados durante la década del 2010-2020. El valor agregado de esta investigación es el uso de Pentaho Community como herramienta

para la generación de reportes, además de la ejecución del proceso ETL, creación del Data Warehouse, el uso de una metodología incremental y la implementación de la herramienta Saiku Analytics. Finalmente, se concluyó que el uso de la metodología incremental permite reducir el riesgo en las etapas finales del proyecto y se demuestra su compatibilidad con la metodología Ralph Kimball; además, la creación de reportes permite un análisis con mayor comprensión sobre los factores determinantes. Esta investigación fue relevante dado que no solo permitió contrastar la aplicación de la metodología dentro de un ciclo incremental, sino que además sirvió para comparar las herramientas empleadas.

2.1.2. Antecedentes nacionales

Según Bellido [13], en su tesis expone las problemáticas existentes al momento de tomar decisiones en una agencia de promociones, por ello, usando a Ralph Kimball como metodología, propuso apoyar en dicho proceso aplicando business intelligence. El valor agregado es que hace énfasis en el requerimiento que existe, por parte de los altos mandos de las compañías, de un sistema que les permita identificar data histórica y analizarla para lograr tomar decisiones de manera oportuna y que estas sean acertadas. Finalmente, mostró pruebas de la efectividad del uso de un BI cuando de potenciar cualquier proceso se trata. Cabe enfatizar que se tuvo en cuenta esta investigación como antecedente pues ejecuta una comparación de metodologías cuyo propósito es desarrollar un BI, adicionalmente ratifica el hecho de que los BI dan soporte a los procesos y la toma de decisiones.

Acosta [14], en su investigación expuso los problemas existentes al momento de tomar decisiones el personal encargado del área de distribución y ventas de la compañía Farmacéutica La Libertad S.R.L. Se aplicó la metodología de Larissa Moss con el propósito de lograr analizar los datos y brindar respaldo al proceso de toma de decisiones. El valor agregado fue el uso de herramientas Microsoft

Azure. Finalmente concluyó que el desarrollo del BI se llevó a cabo haciendo uso de herramientas de SQL Server 2014, como son Analysis Services, SQL Server Data Tools con Integración de Servicios y Microsoft Azure para la distribución. Se diseñó un modelo estrella para las áreas de ventas y distribución. Fueron 10 los reportes que se implementaron y se basaron en lo que el usuario requirió, lo cual avaló el correcto desempeño del BI. Dichos reportes se crearon empleando la herramienta de PoweBI. Esta investigación es relevante para la propia, no solo porque la tecnología cloud empleada es similar a la que se usó al momento de la distribución, sino además porque permitió comparar la metodología seleccionada con la de Larissa Moss.

Pereyra [15], en su investigación expone que la gerencia de la empresa está teniendo inconvenientes con el control de costos en el portafolio de sus proyectos. Por ello, haciendo uso del método experimental, propuso emplear la herramienta Power BI con la finalidad de dar soporte en el proceso de toma de decisiones teniéndose como valor agregado la validación de las hipótesis específicas que planteaban la disminución en tiempos y costos, además de la conexión a la base de datos ubicada en Sharepoint. Finalmente se concluyó que la disminución no solo en el tiempo de la toma de decisiones, sino además en el costo y la aceptación de la herramienta fue elevado. Esta investigación es relevante dado que permitió evaluar la opción de hacer uso de tecnología como sharepoint como repositorio para la data que se usaría como input.

2.1.3. Antecedentes locales

Según [16], en su tesis explica la necesidad de tener un correcto monitoreo del proceder de los casos COVID-19 en el Perú. Para lo cual propuso el desarrollo de una plataforma de Business Intelligence que se basara en un análisis multidimensional. El valor agregado es que no solo implementó la solución sino que además llevó a cabo una evaluación de satisfacción y aceptación de la

misma. Concluyendo así que la plataforma BI había sido aceptada con un 87% muy bueno y un 17% bueno, por parte de los encargados de los sectores tanto privados como públicos de salud, lo cual significa que fue de gran ayuda al momento de ofrecer información sumariada y por ende dio soporte a una toma de decisiones ideal para prevenir y estar al tanto sobre los casos de defunciones o confirmaciones. Además, obtuvo un 100 %, a nivel de entendimiento, rapidez y de fácil uso. Finalmente, esta investigación fue relevante para la mía dado que recomienda también que la comunidad haga uso del entorno de la plataforma BI, con el objetivo de que la información pueda llegar a más personas en cualquier lugar.

Fuentes [17], en su investigación expone la problemática existente en las MYPES al momento de tomar decisiones que les permita crecer de forma constante y a la vez mantener la ventaja competitiva. Para que la situación sea solucionada, se planteó el uso de metodologías de BI con el objetivo de implementar un modelo integrado de inteligencia de negocios que sirva de soporte de decisiones e incluso permita renovar las estrategias usadas con el fin de mejorar los resultados de los procesos. El valor agregado fue que se construyó el modelo enfocado en seis dimensiones: estrategias del negocio, preparador dimensional, requerimientos y análisis dimensional, diseño dimensional, integración de datos y explotación de datos. Finalmente se concluyó que el desarrollo basado en las 6 etapas propuestas del modelo, permitió mejorar la aceptación para la toma de decisiones. Esta tesis es importante dado que expone el uso de una metodología pura para la ejecución del modelo a diferencia del uso de una metodología que se enfoca en la data como tal.

Villanueva [18], en su investigación presenta la problemática existente en una empresa farmacéutica al momento de tomar decisiones en su proceso de compra y venta. Para dar solución a esta situación, se planteó el desarrollo de una aplicación de inteligencia

de negocios que se apoyó en un algoritmo de serie temporal implementado haciendo uso del módulo R de Power BI. El valor agregado fue el uso del módulo de R, además del uso de la metodología de Ralph Kimball. Finalmente, se concluyó que los resultados fueron satisfactorios dado que gracias a los dashboards interactivos construidos, en los cuales se lograban visualizar las tendencias de las ventas, la rotación de los productos, entre otros indicadores, se pudo mejorar la satisfacción de los usuarios. Esta investigación es relevante para la propia, dado que muestra la forma en la cual se debe de emplear la metodología de Ralph Kimball, además de servir de referencia para definir los indicadores idóneos para la toma de decisiones en un proceso de ventas.

2.2. Bases teórico científicas

2.2.1. Business Intelligence

Los sistemas de Business Intelligence permiten un mejor acceso a la información de la empresa y un mejor análisis para apoyar la toma de decisiones de sus diferentes tipos de usuarios. Con esto, la organización, los clientes y los proveedores adquieren más conocimiento, integran la cadena de valor y obtienen más fácilmente una ventaja competitiva de este conocimiento. [19].

2.2.1.1. Data Mart

Se denomina así al subconjunto de datos que constituyen un Data Warehouse. Esta clase de “almacén” se limita a fines generales, y por ende debe de ser implementado a la medida del requerimiento. Por otro lado, esta clase de “almacén” permite extraer y preparar los datos necesarios directamente de fuentes locales, lo que proporciona un acceso más rápido a los datos en lugar de tener que acceder a todos los datos del almacén de datos.[20].

Hay tres formas de determinar la implementación de un data mart:

- Se puede lograr un mejor rendimiento y escalabilidad dividiendo los datos en el almacén de datos y cargándolos en bases de datos con esquemas físicos similares.
- Cuando el almacén de datos es inaccesible por algún motivo, los data marts se pueden crear de forma independiente sin extracción.
- En algunos casos, comience creando un data mart y el data warehouse se crea con la finalidad de consolidar la data en cada uno de estos [20].

2.2.1.2. Data Flow

El flujo de datos, como dice su nombre en español, es la configuración del almacenamiento de datos dentro de un sistema de almacén de datos y la disposición de cómo fluyen los datos desde los sistemas de origen a través del almacén de datos. Las aplicaciones utilizan datos con los que interactúan los usuarios finales. El flujo de datos incluye control de datos, métodos de registro e incluso mecanismos para garantizar la calidad de los datos almacenados. [21].

2.2.2. Cloud Computing

Es la posibilidad de hacer uso de hardware y software tercerizado a través de internet. Fue con la llegada de Google Docs y AWS que la computación en la nube se puso a la vanguardia, permitiendo así a las empresas no solo tener una mayor capacidad de procesamiento de información sino además acceder a ésta en tiempo real [22].

2.2.2.1. Azure

Se considera una plataforma en la nube completa, ya que puede alojar aplicaciones de Microsoft existentes y simplificar la implementación de nuevas aplicaciones. Azure puede incluso optimizar las aplicaciones locales.

Azure combina los servicios en la nube que necesita para desarrollar, probar, implementar y administrar aplicaciones mientras aprovecha la computación en la nube.

Con el hospedaje de aplicaciones en Azure, puede comenzar poco a poco y escalar fácilmente sus aplicaciones a medida que aumentan las necesidades de sus clientes. Azure también proporciona la confiabilidad requerida para una alta disponibilidad de las aplicaciones, incluida la conmutación por error entre regiones. Azure Portal le permite gestionar con facilidad todos sus servicios de Azure. También puede gestionar los servicios mediante programación utilizando API y plantillas específicas del servicio [23]

2.2.2.2. Azure Database for MySQL

Es un servicio de base de datos del tipo relacional, de Microsoft Cloud, con soporte de la tecnología MySQL Community Edition. Permite la posibilidad de que la base sea alojada en un servidor que bien puede ser único o flexible. Esta base de datos es un servicio totalmente administrado que permite manejar cargas de trabajo críticas con una escalabilidad dinámica y un rendimiento predecible [24].

2.2.2.3. Azure Data Factory

Es un servicio ETL en la nube de Azure para la transformación e integración de datos sin necesidad de escalar servidores de forma horizontal. Proporciona una interfaz con la que los usuarios pueden interactuar sin usar código, lo que permite la creación, el control y la gestión intuitiva desde un único panel.

Otro beneficio de Azure Data Factory es la capacidad de extraer y editar paquetes SSIS existentes y ejecutarlos de manera totalmente compatible [25].

2.2.3. Metodología de Ralph Kimball

Se enfoca en el dominio empresarial para implementar un DataMart, que luego constituirá un almacén de datos. En comparación con otros métodos, lleva menos tiempo desarrollarlo y construirlo y, debido al tipo de método, también cuesta menos implementarlo.

El método se divide en 10 etapas con un total de 26 tareas para realizar la elaboración de reportes o cuadros de mando requeridos por el proceso ETL y el área de la empresa. [26].

III. Materiales y métodos

3.1. Tipo de investigación

La presente investigación es del tipo aplicada puesto que se implementará un sistema de inteligencia de negocio construido a partir de la información obtenida del minimarket y el mismo se pondrá en uso por la gerencia comercial para apoyar el proceso de toma de decisiones [27].

3.2. Métodos de investigación

Los métodos de investigación empleados serán los siguientes:

Tabla I
MÉTODOS DE INVESTIGACIÓN

Método	Descripción
Analítico	Este método se empleará con la finalidad de inicialmente analizar la situación actual de la empresa y posteriormente la información relevante para el sistema a proponer.
Deductivo	Este método permitirá determinar los indicadores relevantes para la toma de decisiones del negocio
Implementación	Este método permitirá desarrollar el sistema de inteligencia de negocio y ponerlo en funcionamiento.

3.3. Técnicas e instrumentos de recolección de datos

A continuación, en la siguiente tabla se muestra las técnicas e instrumentos que fueron útiles para la recolección de datos.

Tabla II
TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

Técnicas	Instrumentos	Elementos de la población	Propósito
Entrevista	Cuestionario	Jefe de ventas	Conocer las deficiencias en el proceso de toma de decisiones.
Encuesta	Cuestionario	Expertos	Conocer la validez de la herramienta implementada

(ver anexo N° 02)

3.4. Procedimientos

3.4.1. Metodología de desarrollo

Para la ejecución de la propuesta de investigación se planificó hacer uso de la metodología de Ralph Kimball puesto que contempla las fases necesarias para poder desarrollar de forma satisfactoria un sistema de inteligencia de negocios. A continuación, se detallarán las fases de la metodología:

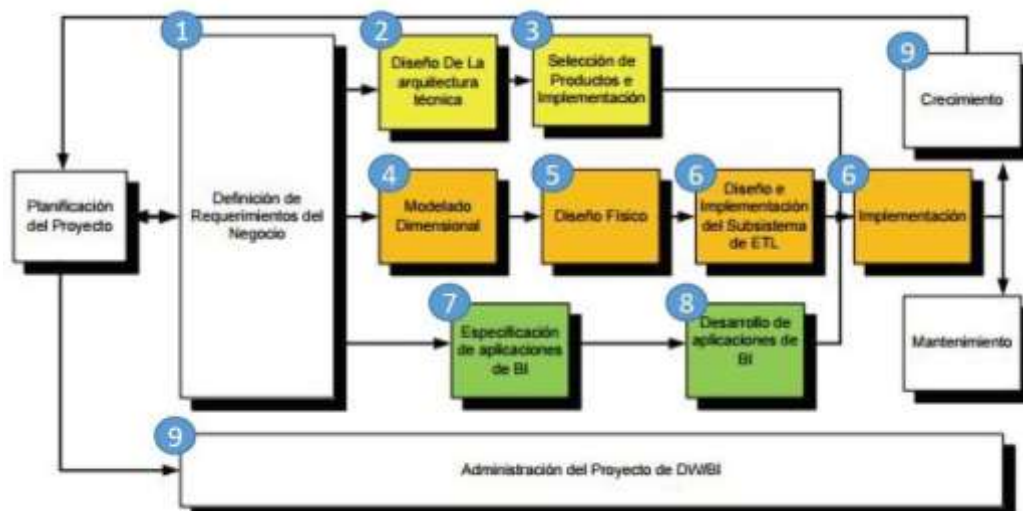


Fig. 1 Fases Ralph Kimball

Como se logra visualizar en la imagen, el input necesario para ejecutar de forma satisfactoria la metodología son los requisitos comerciales (**Fase #1: Definición de requerimientos del negocio**) y a partir de ahí, parten tres caminos que agrupan en su totalidad a las fases restantes.

El camino superior se enfoca en el ordenamiento de la tecnología que se empleará durante el desarrollo del proyecto.

- **Fase #2:** Diseño de la arquitectura técnica
- **Fase #3:** Selección de productos e implementación

El camino medio busca hacer uso de los datos con los que se cuenta y plantear la estrategia de transformación.

- **Fase #4:** Modelado dimensional
- **Fase #5:** Diseño físico
- **Fase #6:** Diseño e implementación del subsistema de ETL

El camino inferior se centra en la solución BI que se brindará.

- **Fase #7:** Especificación de aplicaciones de BI
- **Fase #8:** Desarrollo de aplicaciones de BI

Finalmente, como output se tendrá la documentación propiamente del DWBI que nos permitirá administrar el proyecto.

- **Fase #9:** Administración del Proyecto de DWBI

3.4.2. Producto acreditable

1. Interfaces

Para la construcción de las interfaces del sistema de inteligencia de negocios se hizo uso de la herramienta Power BI. Las mismas que se presentan en el *ítem 4.1.5. Fase #8: Desarrollo de aplicaciones BI, en el Capítulo IV. Resultados.*

2. Arquitectura

De diseñó una arquitectura idónea para el funcionamiento del sistema de inteligencia de negocios, el cual se detalla en el *ítem 4.1.5. Fase #2: Diseño de la arquitectura técnica, en el Capítulo IV. Resultados.*

3. Infraestructura tecnológica

Considerando la arquitectura anteriormente descrita, se definen las características de cada uno de sus componentes en el *ítem 4.1.5. Iteración #3: Selección de productos e implementación, en el Capítulo IV. Resultados.*

3.4.3. Manual de usuario

Se elaboró un manual de usuario con la finalidad de ayudar a los usuarios en el uso del sistema de inteligencia de negocios que se implementó, el cual se muestra en el *Anexo N° 03.*

3.5. Matriz de consistencia

Tabla III
MATRIZ DE CONSISTENCIA

<u>FORMULACIÓN DEL PROBLEMA</u>		<u>MÉTODOLÓGÍA DE INVESTIGACIÓN</u>			
¿Cómo se podría apoyar en la toma de decisiones comerciales de un minimarket de Chiclayo?		<u>TIPO DE INVESTIGACIÓN</u> Aplicada			
<u>OBJETIVO GENERAL</u>	<u>MÉTODO</u>	<u>DESCRIPCIÓN</u>			
Implementar un sistema de inteligencia de negocios que apoye en la toma de decisiones comerciales de un minimarket de Chiclayo.	Analítico	Este método se empleará con la finalidad de inicialmente analizar la situación actual de la empresa y posteriormente la información relevante para el sistema a proponer.			
	Deductivo	Este método permitirá determinar los indicadores relevantes para la toma de decisiones del negocio			
	Implementación	Este método permitirá desarrollar el sistema de inteligencia de negocio y ponerlo en funcionamiento.			
	<u>TÉCNICAS</u>	<u>INSTRUMENTOS</u>	<u>ELEMENTOS DE LA POBLACIÓN</u>	<u>PROPÓSITO</u>	
	Entrevista	Cuestionario	Jefe de ventas	Conocer las deficiencias en el proceso de toma de decisiones.	
Encuesta	Cuestionario	Expertos	Conocer la validez de la herramienta implementada		
<u>OBJETIVOS ESPECÍFICOS</u>	<u>DESCRIPCIÓN DEL LOGRO DE LOS OBJETIVOS ESPECÍFICOS</u>		<u>INDICADORES</u>		
Determinar los datos relevantes para comprender el comportamiento de las ventas en las diferentes sucursales del minimarket.	Listado de hechos y dimensiones		Total de hechos y dimensiones		
Analizar y limpiar la data resultante del sistema de ventas.	Datos normalizados		Total de datos atípicos igual a 0		
Construir reportes analíticos para dar soporte en la toma de decisiones comerciales del jefe de ventas.	Dashboards con información general		Total de dashboards		
	Dashboard de ventas		Total de KPIs		
Validar los reportes analíticos mediante el juicio de expertos.	Aprobación de la solución brindada		Coeficiente de V-Aiken		

3.6. Consideraciones éticas

La lista que se muestra a continuación es aquella que fue considerada para la protección del almacenamiento de los datos:

- ✓ Uso de técnicas de recolección de datos: Encuestas, entrevistas, etc.
- ✓ Protección de la información.
- ✓ Resguardo de los datos y secreto de la información.

IV. Resultados y discusión

4.1. En base a la metodología utilizada

4.1.1. Fase #1: Definición de requerimientos del negocio

En base a las entrevistas con el jefe de ventas, se llegó a la conclusión de que la información requerida en los dashboards, debería de contemplar las siguientes características y consideraciones.

Características

- Ingresos detallados en unidades monetarias
- Ingresos analizados por unidad de tiempo (año, trimestre, mes, semana, fecha).
- Niveles de detalle por empleado o personal de venta, tipo de documento y Tiempo.
- Detalle de los ingresos por cliente
- Detalle de los ingresos por producto
- Análisis de ingresos por tipo de venta.

Consideraciones

- Comparativo de ingreso entre unidades de tiempo
- Análisis de ingresos de uno o más años a la vez
- Suma acumulada de ingresos
- Participación porcentual de las variables de ingresos
- Productos estrella y productos hueso.

4.1.2. Fase #2: Diseño de la arquitectura técnica

La arquitectura que se consideró para la implementación del producto fue la siguiente:

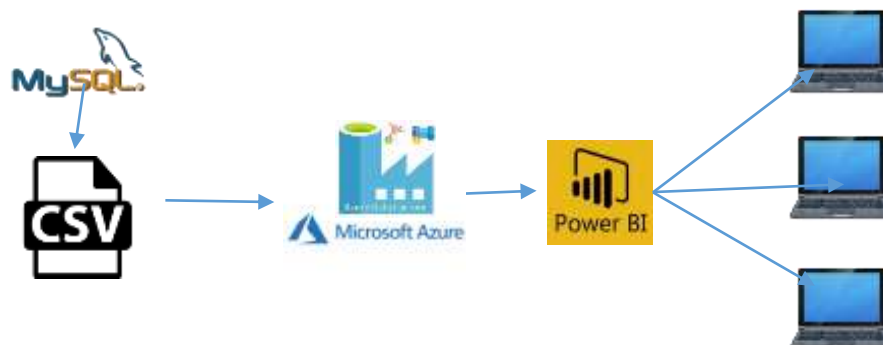


Fig. 2 Arquitectura lógica del sistema de inteligencia de negocio

- **Servidor de base de datos MySQL** = Contendrá la información del sistema de ventas usado por el minimarket.
- **CSV**: Formato en el cual se exportará la data para su fácil almacenamiento en la nube y acceso.
- **Microsoft Azure** = Plataforma cloud que permitirá tener acceso a las herramientas de Data Factory y Azure DB for SQL.
- **Azure DB for SQL** = Base de datos SQL en la nube en la cual se restaurará la BD del sistema de ventas usado por el minimarket.
- **Data Factory** = Servicio de integración de datos serverless que permitirá llevar a cabo el proceso ETL.
- **Power BI** = Herramienta que se conectará a la BD de datos en la nube y en la cual se implementarán los dashboards en base a los indicadores que se definan.
- **Usuario** = Jefe de ventas del minimarket que hará uso de los dashboards para brindar soporte a su proceso de toma de decisiones.

4.1.3. Fase #3: Selección de productos e implementación

- Servidor de datos: MySQL y SQL Server.
- Preparación de los datos: Data Factory.
- Presentación de los datos: Power BI

4.1.4. Fase #4: Modelado dimensional

- Diagrama StarNET

El siguiente diagrama se ha elaborado en base a los requerimientos definidos. Por lo cual se llegó a la conclusión de que era necesario contar con un Hecho_Ingresos, el cual sería alimentado de las dimensiones persona/empleo, producto, tipo de documento, tiempo, tipo de venta, sucursal y cliente.

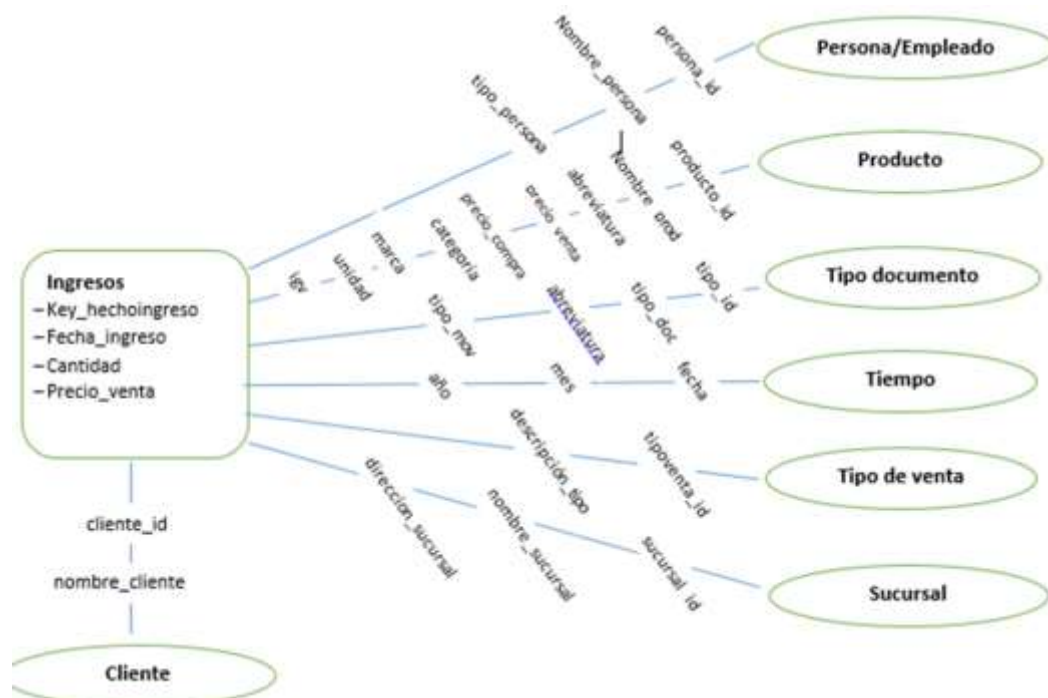


Fig. 3 Diagrama StarNet

4.1.5. Fase #5: Diseño físico

Para una correcta implementación del diseño físico del datamart, compuesto por el hecho ingresos y sus 7 dimensiones. Inicialmente se tuvo que evaluar la data contenida en el sistema de ventas, cuya base de datos está compuesta por 20 tablas que contienen principalmente información sobre los movimientos, sucursales, usuarios, producto, tipo de movimiento y otros. A continuación, se presenta el diagrama de base de datos actual:

En base al análisis de la data y el modelo StarNet propuesto, se diseñó la siguiente base de datos dimensional que brindó soporte a los criterios analíticos.

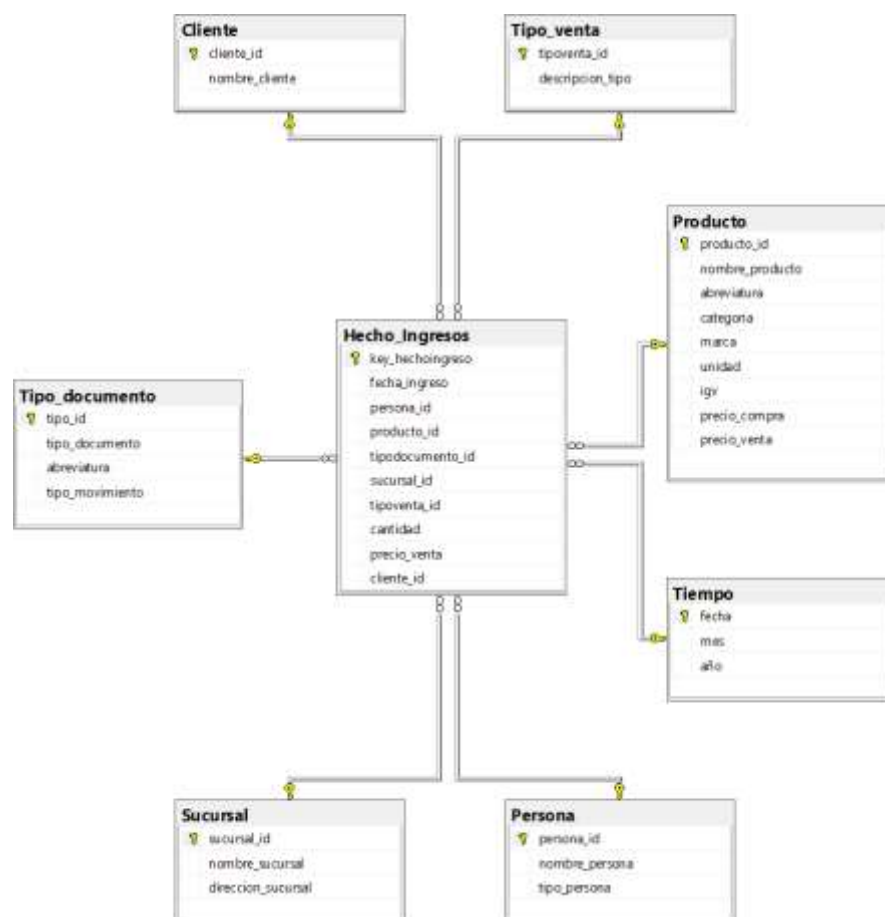


Fig. 5 Base de datos dimensional

A continuación, se detalla la procedencia de los datos y el tipo de datos que se espera.

	Dato	Tipo de dato	Null	Procedencia
dim_cliente	cliente_id	Serial primary key	No	person.id
	nombre_cliente	Varchar(50)	No	person.apellidopaterno person.apellidomaterno person.nombres
dim_tipo_venta	tipoventa_id	Serial primary key	No	autogenerado
	descripcion_tipo	Varchar(255)	No	movimiento.tipoventa
dim_tipo_documento	tipo_id	Serial primary key	No	tipodocumento.id
	tipo_documento	Varchar(50)	No	tipodocumento.nombre

	abreviatura	Varchar(5)	No	tipodocumento.abreviatura
	tipo_movimiento	int	No	tipomovimiento.nombre
dim_sucursal	sucursal_id	Serial primary key	No	sucursal.id
	nombre_sucursal	Varchar(50)	No	sucursal.nombre
	direccion_sucursal	Varchar(255)	No	sucursal.direccion
dim_persona	persona_id	Serial primary key	No	person.id
	nombre_persona	Varchar(100)	No	person.apellidopaterno person.apellidomaterno person.nombres
	tipo_persona	int	No	person.tipo
dim_tiempo	fecha	date primary key	No	movimiento.fecha
	mes	int	No	movimiento.month(fecha)
	año	int	No	movimiento.year(fecha)
dim_producto	producto_id	Serial primary key	No	producto.id
	nombre_producto	Varchar(50)	No	producto.nombre
	abreviatura	Varchar(50)	No	producto.abreviatura
	categoria	Varchar(50)	No	categoria.nombre
	marca	Varchar(50)	No	categoria.nombre
	unidad	Varchar(30)	No	unidad.nombre
	igv	Varchar(1)	No	producto.igv
	precio_compra	float	No	producto.precio compra
precio_venta	float	No	producto.precioventa	
hecho_ingresos	key_hechoingreso	int IDENTITY PK	No	autogenerado
	fecha_ingreso	date primary key	No	movimiento.fecha
	persona_id	int	No	movimiento.respnsable_id
	producto_id	int	No	detallemovimiento.producto_id
	tipodocumento_id	int	No	movimiento.tipodocumento_id
	sucursal_id	int	No	movimiento.sucursal_id
	tipoventa_id	int	No	movimiento.tipoventa
	cantidad	int	No	movimiento.cantidad
	precio_venta	float	No	detallemovimiento.precioventa
cliente_id	int	No	movimiento.persona id	

4.1.6. Fase #6: Diseño e implementación del subsistema de ETL

Esta actividad en el desarrollo de la solución de Business Intelligence permitió el poblamiento de la base dimensional o Data mart, para ello, se utilizó el Data Factory, servicio de Microsoft Azure.

Por el mismo diseño del Data mart, en la que las tablas dimensiones son referenciadas por la tabla hecho, se determinó primero la carga de las dimensiones y luego la carga de la fact table o tabla hecho.

Extracción para el poblamiento de las tablas dimensión

▪ Extracción para el poblamiento de la dimensión Producto

Para lograr poblar la dimensión Producto, se hizo uso de las tablas producto, unidad y marca. Se emplearon 3 left joins que permitieron pasar obtener los siguientes datos: producto.id, producto.nombre, producto.abreviatura, categoría.nombre como categoría, unidad.nombre, producto.igv, producto.precio compra y producto.precio venta. Una vez obtenidos los datos requeridos para poblar la dimensión, se procedió a efectuar un left join con la tabla de la dimensión, con la finalidad de que únicamente pasen los datos nuevos y, finalmente se filtraron aquellas filas que tuvieran la columna nombre del producto, vacía.

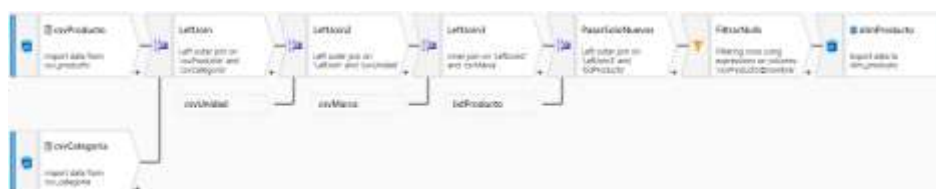


Fig. 6 Pipeline Dim_producto

▪ Extracción para el poblamiento de la dimensión Persona

Para poblar la dimensión Persona, se hizo uso de las tablas person y rolpersona. En primer lugar, se aplicó un inner join entre ambas tablas, tomando como valor común el id de la persona. Luego se procedió a usar un left join para impedir que pasen datos repetidos y finalmente, fue necesario que se

concatenen el person.apellidopaterno, person.apellidomaterno y person.nombres, que se filtre por el rol_id con valor 1 y que se agrupen con el objetivo de no tener valores repetidos.



Fig. 7 Pipeline Dim_persona

- **Extracción para el poblamiento de la dimensión Cliente**

El poblamiento de la dimensión cliente consistió en realizar un left join entre las tablas person y rolpersona. Al igual que en los casos anteriores, se validó que solo pasen datos nuevos y finalmente se procedió a concatenar los atributos person.apellidopaterno, person.apellidomaterno y person.nombres, se filtró por el rol_id con valor 3 y se agruparon los datos con el objetivo de no tener valores repetidos.



Fig. 8 Pipeline Dim_cliente

- **Extracción para el poblamiento de la dimensión Sucursal**

Poblar la dimensión sucursal consistió en seleccionar los atributos id, nombre y direccion de la tabla sucursal, validar que únicamente pasen los datos nuevos a través de un left join y finalmente filtrar que no pase ningún id que sea null.



Fig. 9 Pipeline Dim_sucursal

- **Extracción para el poblamiento de la dimensión Tiempo**

La dimensión tiempo se pobló haciendo uso de la columna fecha de la tabla movimiento. Para ello se agruparon, se validó que solo pasaran las nuevas fechas, se filtraron aquellas filas que contenían valores nulos y finalmente se crearon dos columnas derivadas con el objetivo de separar el año y el mes.



Fig. 10 Pipeline Dim_tiempo

- **Extracción para el poblamiento de la dimensión Tipo_documento**

Inicialmente se tuvo que hacer un left join entre las tablas tipodocumento (tdo) y tipomovimiento (tmo), obteniéndose los valores tdo.id, tdo.nombre, tdo.abreviatura y tmo.nombre. Luego se procedió con la validación que permite que solo los datos nuevos pasen y finalmente se corroboró que no pasen aquellas filas con valores nulos.



Fig. 11 Pipeline Dim_tipo_documento

- **Extracción para el poblamiento de la dimensión Tipo_venta**

Para poblar la dimensión tipo_venta, se procedió con la validación de que solo pasen valores nuevos, se agrupó por el tipoventa de la tabla movimiento y finalmente se filtraron aquellas filas con valores nulos.



Fig. 12 Pipeline Dim_tipo_venta

Extracción para el poblamiento de la tabla hecho

- **Extracción para el poblamiento de la tabla Hecho**

Para el poblamiento de la tabla Hecho, se llevó a cabo un inner join entre las tablas movimiento y detallemovimiento. Luego se procedió a realizar dos left join con la tabla rolpersona, la cual previamente fue duplicada con el objetivo de obtener un grupo de todos aquellos usuarios cuyo rol_id sea 1 y cuyo rol_id fuera 3. De dichas sub-consultas solo se obtuvieron los person_id.

Una vez obtenidos los datos movimiento.fecha, movimiento.responsable_id, detallemovimiento.producto_id, movimiento.tipodocumento_id, movimiento.sucursal_id, movimiento.tipoventa, movimiento.cantidad, detallemovimiento.precioventa y movimiento.persona_id se procedió con el filtrado por tipomovimiento_id = 2, se validó que únicamente pasen valores nuevos y se creó una columna derivada con el objetivo de separar los id de los clientes y de los empleados.



Fig. 13 Pipeline Hecho_ingresos

4.1.7. Fase #7: Especificación de aplicaciones de BI

Los diseños de los reportes se ajustaron al tipo de información o indicadores requeridos por los usuarios. También se consideró los visualizadores más adecuados para la información a presentar.

Se consideraron algunos dashboard para analizar la información por diferente criterio o perspectiva, es decir, analizar los ingresos por la perspectiva del producto, tiempo, empleado o vendedor, sucursal, etc., que corresponde a cada una de las dimensiones del modelo dimensional.

Se utilizó Power BI para el desarrollo de los dashboard y reportes analíticos y el servicio de Power BI Services de Microsoft para la publicación de la información y pueda ser accedida en la nube.

4.1.8. **Fase #8:** Desarrollo de aplicaciones de BI

Dashboard de Análisis general



Fig. 14 Dashboard de análisis general

Con este dashboard, se considera comparativos entre unidades de tiempo y evolución de ingresos.



Información que muestra el monto de ingreso del mes seleccionado y el ingreso del mes anterior. Así como el número de transacciones realizadas en el mes seleccionado y el número de transacciones del mes anterior.

KPIs, que muestran los ingresos y número de transacciones comparado con el mes anterior y paralelo.

En este gráfico se puede observar la evolución de los ingresos diarios del mes seleccionado.



Otro requerimiento general es visualizar los ingresos por la perspectiva o dimensión cliente, pudiendo observar que su mayor ingreso lo genera el público en general.



Otro requerimiento, es visualizar los ingresos por la dimensión Producto, en este gráfico, por el nivel de cada categoría. Se muestra el Top 20.

Dashboard de Análisis por el criterio de Producto



Fig. 15 Dashboard de Análisis por el criterio de Producto

Este requerimiento muestra los ingresos analizados por la perspectiva de Producto (Dimensión Producto), mostrando la variación de los ingresos de la categoría seleccionada, su evolución y la distribución de los productos de esa categoría comparado con el mes anterior y mes paralelo del año anterior.



Los indicadores presentados, son los mismos que el dashboard anterior, analizados bajo la dimensión Producto, en su nivel de categoría.

En este gráfico se puede observar la evolución de los ingresos diarios del mes y categoría seleccionada

Ingresos analizados bajo la dimensión Producto, en su nivel de granularidad (Producto), considerando los KPIs de variación con respecto al mes anterior y mes paralelo del año anterior.

Se muestra los ingresos y el número de unidades vendidas por la categoría seleccionada (Dimensión Producto) y por día de semana (Dimensión Tiempo).

Dashboard de ranking de ingresos

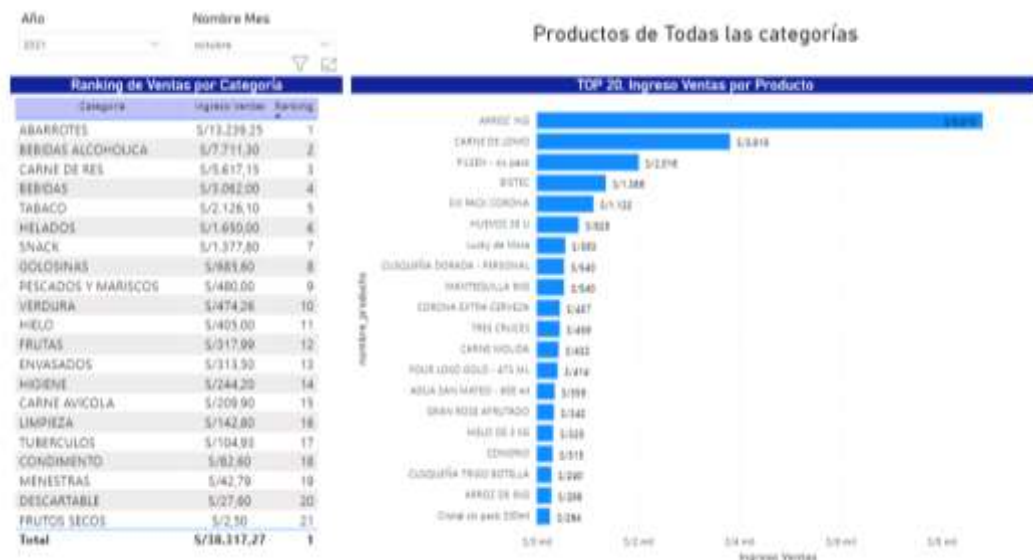


Fig. 16 Dashboard de Ranking de ingresos

En este reporte se muestra el ranking de ingresos por categoría y producto por año y mes seleccionado, considerando los productos estrella y productos hueso.

En la grilla de la izquierda se muestra el ranking por categoría, al seleccionar una de ella muestra en la parte derecha el TOP 20 de los productos que generan más ingresos de la categoría seleccionada.

Dashboard de tendencia de ingresos

Año: 2021

Tendencia de Ingresos Mensual por Categoría/Producto													
Categoría	1	2	3	4	5	6	7	8	9	10	11	12	Total
ASARROTES	\$/2,570	\$/4,037	\$/7,826	\$/3,964	\$/5,002	\$/6,992	\$/15,080	\$/11,019	\$/7,301	\$/13,239	\$/6,846	\$/6,155	\$/92,832
BEBIDAS ALCOHOLICA	\$/1,410	\$/7,238	\$/9,168	\$/4,038	\$/5,645	\$/6,519	\$/9,799	\$/6,432	\$/5,210	\$/7,711	\$/10,088	\$/7,031	\$/86,728
BEBIDAS	\$/1,031	\$/2,413	\$/3,631	\$/2,359	\$/2,349	\$/1,853	\$/3,402	\$/2,640	\$/2,210	\$/3,082	\$/3,170	\$/4,827	\$/33,068
CARNE DE RES	\$/396	\$/1,859	\$/1,879	\$/1,810	\$/2,189	\$/1,833	\$/4,481	\$/4,454	\$/1,181	\$/5,817	\$/1,391	\$/2,114	\$/29,664
VERDURA	\$/3,911	\$/3,217	\$/4,301	\$/1,341	\$/1,811	\$/2,078	\$/3,396	\$/1,330	\$/730	\$/474	\$/378	\$/136	\$/25,548
TABADO	\$/471	\$/1,444	\$/1,369	\$/1,105	\$/1,792	\$/2,285	\$/2,276	\$/1,640	\$/2,086	\$/2,126	\$/2,058	\$/2,388	\$/21,214
HELADOS	\$/746	\$/1,809	\$/1,865	\$/1,049	\$/994	\$/781	\$/1,443	\$/1,026	\$/625	\$/1,650	\$/1,808	\$/1,882	\$/15,594
SNACK	\$/461	\$/774	\$/1,211	\$/1,048	\$/1,168	\$/1,168	\$/1,388	\$/1,382	\$/960	\$/1,378	\$/1,512	\$/1,085	\$/14,823
PESCADOS Y MARISCOS	\$/775	\$/338	\$/1,620	\$/653	\$/1,883	\$/874	\$/1,318	\$/723	\$/537	\$/400	\$/713	\$/613	\$/10,940
FRUTAS	\$/1,321	\$/1,736	\$/1,081	\$/645	\$/745	\$/769	\$/1,314	\$/473	\$/438	\$/318	\$/262	\$/223	\$/9,708
GOCOSINAS	\$/62	\$/270	\$/474	\$/432	\$/66	\$/768	\$/727	\$/633	\$/533	\$/688	\$/778	\$/1,099	\$/7,143
MENESTRAS	\$/1,274	\$/950	\$/1,032	\$/380	\$/209	\$/683	\$/1,841	\$/777	\$/70	\$/43	\$/126	\$/27	\$/6,632
CARNE AVICOLA	\$/250	\$/713	\$/720	\$/317	\$/911	\$/708	\$/1,380	\$/541	\$/412	\$/210	\$/100	\$/247	\$/6,624
ENVASADOS	\$/212	\$/290	\$/545	\$/747	\$/684	\$/573	\$/533	\$/944	\$/432	\$/314	\$/335	\$/375	\$/5,456
HIGIENE	\$/145	\$/338	\$/342	\$/404	\$/319	\$/436	\$/429	\$/279	\$/321	\$/244	\$/488	\$/542	\$/4,507
HIELO	\$/46	\$/485	\$/587	\$/328	\$/211	\$/182	\$/472	\$/663	\$/326	\$/405	\$/224	\$/180	\$/4,463
TUBERCULOS	\$/840	\$/487	\$/627	\$/278	\$/311	\$/390	\$/681	\$/322	\$/101	\$/105	\$/93	\$/70	\$/4,150
LIMPIEZA	\$/144	\$/252	\$/685	\$/136	\$/389	\$/207	\$/632	\$/983	\$/219	\$/143	\$/140	\$/232	\$/3,563
EMBUTIDOS	\$/13	\$/111	\$/214	\$/114	\$/248	\$/231	\$/134	\$/28	\$/17	\$/171	\$/204	\$/154	\$/1,504
DESCARTABLE		\$/52	\$/246	\$/19	\$/63	\$/87	\$/177	\$/58	\$/16	\$/28	\$/35	\$/68	\$/880
FRUTOS SECOS	\$/20	\$/154	\$/291	\$/207	\$/49	\$/85	\$/22	\$/0	\$/19	\$/3	\$/5	\$/13	\$/849
CONDIMENTO	\$/96	\$/23	\$/69	\$/32	\$/97	\$/94	\$/93	\$/66	\$/38	\$/83	\$/46	\$/72	\$/657
Total	\$/18,937	\$/29,072	\$/40,478	\$/21,394	\$/28,334	\$/29,817	\$/51,110	\$/36,552	\$/24,025	\$/38,117	\$/31,150	\$/36,570	\$/385,754

Fig. 17 Dashboard de tendencia de ingresos

Otro requerimiento solicitado es la tendencia de ingresos por categoría y producto. Se implementó una jerarquía categoría/producto para mostrar dentro de una categoría que productos son los que aportan más ingresos.

Este reporte, considera el requerimiento de visualizar un detalle de los ingresos mensualmente por la dimensión Producto (categoría y producto)

Link de acceso: [Power BI](#)

4.1.9. **Fase #9:** Administración del Proyecto de DWBI

Entre los indicadores o Kpi's que se han implementado en los diferentes dashboard y reportes para analizar sus ingresos se han considerado los siguientes:

1. Indicadores de variación porcentual % con el mes anterior.
2. Indicadores de variación porcentual % con el mes paralelo del año anterior.
3. Indicadores de participación porcentual, para ver que tanto en porcentaje representa un mes, cliente, empleado con respecto al total.
4. Número de transacciones de ventas.
5. Evolución de ventas

4.2. Impactos esperados

4.2.1. Impactos económicos

Reducción de pérdidas económicas debido a una mala gestión de promociones y en su lugar, un aumento en los ingresos dado a una correcta toma de decisiones.

4.2.2. Impactos sociales

Clientes satisfechos debido a que sus necesidades se ven atendidas, no solo al momento de brindarles promociones acordes a sus necesidades, sino que además al tener información detallada de las ventas, el jefe podrá determinar qué productos no deben faltar.

4.2.3. Impactos en tecnología

La tecnología cloud permite que la información y modelos se tenga a disposición en cualquier momento y desde cualquier dispositivo. Se espera que posteriormente, el sistema transaccional también pueda ser migrado a la nube.

V. **Discusión**

Con la finalidad de que se comprenda en su totalidad el proceso de ventas de la compañía y se logre completar el primer objetivo de la investigación “Determinar los datos relevantes para comprender el comportamiento de las ventas en las diferentes sucursales del minimarket” se hizo necesario que llevara a cabo una entrevista con el jefe de ventas y además se analizara la data resultante de su sistema transaccional en busca de aquella información relevante que se posee pero no se está tomando en cuenta al momento de la toma de decisiones. Para ello se siguió lo planteado por Villanueva [18] al momento del uso de la metodología de Ralph Kimball en su etapa de definición de requerimientos y se obtuvo como resultado que eran 6 las características que se deberían de tener en cuenta, además de 6 consideraciones que darían como resultado un hecho ingresos y 7 dimensiones con sus respectivos niveles de información (Fig. 3).

El segundo objetivo de la investigación fue “Analizar y limpiar la data resultante del sistema de ventas” con el objetivo de inicialmente definir el estado actual en el que se encontraban los datos y en base a ello plantear una estrategia de limpieza que permitiera normalizar la data y reducir a 0 los datos atípicos. Con esto se confirmó lo indicado por Morales [11], es decir que la fase más importante en el desarrollo de un BI es el momento de la implementación del ETL dado que de esto dependerá cuán precisa es la información que se muestra. Para lograr cumplir con este objetivo, se hizo uso de Azure Data Factory dado que ofrece una mayor capacidad para procesar grandes volúmenes de data, lo cual se ve reflejado en la rapidez al momento de ejecutar las consultas de limpieza implementadas. Además, al usar tecnología cloud, la distribución de la solución fue más rápida como lo menciona [14] y se descartó el uso de otros repositorios de data, como el propuesto por [15] dado que Azure ofrece su propio blob storage

Para poder cumplir con el tercer objetivo “Construir reportes analíticos para dar soporte en la toma de decisiones comerciales del jefe de ventas”, se hizo uso de la herramienta Power BI integrada con la tecnología de Azure con el objetivo de brindar una fácil distribución de la información como lo indica [10] y además se construyeron 5 dashboards orientados a la información específica de las ventas,

además de 1 dashboard que contenía información general y le permitió al jefe de ventas tener una visión panorámica de su proceso. Finalmente, es necesario recalcar, que gracias a la herramienta seleccionada, se obtuvo la flexibilidad mencionada por [16] al momento de poder visualizar los dashboards tanto desde un aplicativo de escritorio como desde la versión web o móvil.

Finalmente, el último objetivo fue “Validar los reportes analíticos mediante el juicio de expertos” y para ello se llevó a cabo la aplicación de un cuestionario a 3 ingenieros de sistemas. Dicho cuestionario evaluó la propuesta desde 5 dimensiones: a) facilidad de aprendizaje, b) recuerdo en el tiempo, c) comprensibilidad, d) atraktividad y e) satisfacción. Las respuestas brindadas por los jueces permitieron determinar el coeficiente de V-Aiken, el cual tiene como propósito identificar cuán relevantes son los ítems evaluados en relación a un contenido. El resultado obtenido fue de 0.85, lo cual indica que los jueces estaban de acuerdo entre sí con las respuestas brindadas.

El coeficiente de fue calculado tanto de forma general, como por dimensión y los resultados obtenidos fueron los siguientes:

FACILIDAD DE APRENDIZAJE	0.68
RECUERDO EN EL TIEMPO	0.92
COMPRESIBILIDAD	0.88
ATRATIVIDAD	0.87
SATISFACCIÓN	0.92

Dichos resultados se traducen en que en dichas 5 dimensiones, los jurados se encuentran de acuerdo con la relevancia de estas para el producto. Es decir, se confirma que la herramienta tiene una alta facilidad para ser recordada en el tiempo, para ser comprendida, atractiva y es satisfactoria para el usuario; además, es de fácil aprendizaje, pero no en un alto grado (Anexo N° 03).

VI. Conclusiones

- 1.** El relevamiento de la información fue fundamental que se haga a detalle, dado que de éste dependía que los datos almacenados en el datamarte, correspondan a indicadores que fueran de utilidad, en este caso, para el jefe de ventas y que se mostrarían en los dashboards,
- 2.** La única forma de mostrar información relevante, es si los datos de ingreso son precisos y es por ello que el proceso de la implementación del ETL es uno de los más importantes. Gracias al Data Factory, la limpieza de la data se pudo llevar a cabo de forma de precisa y se pudo procesar con rapidez debido a los servidores en la nube de Azure.
- 3.** Fueron 4 los dashboards que se conectaron a las 7 dimensiones y 1 hecho, con la finalidad de mostrar las 6 características y consideraciones definidas durante el desarrollo del primer objetivo.
- 4.** En base al resultado obtenido por el coeficiente de V-Aiken, se puede concluir que la solución brindada es acorde a las necesidades de la empresa, es decir, que los dashboards muestran información que sirve de apoyo al jefe de ventas para la toma de decisiones.

VII. Recomendaciones

- 1.** Se recomienda evaluar otras tecnologías cloud y principalmente herramientas que den soporte a la elaboración del ETL como Data Prep de Trifacta.
- 2.** Para poder profundizar en la investigación, es recomendable implementar un modelo predictivo que permita tener una visión sobre las ventas que se pueden esperar a futuro.
- 3.** Se considera idóneo, la evaluación de nuevas metodologías emergentes, que se centran exclusivamente en desarrollo de BI y ya no tanto en minería de datos.
- 4.** Se recomienda considerar una api (interfaz de programación de aplicaciones) para el usuario, donde pueda personalizar las consultas.

VIII. Referencias

- [1] G. Martínez Luna, «Minería de datos: cómo hallar una aguja en un pajar,» *Ciencia*, pp. 18-28, 2011.
- [2] J. Lluís Cano, *Business intelligence: competir con información*, Banesto, Fundación Cultural, 2007.
- [3] C. Howson, *Business intelligence: estrategias para una implementación exitosa*, McGraw-Hill, 2012.
- [4] Tableau. A Salesforce Company, «Inteligencia de negocios o análisis de negocios. ¿Cuál es la diferencia y cuáles necesitas?,» [En línea]. Available: <https://www.tableau.com/es-mx/learn/articles/business-intelligence/bi-business-analytics>. [Último acceso: 18 Marzo 2022].
- [5] Marqués, «Beneficios del Business Intelligence en pequeñas y medianas empresas,» 26 Abril 2021. [En línea]. Available: <https://www.marquesme.com/beneficios-del-business-intelligence-en-pequenas-y-medianas-empresas>. [Último acceso: 18 Marzo 2022].
- [6] EY, «Analítica de datos, una de las tendencias digitales más fuertes que las empresas deben considerar en 2022,» 21 Diciembre 2021. [En línea]. Available: https://www.ey.com/es_pe/news/2021/12/analitica-datos-tendencias-digitales-fuertes-considerar. [Último acceso: 18 Marzo 2022].
- [7] INEI, «EN EL PERÚ EXISTEN MÁS DE 2 MILLONES 838 MIL EMPRESAS,» INEI, Lima, 2021.
- [8] C. Silva, «CADE 2021: Perú tiene la tasa más alta del mundo de empresas por habitantes,» *El Comercio*, 16 Noviembre 2021.
- [9] BBVA, «¿Cuál será el futuro de las pymes peruanas en 2021?,» 11 Diciembre 2020. [En línea]. Available: <https://www.bbva.com/es/pe/cual-sera-el-futuro-de-las-pymes-peruanas-en-2021/>. [Último acceso: 18 Marzo 2022].
- [10] F. Zapata Barreno, «Aplicación móvil con implementación de un modelo de inteligencia de negocios para el proceso de facturación de la ferretería “Del Ahorro” usando la herramienta Microsoft Power BI,» Universidad Técnica del Norte, Ecuador, 2019.
- [11] S. Morales, «Metodología para procesos de inteligencia de negocios con mejoras en la extracción y transformación de fuentes de datos, orientado a la toma de decisiones,» Universidad de Alicante, Alicante, 2019.
- [12] J. F. Coba Medina y E. A. Revelo Paz, «Implementación de una solución de inteligencia de negocios para el análisis de datos relacionados con los proyectos de software del sector público en el Ecuador en la última década,» UNIVERSIDAD POLITÉCNICA SALESIANA, Quito, 2021.
- [13] L. B. Bellido Camacho, «Propuesta de mejora en la toma de decisiones aplicando business intelligence caso: agencia de promoción de inversiones.,» Lima, 2019.
- [14] D. Acosta Lescano y S. Vásquez Ávila, «Analítica de datos para el soporte en la toma de decisiones en el área de distribución y ventas de la Distribuidora Farmacéutica la Libertad S.R.L. utilizando Microsoft Azure y la metodología de Larissa Moss,» Universidad Privada Antenor Orrego, Trujillo, 2020.

- [15] J. A. Pereyra Salazar, «Implementación de business intelligence para la optimización de toma de decisiones en la gerencia de operaciones en una empresa de instalación de sistemas contra incendios,» UNMSM, Lima, 2021.
- [16] E. Nauca Torres y J. Gómez Martínez, «Implementación de una plataforma de Business Intelligence basado en análisis multidimensional para monitorear el comportamiento de casos COVID-19 en el Perú, periodo marzo-julio 2020,» Universidad de Lambayeque, Lambayeque, 2020.
- [17] D. J. Fuentes Adrianzén, «Modelo integrado de inteligencia de negocios como soporte a la toma de decisiones en la gestión comercial de las MYPES,» USS, Pimentel, 2021.
- [18] S. M. Villanueva Callirgos, «Implementación de una solución de inteligencia de negocios para apoyar la toma de decisiones en el proceso de compra y venta en una empresa farmacéutica en la ciudad de Chiclayo,» Universidad Católica Santo Toribio de Mogrovejo, Chiclayo, 2019.
- [19] Microsoft, «Microsoft,» 2020. [En línea]. Available: <https://info.microsoft.com/ww-landing-2020-gartner-magic-quadrant-for-analytics-and-business-intelligence.html?LCID=EN-US>.
- [20] C. J. Date, Introducción a los sistemas de bases de datos, México: Pearson Educación, 2001.
- [21] J. Martínez Mostazo, UF1890 - Desarrollo de componente software y consultas dentro del sistema de almacén de datos, España: Editorial Elearning , 2015.
- [22] MakeSoft Technologies, «Cloud Computing: breve historia,» [En línea]. Available: <https://www.makesoft.es/cloud-computing-breve-historia/>.
- [23] Microsoft, «Microsoft,» 2021. [En línea]. Available: <https://docs.microsoft.com/es-es/azure/guides/developer/azure-developer-guide>.
- [24] Microsoft, «¿Qué es Azure Database for MySQL?,» 2022. [En línea]. Available: <https://docs.microsoft.com/es-es/azure/mysql/overview>. [Último acceso: 2022 03 18].
- [25] Microsoft, «Microsoft,» 28 Setiembre 2021. [En línea]. Available: <https://docs.microsoft.com/en-us/azure/data-factory/introduction>.
- [26] J. Villegas, «Implementación de un Datamart como solución de inteligencia de negocios bajo la metodología de Ralph Kimball para la empresa Corporación Corrales SAC-LIMA,» Lima, 2018.
- [27] R. Hernández Sampieri, C. Fernández Colla y P. Baptista Lucio, Metodología de la investigación, McGraw-Hill Education, 2014.
- [28] J. Salkind, de *Metodos de investigación*, México, Prentice Hall, 1999.
- [29] R. Rizo, «Metodología para procesos de Inteligencia de Negocios con mejoras en la extracción y transformación de fuentes de datos, orientado a la toma de decisiones,» Alicante, 2019.

IX. Anexos

Anexo N° 01. Instrumentos de recolección de datos

Objetivo de la entrevista: Determinar las falencias que se tienen en el área de ventas, principalmente lo relacionado con la toma de decisiones.

Cargo: Jefe de ventas

1. ¿Actualmente cuentan con algún sistema para el control de las ventas?
A raíz de la pandemia, nos vimos forzados a “modernizarnos” y por ende contratamos un sistema comercial que nos permite registrar los productos, precios, categorías, entre otros datos relevantes. Además, claro está, contamos con el registro del detalle de la venta como tal y el stock de cada producto.
2. ¿En qué les beneficia dicho sistema?
Gracias a este sistema hemos podido comprender un poco más sobre los productos que estamos vendiendo y la acogida de los mismos, pero lamentablemente el sistema no nos permite obtener reportes específicos. Solo nos brinda información general que resulta muy difícil de procesar y eso se refleja en los desaciertos que estamos teniendo en nuestras estrategias últimamente.
3. ¿Cuáles son los retos que tiene como jefe de ventas?
Intento proponer estrategias de ventas, es decir ofertas que podrían causar un gran impacto en las ventas, pero ese proceso no está siendo muy preciso. Cuando la estrategia es la correcta, nuestras ventas pueden llegar a aumentar hasta en un 20%, pero cuando no lo es, las pérdidas logran superar hasta el 50% de lo que se esperaba en el día. Muchas veces nos hemos sobre stockeado para ofrecer las promociones y si éstas no son exitosas, se nos hace muy complicado recuperarnos.
4. ¿Cuenta con medidas que le permitan medir el estado de su área y éxito de sus planes?
No tenemos indicadores propiamente, pero contamos con un Excel en el cual vamos registrando información relevante a las estrategias que implementamos y cuál fue su impacto.
5. ¿Considera importante contar con una herramienta que le dé soporte a su proceso de toma de decisiones?
En definitiva, si tuviéramos algo que nos permitiera visualizar toda la información, que recabamos con nuestro sistema comercial, de una forma más ordenada y precisa, estoy seguro de que las ventas aumentarían y los errores que cometo al momento de definir las estrategias, serían menores. Además, de que

podría establecer indicadores precisos que me permitieran medir con exactitud el estado de mi área y gestión.

Objetivo de la encuesta: Determinar la validez y usabilidad de la herramienta.

Encuestado: Ingeniero experto.

	No aplica	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
VALOR	0	1	2	3	4

CRITERIOS DE USABILIDAD	NA	MED	ED	DE	MDE
FACILIDAD DE APRENDIZAJE	Escala de likert				
El software muestra resultados fáciles de entender					
No existe un punto del cual no pueda salir el usuario y por ende puede navegar por el sistema con tranquilidad					
Es posible retroceder en el software para poder modificar los datos ingresados					
El software cuenta con mensajes de alerta que le permitan al usuario completar las acciones de forma adecuada					
El usuario puede identificar fácilmente las figuras, tablas, elementos clickeables y el tipo de acción que debe ejecutar cada elemento del software					
RECUERDO EN EL TIEMPO	Escala de likert				
Recordar el funcionamiento del software no requiere de mucho esfuerzo por parte del usuario					
Es muy fácil recordar cómo ejecutar las acciones disponibles dentro del software					
COMPRESIBILIDAD	Escala de likert				
El texto contenido en el software es entendible y mantiene uniformidad a través de todas sus secciones					
No es necesario un mayor esfuerzo para comprender el significado de las imágenes e íconos que se usan en el software					
Se hace uso de un lenguaje amigable, familiar y/o cercano en el software					
Los tipos y tamaños de letra usados dentro del software son legibles y distinguibles					
ATRACTIVIDAD	Escala de likert				
Los títulos de las figuras y tablas son íntegramente descriptivos y distintivos					
Las interfaces no están atiborradas de información desordenada					
El software utiliza notificaciones o emergentes para mostrar información importante					
Se hace uso de una adecuada posición de los elementos (botones, mensajes, información, etc.) dentro del software					
La apariencia del software es visualmente agradable y sencilla					
SATISFACCIÓN	Escala de likert				
La utilidad del software es auténtica para el usuario					
La herramienta no provoca frustración en el usuario al ser usada					
El software cumple con todas las expectativas que el usuario tiene sobre éste					
El usuario considera útil la repetición del uso de la herramienta para sus labores					

Anexo N° 02. Respuestas juicio de expertos

CRITERIOS DE USABILIDAD	NA	MED	ED	DE	MDE
FACILIDAD DE APRENDIZAJE Escala de likert					
El software muestra resultados fáciles de entender	0	0	0	1	2
No existe un punto del cual no pueda salir el usuario y por ende puede navegar por el sistema con tranquilidad	0	0	0	2	1
Es posible retroceder en el software para poder modificar los datos ingresados	3	0	0	0	0
El software cuenta con mensajes de alerta que le permitan al usuario completar las acciones de forma adecuada	0	0	0	3	0
El usuario puede identificar fácilmente las figuras, tablas, elementos clickeables y el tipo de acción que debe ejecutar cada elemento del software	0	0	0	1	2
RECUERDO EN EL TIEMPO Escala de likert					
Recordar el funcionamiento del software no requiere de mucho esfuerzo por parte del usuario	0	0	0	2	1
Es muy fácil recordar cómo ejecutar las acciones disponibles dentro del software	0	0	0	2	1
COMPREENSIBILIDAD Escala de likert					
El texto contenido en el software es entendible y mantiene uniformidad a través de todas sus secciones	0	0	0	2	1
No es necesario un mayor esfuerzo para comprender el significado de las imágenes e iconos que se usan en el software	0	0	0	1	2
Se hace uso de un lenguaje amigable, familiar y/o cercano en el software	0	0	0	1	2
Los tipos y tamaños de letra usados dentro del software son legibles y distinguibles	0	0	0	2	1
ATRACTIVIDAD Escala de likert					
Los títulos de las figuras y tablas son integralmente descriptivos y distintivos	0	0	0	1	2
Las interfaces no están atiborradas de información desordenada	0	0	0	1	2
El software utiliza notificaciones o emergentes para mostrar información importante	0	0	0	2	1
Se hace uso de una adecuada posición de los elementos (botones, mensajes, información, etc.) dentro del software	0	0	0	2	1
La apariencia del software es visualmente agradable y sencilla	0	0	0	2	1
SATISFACCIÓN Escala de Likert					
La utilidad del software es auténtica para el usuario	0	0	0	0	3
La herramienta no provoca frustración en el usuario al ser usada	0	0	0	1	2
El software cumple con todas las expectativas que el usuario tiene sobre éste	0	0	0	1	2
El usuario considera útil la repetición del uso de la herramienta para sus labores	0	0	0	2	1

Las respuestas dadas por los jueces, fueron tabuladas y contabilizadas en base a la siguiente tabla leyenda:

	No aplica	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
VALOR	1	2	3	4	5

Los resultados que se obtuvieron fueron los siguientes:

JUEZ	FACILIDAD DE USO					RECUERDO EN EL TIEMPO	
	P1	P2	P3	P4	P5	P6	P7
JUEZ 1	4	4	1	4	4	5	5
JUEZ 2	5	4	1	4	5	4	5
JUEZ 3	5	5	1	4	5	5	4

JUEZ	COMPRESIBILIDAD				ATRACTIVIDAD						SATISFACCIÓN			
	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	
JUEZ 1	4	5	5	5	5	4	4	4	4	5	4	5	4	
JUEZ 2	5	5	5	4	4	5	5	4	5	5	5	4	5	
JUEZ 3	4	4	4	4	5	5	4	5	4	5	5	5	4	

Una vez hallados los puntajes, se calculó el valor mínimo y el nuevo número de categorías:

Valor mínimo	1	1
Número categorías	5	4

Tabla IV
Coeficiente de V-Aiken

JUEZ	FACILIDAD DE USO					RECUERDO EN EL TIEMPO		COMPRESIBILIDAD				ATRACTIVIDAD					SATISFACCIÓN			
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20
JUEZ 1	0.75	0.75	0	0.75	0.75	1	1	0.75	1	1	1	1	0.75	0.75	0.75	0.75	1	0.75	1	0.75
JUEZ 2	1	0.75	0	0.75	1	0.75	1	1	1	1	0.75	0.75	1	1	0.75	1	1	1	0.75	1
JUEZ 3	1	1	0	0.75	1	1	0.75	0.75	0.75	0.75	0.75	1	1	0.75	1	0.75	1	1	1	0.75
PROMEDIO INDIVIDUAL	0.92	0.83	0.00	0.75	0.92	0.92	0.92	0.83	0.92	0.92	0.83	0.92	0.92	0.83	0.83	0.83	1.00	0.92	0.92	0.83
PROMEDIO PARCIAL	0.68					0.92		0.88				0.87					0.92			
COEFICIENTE V-IKEN	0.85																			

Anexo N° 03. Manual de usuario

1. Implementación del sistema

A. Requerimientos de Hardware

Para que el sistema funcione y rinda adecuadamente, se deberá de cumplir con los siguientes requerimientos a nivel de hardware que a continuación se detallan:

Sugerido de Microprocesador

- Idóneo: Procesador Intel Core I7
- Mínimo: Procesador Intel Core I3

Sugerido de Memoria RAM

- Idóneo: 8 Gb
- Mínimo: 4 Gb

De no contarse con estos requerimientos, el sistema podrá funcionar pero se deberá de tener en cuenta que pueden tenerse resultados erróneos.

B. Requerimiento de Software

Para que el sistema funcione y rinda adecuadamente, los requerimientos de software a cumplir son los siguientes:

Sistema Operativo

- Microsoft Win10 o superiores.
- Linux (Todas las versiones)

Navegador Web

- Mozilla Firefox versión 100 a +
- Google Chrome versión 101 a +



C. Ruta de acceso

El enlace principal de ingreso al sistema es: [link](#)

2. Configuraciones cloud

Para poder implementar el sistema, previamente se debió de habilitar la plataforma cloud que daría soporte, Para ello se llevaron a cabo las siguientes configuraciones en Azure:

Creación de un grupo de recursos llamado tesisTatiana, el cual se encuentra en la región “**West US 2**”. En este, se encuentran contenidos los siguientes servicios:

Name ↑	Type ↑↓	Location ↑↓
 containertatiana	Storage account	West US 2
 dbtesistatiana (servertatiana/dbtesistatiana)	SQL database	East US 2
 dftesistatiana	Data factory (V2)	West US 2
 servertatiana	SQL server	East US 2

A continuación, se describirán las características de cada servicio:

- Blob Storage (containertatiana): es el repositorio en el cual se encuentran alojados los archivos .csv que se usan para el ETL (data factory)

Name	Last modified	Public access level	Lease state
<input type="checkbox"/> \$logs	5/13/2022, 12:49:15 PM	Private	Available
<input type="checkbox"/> caja	5/13/2022, 12:50:16 PM	Private	Available
<input type="checkbox"/> categoria	5/13/2022, 12:50:26 PM	Private	Available
<input type="checkbox"/> concepto	5/13/2022, 12:50:30 PM	Private	Available
<input type="checkbox"/> detallmovimiento	5/13/2022, 12:50:56 PM	Private	Available
<input type="checkbox"/> marca	5/13/2022, 12:51:03 PM	Private	Available
<input type="checkbox"/> movimiento	5/13/2022, 12:51:09 PM	Private	Available
<input type="checkbox"/> person	5/13/2022, 12:51:17 PM	Private	Available
<input type="checkbox"/> producto	5/10/2022, 10:35:14 PM	Private	Available
<input type="checkbox"/> rol	5/13/2022, 12:51:21 PM	Private	Available
<input type="checkbox"/> rolpersona	5/13/2022, 12:51:24 PM	Private	Available
<input type="checkbox"/> sucursal	5/13/2022, 12:51:32 PM	Private	Available
<input type="checkbox"/> tipodocumento	5/13/2022, 12:51:38 PM	Private	Available
<input type="checkbox"/> tipomovimiento	5/13/2022, 12:51:42 PM	Private	Available
<input type="checkbox"/> unidad	5/13/2022, 12:51:50 PM	Private	Available

- Servidor SQL (servertatiana): debido a que el datamart se encuentra alojado en una base de datos SQL, fue necesario contar con un servidor en el cual alojarla.

Essentials		Server admin	
Resource group (move)	: tesisTatiana	Server admin	: adminf
Status	: Available	Networking	: Show networking settings
Location	: East US 2	Active Directory admin	: Not configured
Subscription (move)	: Azure for Students	Server name	: servertatiana.database.windows.net

Este servidor cuenta con una configuración de red que permite impedir la conexión de dispositivos no autorizados:

Firewall rules

Allow certain public internet IP addresses to access your resource. [Learn more](#)

+ Add your client IPv4 address (2800:200:f100:214:d43f:5add:6f75:97f2) + Add a firewall rule

Rule name	Start IPv4 address	End IPv4 address
ClientIPAddress_2022-5-22_9-46-28	190.236.31.53	190.236.31.53
query-editor-91ab0b	179.6.46.253	179.6.46.253

- **BD SQL (dbtesistatiana):** es una base de datos SQL con las siguientes características:

Essentials		Server name	
Resource group (move)	tesistatiana	Server name	tesistatiana.database.windows.net
Status	Online	Elastic pool	No elastic pool
Location	East US 2	Connection strings	Show database connection strings
Subscription (move)	Azure for Students	Pricing tier	Basic
Subscription ID		Earliest restore point	2022-05-20 04:50 UTC

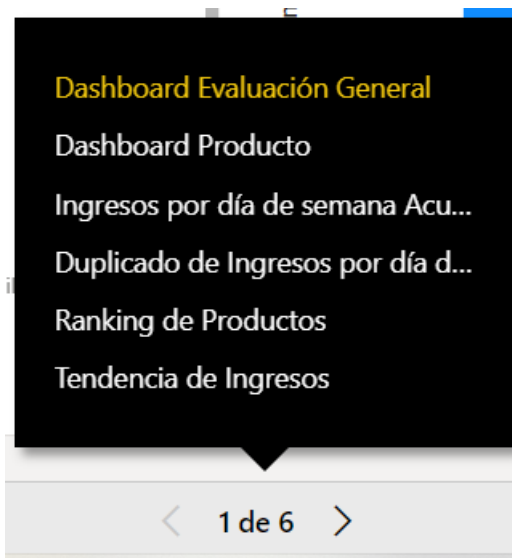
- **Data Factory (dfesistatiana):** está compuesto por 1 pipeline general que permite la ejecución de todos los data flows. Además, 2 grupos de datasets, los cuales generan la conexión entre los archivos csv ubicados en el container y las tablas de la base de datos sql. Finalmente, se implementó un data flow por cada dimensión y hecho que se debía poblar.

Factory Resources	
Filter resources by name	
Pipelines	1
PoblarDimyHecho	
Datasets	22
bd_sql	8
CSV_BlobStorage	14
Data flows	8
hecho_Ingresos	
Dim_cliente	
Dim_persona	
Dim_producto	
Dim_sucursal	
Dim_tiempo	
Dim_tipo_documento	
Dim_tipo_venta	

3. Interacción con los dashboards

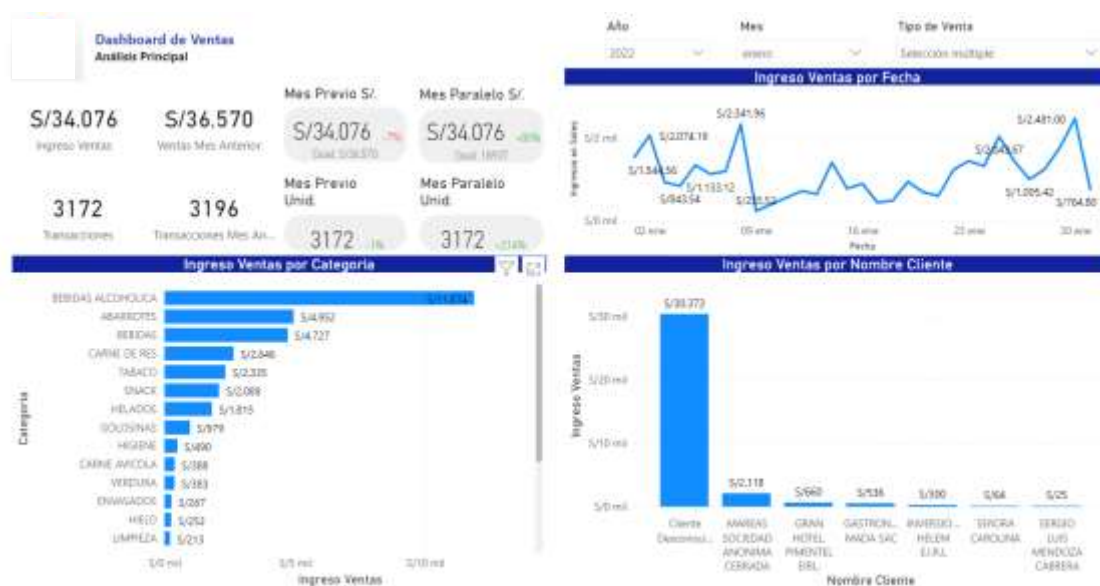
En la parte inferior de la interfaz inicial, se puede visualizar una barra que indica el total de reportes y dashboards con los que se puede interactuar.

En total son 6 las páginas por las que se podrá navegar para acceder a distinta información relevante para la toma de decisiones.



La primera página nos mostrará el dashboard de evaluación general de las ventas, el cual se divide en 4 secciones que se ven afectadas por los filtros de año, mes y tipo de venta.

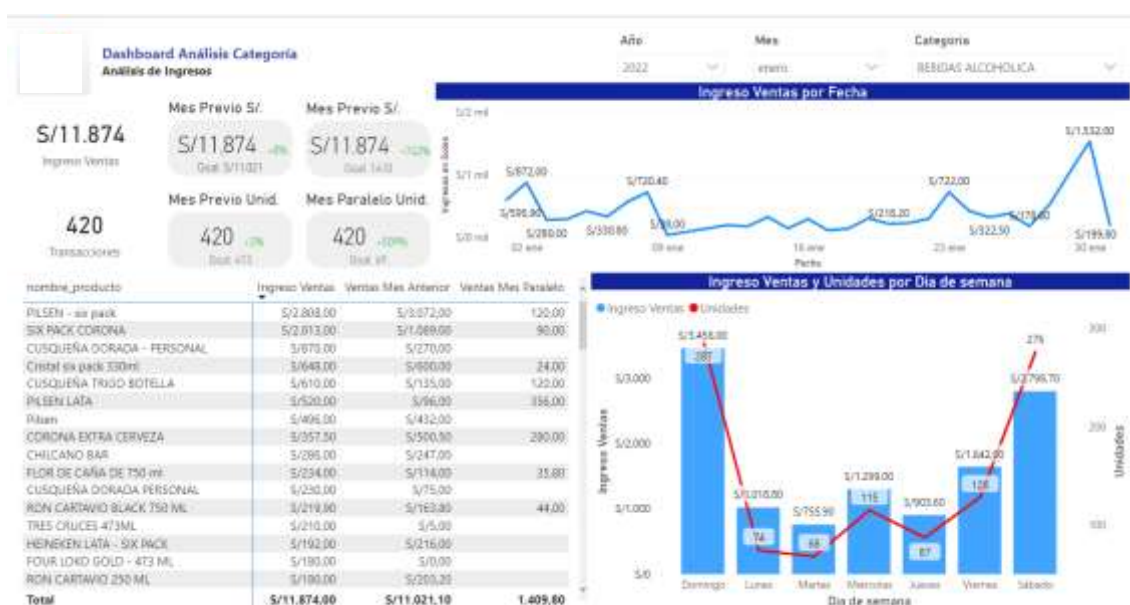
La sección ubicada en la parte superior izquierda, muestra el resumen de ingresos obtenidos por las ventas correspondientes a los filtros seleccionados, además del total de transacciones, ventas y transacciones del mes anterior con el objetivo de llevar a cabo una comparación y finalmente el porcentaje de crecimiento o disminución.



En las siguientes secciones podremos apreciar el ingreso de ventas por categoría, las ventas por fecha y finalmente las ventas por cliente.

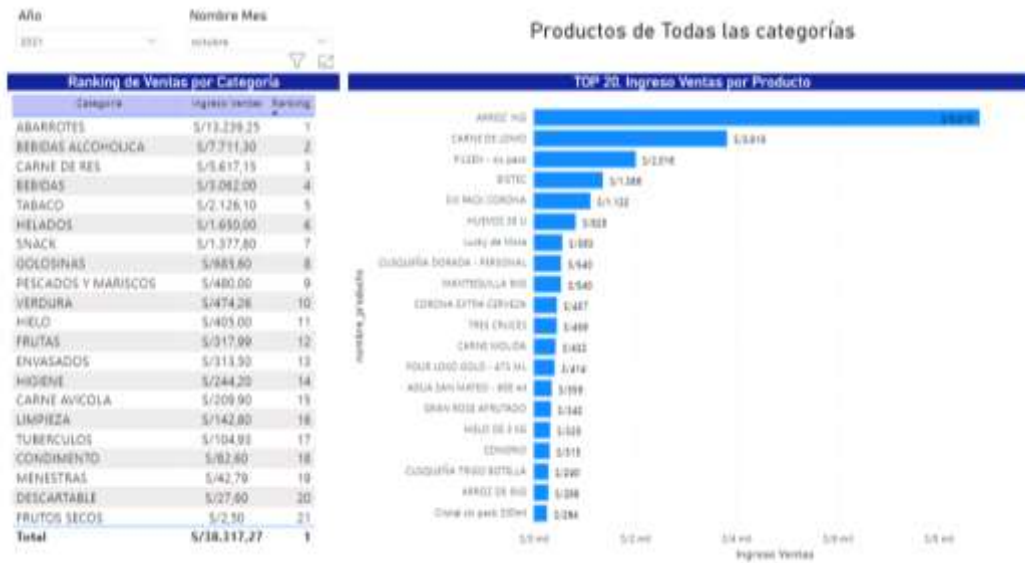
La segunda página nos muestra un dashboard orientado a los productos. El cual al igual que el anterior, depende de los filtros año, mes y categoría, para poder mostrar la información.

En la primera sección también encontraremos un resumen del ingreso de ventas, transacciones y su comparación con los períodos anteriores. Pero, además, podemos encontrar un gráfico de líneas que nos muestra el ingreso de las ventas según la fecha, un ranking de los productos que más se han vendido y finalmente, una gráfica que nos permite visualizar la relación existente entre la cantidad de productos vendidos y cuánto representa en soles.



La tercera página contiene un cuadro del acumulado de ventas de la semana, en el cual, al seleccionar cada categoría, se podrán visualizar los productos que la conforman. Finalmente, se debe de tener en cuenta que en la parte superior izquierda, se cuenta con un filtro de fechas.

En la quinta página podemos apreciar que existen dos filtros en la parte superior, el del año y el del nombre del mes. Además, contamos con un listado de categorías ordenadas de mayor a menor en base a sus ventas, que a la vez también funciona como filtro dado que si seleccionamos alguna, el gráfico de barras se verá modificado y mostrará únicamente aquellos productos que la componen.



En la última página, se podrá visualizar el filtro del año y a la vez una tabla que permitirá visualizar si ha habido un crecimiento o decrecimiento de las ventas a lo largo de los meses y la vez también se puede ver el mismo detalle a nivel de productos.

Año: 2021

Tendencia de Ingresos Mensual por Categoría/Producto

Categoría	1	2	3	4	5	6	7	8	9	10	11	12	Total
ABARRIOTES	5/2.570	5/4.037	5/7.828	5/13.884	5/19.802	5/26.992	5/35.080	5/43.219	5/51.301	5/59.339	5/67.446	5/75.513	5/83.632
BEBIDAS ALCOHÓLICA	5/1.410	5/7.238	5/9.188	5/4.028	5/9.643	5/6.519	5/9.789	5/8.452	5/9.212	5/7.711	5/10.598	5/11.031	5/86.728
BEBIDAS	5/1.231	5/2.413	5/3.631	5/2.359	5/2.349	5/1.893	5/3.402	5/2.640	5/2.210	5/3.062	5/3.170	5/4.827	5/33.068
CARNE DE RES	5/3.996	5/1.888	5/1.978	5/1.810	5/2.189	5/1.893	5/4.461	5/4.454	5/1.181	5/5.877	5/1.281	5/2.114	5/29.884
VERDURA	5/3.811	5/3.217	5/4.501	5/1.341	5/1.611	5/2.078	5/3.399	5/1.550	5/750	5/474	5/378	5/138	5/25.548
TABACO	5/471	5/1.444	5/1.583	5/1.105	5/1.792	5/2.285	5/2.276	5/1.840	5/2.086	5/2.126	5/2.058	5/2.388	5/21.214
HELADOS	5/748	5/1.829	5/1.363	5/1.049	5/994	5/791	5/1.443	5/1.226	5/925	5/1.850	5/1.808	5/1.392	5/15.594
SNACK	5/481	5/774	5/1.211	5/1.048	5/1.188	5/1.188	5/1.368	5/1.082	5/960	5/1.379	5/1.512	5/1.895	5/14.823
PESCADOS Y MARISCOS	5/775	5/108	5/1.620	5/653	5/1.880	5/874	5/1.318	5/722	5/537	5/400	5/712	5/613	5/10.940
FRUTAS	5/1.321	5/1.736	5/1.081	5/645	5/745	5/789	5/1.214	5/473	5/439	5/318	5/282	5/225	5/9.709
GOLOSINAS	5/62	5/270	5/474	5/432	5/606	5/766	5/727	5/633	5/533	5/688	5/778	5/1.099	5/7.143
MEÑESTRAS	5/1.274	5/950	5/1.032	5/380	5/229	5/663	5/1.841	5/777	5/710	5/43	5/126	5/27	5/6.832
CARNE AVICOLA	5/350	5/713	5/720	5/317	5/951	5/723	5/1.380	5/541	5/412	5/210	5/100	5/247	5/6.624
ENVASADOS	5/212	5/290	5/346	5/747	5/684	5/673	5/533	5/944	5/432	5/314	5/325	5/375	5/5.456
HIGIENE	5/145	5/338	5/542	5/404	5/319	5/436	5/429	5/279	5/321	5/244	5/488	5/562	5/4.507
HIELO	5/46	5/432	5/587	5/326	5/211	5/182	5/672	5/863	5/326	5/405	5/234	5/180	5/4.463
TUBERCULOS	5/940	5/437	5/627	5/278	5/211	5/380	5/681	5/222	5/101	5/105	5/93	5/73	5/4.150
LIMPIEZA	5/144	5/252	5/445	5/156	5/369	5/207	5/632	5/963	5/219	5/143	5/140	5/232	5/3.563
ENVASADOS	5/13	5/111	5/214	5/114	5/248	5/251	5/134	5/28	5/17	5/171	5/254	5/1.594	
DESCARTABLE	5/52	5/246	5/19	5/83	5/87	5/777	5/58	5/16	5/28	5/35	5/88	5/880	
FRUTOS SECOS	5/28	5/154	5/291	5/207	5/49	5/85	5/22	5/2	5/19	5/3	5/5	5/13	5/848
CONDIMENTO	5/36	5/23	5/88	5/32	5/87	5/34	5/53	5/66	5/38	5/83	5/46	5/72	5/657
Total	5/18.557	5/29.872	5/48.478	5/21.394	5/28.334	5/29.817	5/51.110	5/36.552	5/24.825	5/38.317	5/31.150	5/36.578	5/385.754