División de Posgrado
Trabajo de Titulación

“Aplicación de técnicas de inteligencia de negocios para segmentación de pagos online en una entidad financiera”

Proyecto de Desarrollo que se presenta como opción al título de Magister en Sistemas de Información con mención en Inteligencia de Negocios.

Autor:
Mario Pincay Cedeño.

Tutor:
Fabricio Echeverría

Guayaquil
2019

Dedicatoria

Primeramente, darle gracias a Dios que siempre guía mi camino, me da fortaleza y entendimiento para lograr mis objetivos deseados.

A mis padres Cruz Cedeño y Nicolás Pincay, quienes no dudaron ni un segundo en darme todo su apoyo, para hacer realidad todos mis sueños.

A mis hermanos Jessica, Lorena y Fernando que siempre están apoyándome en todos mis proyectos que me propongo.

Agradecimientos

A todos los docentes de la Universidad y personal en general que de alguna u otra manera me brindaron su apoyo incondicional en las diversas etapas de mi vida universitaria y profesional.

Universidad Ecotec

TRABAJO DE TITULACIÓN

CERTIFICACION DE REVISION FINAL

QUE EL PRESENTE PROYECTO DE DESARROLLO TITULADO:

APLICACIÓN DE TÉCNICAS DE INTELIGENCIA DE NEGOCIOS PARA SEGMENTACIÓN DE PAGOS ONLINE EN UNA ENTIDAD FINANCIERA

, FUE REVISADO, SIENDO SU CONTENIDO ORIGINAL EN SU TOTALIDAD, ASÍ COMO EL CUMPLIMIENTO DE LOS REQUERIMIENTOS QUE SE DICTAN EN LA GUÍA DE ELABORACIÓN DEL TRABAJO DE TITULACIÓN, POR LO QUE SE AUTORIZA A: GERARDO MARIO PINCAY CEDEÑO 0919802215**, QUE PROCEDA A SU PRESENTACION.**

Samborondón, 24 de Julio de 2019

TUTOR: Pedro Fabricio Echeverría Briones

Universidad Ecotec

DIVISIÓN DE POSGRADO
PROCESO DE TITULACIÓN
CERTIFICADO DEL PORCENTAJE DE COINCIDENCIAS

Habiendo sido nombrado PEDRO FABRICIO ECHEVERRIA BRIONES tutor del trabajo de titulación " APLICACIÓN DE TÉCNICAS DE INTELIGENCIA DE NEGOCIOS PARA SEGMENTACIÓN DE PAGOS ONLINE EN UNA ENTIDAD FINANCIERA" elaborado por GERARDO MARIO PINCAY CEDEÑO, 0919802215, con mi respectiva supervisión como requerimiento parcial para la obtención del título de **Maestría en Sistemas de Información Mención Inteligencia de Negocios.**

Se informa que el mismo ha resultado tener un porcentaje de coincidencias (9 %) mismo que se puede verificar en el siguiente link:
https://secure.urkund.com/view/53127271-751880-133087#q1bKLVayijY3iNVRKs5Mz8tMy0xOzEtOVbIy0DMwMDK1MLcwNrUAMgyMLE1MagE=

Adicional se adjunta print de pantalla de dicho resultado.

URKUND

Documento	GERARDO MARIO PINCAY CEDEÑO.docx (D54562014)
Presentado	2019-07-24 19:55 (-05:00)
Presentado por	gpincay@mgs.ecotec.edu.ec
Recibido	nvillena.ecotec@analysis.urkund.com
Mensaje	Tesis Documento final Mario Pincay Cedeño Mostrar el mensaje completo
	9% de estas 24 páginas, se componen de texto presente en 16 fuentes.

Pedro Fabricio Echeverria Briones
PROFESOR TUTOR DE LA CIENCIA

Resumen

Uno de los factores que más impactan de forma negativa a las entidades bancarias es cuando desean ofrecer nuevos productos el cual no tienen sus clientes segmentados, la correcta y oportuna segmentación de clientes mediante sus pagos On-Line traerá consigo crear estrategias para la captación de clientes y ofrecer productos innovadores.

En las últimas campañas de mercadeo se ha evidenciado que los resultados obtenidos no son los esperados ya que el perfil de los clientes que se les ofrece los productos no cumplen con sus necesidades, es por eso que se analizara información de los pagos Online para poder tener una perspectiva de la demanda de cada cliente.

Siguiendo la metodología CRISP-DM en base al análisis RFM (Recencia, Frecuencia, Monto de pagos Online), se procedió crear un Dashboard con el objetivo de poder analizar la información de sus pagos On-Line, y así crear estrategias de mercado para traer ventaja competitiva a la organización, esta solución que permitirá al área de Mercadeo tener una visión de sus clientes para la toma de decisiones.

Palabras clave: Segmentación de clientes, inteligencia de negocios, CRIPS-DM, Power BI, Dashboard.

Abstract

One of the factors that most negatively impact banking entities is when they want to offer new products which their customers do not have segmented, the correct and timely segmentation of customers through their On-Line payments will bring about creating strategies for customer acquisition and offer innovative products.

In the last marketing campaigns, it has been evidenced that the results obtained are not the expected ones since the profile of the clients that are offered the products do not meet their needs, that is why we analyze information of the Online payments to be able to have a perspective of the demand of each client.

Following the CRISP-DM methodology based on the RFM analysis (Receipt, Frequency, Online Payment Amount), we proceeded to create a Dashboard with the objective of being able to analyze the information of your On-Line payments, and thus create market strategies to bring competitive advantage to the organization, this solution that will allow the Marketing area to have a vision of its clients for decision making

Keywords: Customer segmentation, business intelligence, CRIPS-DM, Power BI, Dashboard.

Tabla de Contenidos

INTRODUCCIÓN 1

Antecedentes 3

Descripción del problema 4

Pregunta del problema. 4

Idea a Defender 4

Objetivos Generales y Específicos. 5

CAPÍTULO 1 6

1. Marco Teórico 6

1.1 Modalidad de Investigación y Metodología a Utilizar 6

1.1.1 Investigación Bibliográfica. 6

1.1.2 Investigación de Campo 6

1.2 Metodología CRISP-DM 7

1.2.1 ¿Por qué Utilizar la metodología CRISP-DM? 9

1.3 Bussiness Intelligence(BI) 10

1.4 Pagos Online 13

1.5 Segmentación 16

1.6 ETL 17

1.7 Dashboard 17

1.8 Data Mart. 18

1.9 Variables 19

1.10 RFM 20

1.11 Power BI 21

1.12 Sap PowerDesigner 22
1.13 SQL Server 22
CAPÍTULO 2 24
2. Solución de la Propuesta 24
2.1 Análisis de la situación actual. 24
2.2 Entendimiento de los objetivos. 25
2.3 Entendimiento de los Datos 26
2.4 Preparación de los datos 27
2.5 Requerimientos Funcionales y no Funcionales 29
2.5.1 REQ001 - Diagnosticar el estado actual del Proceso 29
2.5.2 REQ002 – Aplicación del modelo RFM 30
2.5.3 REQ003 – Desarrollo del Modelo de Datos. 30
2.5.4 REQ004 – Diseño/Desarrollo de ETL. 31
2.5.5 REQ005 – Desarrollo del Dashboard. 32
2.6 Roles y Responsabilidades. 33
2.7 Prototipos. 35
2.8 Componentes de la Arquitectura Tecnológica. 36
2.9 Diseño del modelo de Datos. 36
2.9.1 Seleccionar el proceso del negocio a modelar 37
2.9.2 Definir el nivel de granularidad del proceso del negocio 37
2.9.3 Escoger las dimensiones que aplican en cada fila de la tabla de hechos 37
2.9.4 Identificar los hechos que poblaran la tabla de hechos 38
2.10 Diccionario de Datos 42
2.11 Diseño de ETL 45

2.11.1 ETL Grabar DIM_ETABLECIMIENTOS 45
2.11.2 ETL Grabar DIM_HORAS 46
2.11.3 ETL Grabar DIM_FECHAS 47
2.11.4 ETL Grabar DIM_CLIENTES 48
2.11.5 Modelo RFM 51
2.12 Desarrollo del módulo de visualización. 53
2.12.1 Reporte de pagos por horas. 54
2.12.2 Reporte de pagos por Establecimientos. 54
2.12.3 Reporte de pagos por Día. 55
2.13 Presupuesto 57
2.14 Plan del Proyecto. 60
2.15 Riesgos y Contingencia. 60
CAPÍTULO 3 62
3. Validación de Resultados 62
3.1 Esquema de validación. 66
3.1.1 Análisis del reporte de pagos por hora. 66
3.1.2 Análisis del reporte de pagos por establecimientos. 67
3.2 Impacto. 68
CONCLUSIONES 69
RECOMENDACIONES 70
REFERENCIAS 71
ANEXOS 74

Lista de tablas

Tabla 2. 1 – Descripción de Matriz Origen destino 28
Tabla 2. 2 – Ficha Técnica. REQ-001 29
Tabla 2. 3 - Ficha Técnica. REQ-002. 30
Tabla 2. 4 - Ficha Técnica. REQ-003. 31
Tabla 2. 5 - Ficha Técnica. REQ-004. 32
Tabla 2. 6 - Ficha Técnica. REQ-005. 33
Tabla 2. 7 – Tabla de Roles. 34
Tabla 2. 8 – Tabla de Actividades y Roles. 34
Tabla 2. 9 - Descripción de Dimensiones y Hechos 38
Tabla 2. 11 – Detalle de tabla DIM_ESTABLECIMIENTOS. 42
Tabla 2. 12 - Detalle de tabla DIM_FECHAS. 42
Tabla 2. 13 - Detalle de tabla DIM_CLIENTES. 43
Tabla 2. 14 - Detalle de tabla DIM_HORA. 43
Tabla 2. 15 - Detalle de tabla HEC_PAGOS_ONLINE. 44
Tabla 2. 16 – Tabla de Frecuencias. 51
Tabla 2. 17 – Tabla Recencia. 52
Tabla 2. 18 – Tabla Monetario. 52
Tabla 2. 19 - Recursos Humanos. 57
Tabla 2. 20 – Servicios. 57
Tabla 2. 21 – Recurso de hardware. 58
Tabla 2. 22 – Recurso de Software. 58
Tabla 2. 23 – Datos. 59
Tabla 2. 24 – Presupuesto Total. 59
Tabla 2. 25 – Riesgos y Contingencia 61

Tabla 3. 1 – Formulario de Encuesta. 62

Lista de figuras

Figura 1. 1 - Secuencia del Proceso metodología CRISP-DM. 7
Figura 1. 2 - Grado de utilización de distintas metodologías. 9
Figura 1. 3 - Fases de BI según (Peña, 2006). 12
Figura 1. 4 – Total de transacciones por año 14
Figura 1. 5 – Total recaudado por año 15
Figura 1. 6 - Grafico se visualiza muchas fuentes diferentes y cargarlos en varios destinos. 17
Figura 1. 7 - Grafico modelo RFM. 21

Figura 2. 1 - Prototipo del Dashboard. 35
Figura 2. 2 - Arquitectura BI. 36
Figura 2. 3 - Modelo Dimensional Copo de Nieve 39
Figura 2. 4 - Grafico de generación de script DDL SQL 40
Figura 2. 5 – Data Definition Language(DDL) - Script SQL 41
Figura 2. 6 - Ejecución del ETL Establecimientos 45
Figura 2. 7 - Tabla DIM_ESTABLECIMIENTOS. 46
Figura 2. 8 - Ejecución del ETL Horas 46
Figura 2. 9 - Tabla DIM_HORA. 47
Figura 2. 10 - Ejecución del ETL Fechas. 47
Figura 2. 11 - Tabla DIM_FECHA. 48
Figura 2. 12 - Ejecución del ETL Clientes. 48
Figura 2. 13 - Tabla DIM_CLIENTES. 49
Figura 2. 14 - Ejecución de Procesos ETL. 49
Figura 2. 15 - Tabla HEC_PAGOS_ONLINE. 50
Figura 2. 16 - Tabla MODELO_RFM. 50
Figura 2. 17 - Dashboard. 53
Figura 2. 18 - Grafico de pagos por Horas. 54
Figura 2. 19 - Grafico de pagos por Establecimientos. 55
Figura 2. 20 - Grafico de pagos por Día. 56
Figura 2. 21 – Planificación de cronograma de actividades 60

Figura 3. 1 – Grafico de porcentajes de encuestados. 63
Figura 3. 2 – Grafico de medios de pagos web. 64
Figura 3. 3 – Grafico de frecuencia de pagos web 64
Figura 3. 4 – Grafico de tipo de pagos realizados en plataformas virtuales 65
Figura 3. 5 – Grafico tendencia de edades de pagos Web 66
Figura 3. 6 – Grafico del reporte de pagos por hora con mayor recaudación 67
Figura 3. 7 – Grafico con mayor recaudación por establecimientos. 68

Introducción

Hoy en día hablar de internet es muy común en la sociedad, la interconexión con la mayoría de empresas ha sido consigo facilitar al usuario común poder realizar transacciones vía web el cual facilitan la vida de las personas, los pagos On-Line es una forma electrónica de operación comercial el cual ambas personas tanto el comprador como el vendedor son beneficiados, este tipo de transacción aporta facilidad, seguridad e integridad.

Según investigaciones se ha detectado que casi el 90% de transacciones son realizadas por muchos medios electrónicos como dispositivos móviles, Laptop, Smart TV, etc.

Dentro de las entidades bancarias los clientes pasan ser uno de los activos más importantes en la organización que son ellos los que garantizan la rentabilidad para mantenerse en el mercado y poder crecer de forma sustentable, uno de los factores que más impactan de forma negativa a las entidades financieras es cuando desean ofrecer nuevos productos el cual no tienen sus clientes segmentados, los departamentos responsables no tienen toda la información consolidada; a su vez al momento de la captación de clientes no sabemos el perfil del cliente, sus intereses, afinidades, entre otros.

Entre las técnicas más utilizadas para ofrecer productos nuevos es la segmentación de clientes por medio de perfiles. Estas iniciativas se basan fundamentalmente en análisis del mercado que no responden en muchos de los casos a la realizada que

hoy requieren los clientes en función de sus preferencias al no tomar en cuenta factores críticos como:

- ✓ Intereses Financieros
- ✓ Beneficios que las entidades bancarias puedan ofrecerles.
- ✓ Incremento de sus Inversiones a corto o largo plazo
- ✓ Productos dirigidos y personalizados

Es fundamental que las entidades bancarias implementen herramientas innovadoras que permitan atraer más clientes por medio de técnicas de análisis de datos para ofrecer productos acordes a sus perfiles. En la presente investigación se propone una herramienta de inteligencia de negocio que a partir de sus pagos On-Line permita crear segmentos de clientes por medio de análisis multivariantes para el análisis de datos aplicando Inteligencia de Negocios y a su vez por medio de un Dashboard verificar perfiles de usuarios, de esta forma tomar acciones proactivas basadas en los aspectos descritos anteriormente.

Con ello implementar estrategias de mercadeo donde se ofrezcan productos o servicios acorde a las necesidades de cada grupo de clientes y traer ventaja competitiva a la organización, al final lo que buscamos con la segmentación es conocer el tipo de clientes que tenemos, diferéncialos, clasificarlos y que estrategia podemos brindarles con ello reducir gastos innecesarios y los resultados serían más eficientes el cual se verá reflejados en los ingresos que aumentaran la rentabilidad de la entidad financiera.

Antecedentes

Dentro de la problemática que las entidades financieras hoy en día mantienen al no poder ofrecer productos o servicios nuevos es no constar con un área de Inteligencia de Negocio el cual analice y evalúe la información transaccional que genera diariamente.

Todos los procesos son llevados en departamentos independientes o en ocasiones de forma manual el cual al momento de solicitar alguna base de cliente segmentados a los departamentos de sistemas, mercadeo o marketing estos se vuelven engorrosos por no tener toda la información consolidada.

El limitado análisis de los factores descritos anteriormente trae consigo que las estrategias de mercadeo para la captación de los clientes no tengan la efectividad deseada, trayendo consigo gastos Financieros, Administrativos y del Recurso Humano de la entidad financiera.

Descripción del problema

Dentro de la entidad financiera al momento de realizar una campaña de mercadeo, estos no tienen la información consolidada de una manera eficiente el cual al solicitar la información de los clientes al departamento de sistema, ellos proceden a facilitar la información de una manera general, los departamento de mercadeo realizan el respectivo análisis para poder enfocar la campaña de captación con la información proporcionada, con esta propuesta de análisis de datos por medio de la información proporcionada por la entidad financiera ellos pueden tomar una mejor decisión ya que se diseñara un Dashboard con reportes acordes a sus necesidades y así ofertar una campaña eficiente y dar un enfoque más personalizado al usuario, con esto ahorraríamos recursos administrativos y financieros para el beneficio de la entidad financiera.

Pregunta del problema.

¿Cómo mejorar la segmentación de clientes en una entidad financiera mediante los pagos online?

Idea a Defender

Si se implementan técnicas de inteligencia de negocio con análisis multivariantes en la creación de un Dashboard se mejorará la segmentación de clientes a partir de sus pagos online

Objetivos Generales y Específicos.

- Diagnosticar el estado actual del proceso de segmentación de cliente de instituciones bancarias.
- Obtener un modelo predictivo de los pagos On-Line utilizando técnicas de Machine Learning para ver el comportamiento de sus clientes.
- Aplicar el modelo RFM (Reciente, Frecuencia, Monetario) en el análisis de los datos generados por los pagos online para la segmentación de clientes.
- Desarrollar un ETL para la estructuración de los datos que serán visualizados en una Dashboard.
- Utilizar herramientas de Inteligencia de Negocios creando un Dashboard el cual muestre reportes de segmentos de clientes en base a sus pagos On-Line que le permitan analizar su información.
- Validar los resultados de la segmentación de clientes aplicando métodos cualitativos y cuantitativos.

Capítulo 1

1. Marco Teórico

Dentro del marco teórico que enunciamos en el presente trabajo de investigación estará enmarcado en algunos conceptos y la metodología que se procederá implementar para poder desarrollar el proyecto.

1.1 Modalidad de Investigación y Metodología a Utilizar

1.1.1 Investigación Bibliográfica.

Un medio importante para el desarrollo y respaldo de este trabajo es la información bibliográfica el cual nos permite iniciar con la investigación científica y recolección de datos y documentos tales como libros, revistas, tesis de grado el cual tengan información relacionada con la segmentación de clientes, siempre y cuando todo este material guarde relación con nuestro problema de estudio.

1.1.2 Investigación de Campo

La modalidad a utilizarse en el presente trabajo es la investigación de campo, la misma que se tendrá la oportunidad de vivir más de cerca lo que sucede en el departamento de Mercadeo de la entidad financiera, dentro de la investigación utilizaremos el recurso humano del área de Mercadeo el cual nos dará a conocer como llevan actualmente el proceso de las campañas de captación de clientes, esta información será recolectada el cual sustentará de mejor manera nuestra investigación, para la recolección de datos aplicaremos técnicas de investigación que serán la encuesta como la entrevista.

1.2 Metodología CRISP-DM

CRISP-DM que son las siglas de Cross-Industry Standard Process for Data Mining, es un método probado para orientar sus trabajos de minería de datos.

- Como **metodología**, incluye descripciones de las fases normales de un proyecto, las tareas necesarias en cada fase y una explicación de las relaciones entre las tareas.
- Como **modelo de proceso,** CRISP-DM ofrece un resumen del ciclo vital de minería de datos

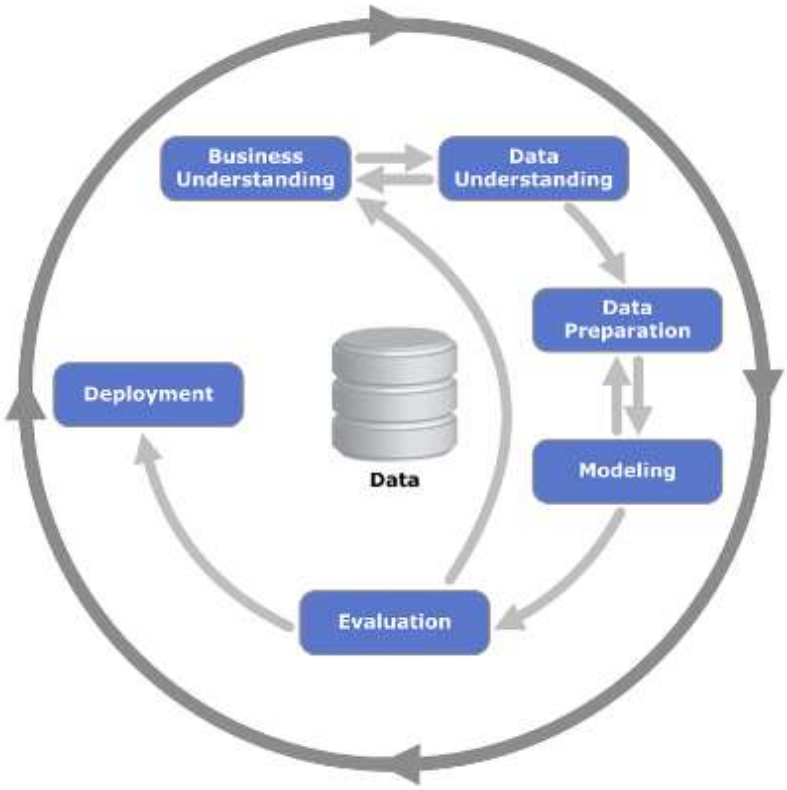


Figura 1. 1 - Secuencia del Proceso metodología CRISP-DM.
(Fuente: (CRISP-DM, 2016))

El ciclo vital del modelo contiene seis fases con flechas que indican las dependencias más importantes y frecuentes entre fases. La secuencia de las fases no es estricta. De hecho, la mayoría de los proyectos avanzan y retroceden entre fases si es necesario. El modelo de CRISP-DM es flexible y se pueden personalizar fácilmente.

CRISP-DM permite crear un modelo de minería de datos que se adapte a sus necesidades concretas.

En tal situación, las fases de modelado, evaluación e implementación pueden ser menos relevantes que las fases de preparación y comprensión de datos. Sin embargo, es muy importante considerar algunas cuestiones que surgen durante fases posteriores para la planificación a largo plazo y objetivos futuros de minería de datos. (IBM, 1994, 2012)

En el caso de los proyectos de implementación de minería de datos una de las metodologías que ha tenido más apoyo de las empresas privadas y organismos públicos es CRISP-DM, como se puede observar en la siguiente gráfica, publicada en kdnuggets.com, y que representa el grado de utilización de las principales guías de desarrollo de proyectos de minería de datos según las encuestas realizadas.

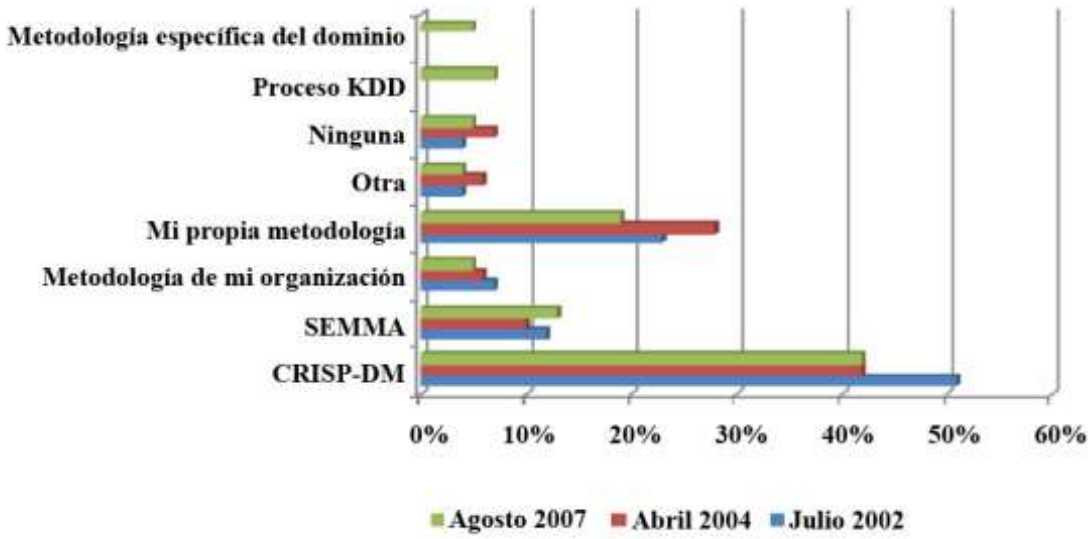


Figura 1. 2 - Grado de utilización de distintas metodologías.
(Fuente: (KDNUGGETS, 2014))

1.2.1 ¿Por qué Utilizar la metodología CRISP-DM?

Utilizamos la metodología CRISP-DM, porque es un modelo estructurado el cual se ha usado bastante para el desarrollo de proyectos de minería de datos, a su vez también se adapta a las herramientas que utilizaremos. En este caso, Power Design y Power BI. Como se ha visto en el punto 2.1.3, los pasos a seguir en el proceso de la minería de datos son:

1. Comprensión del negocio.
2. Comprensión de los datos.
3. Preparación de los datos.
4. Modelado.
5. Evaluación.
6. Implantación.

Power Design y Power BI da soporte a los pasos 4, 5 y 6 del proceso el cual se procederá con el diseño y modelado desde una manera fácil.

Además, Power BI es una herramienta que proporciona al usuario varios algoritmos para diferentes tipos de análisis como son: Clustering, árboles de decisión, clasificador bayesiano, SVM (Máquinas de Vectores de Soporte) y GLM (Modelo Lineal Generalizado), etc.

Otra razón por lo cual se optó por esta herramienta en su fácil uso, porque Power BI tiene una interfaz gráfica muy intuitiva e incluso puede utilizarla usuarios no expertos.

1.3 Bussiness Intelligence(BI)

Es necesario conocer que es Business Intelligence (BI) o Inteligencia de Negocios Implementado por Howard Dresner en 1989, describe una serie de conceptos y métodos enfocados a mejorar la toma de decisiones en las organizaciones (Biere, 2003).

El Business Intelligence surge que a partir de dicha información que puedan generar escenarios, pronósticos e informes, que apoyen a la toma de decisiones, lo que esto conllevaría tomar ventaja competitiva en la organización.

La clave para el BI es la información y uno de sus mayores beneficios es la posibilidad de utilizarla en la toma de decisiones y a su vez es el principal conocimiento que sostiene a la organización.

Otra definición, acuñada por (Peña, 2006), el BI tiene como objetivo:

> *"procura caracterizar una amplia variedad de tecnologías, plataformas de software, especificaciones de aplicaciones y procesos."*

Dentro de todos los conceptos que se han definido podemos decir que el objetivo de Business Intelligence, es poder contribuir a la organización para tomar decisiones de manera rápida, objetiva y mejorar su desempeño para obtener ventaja competitiva con otras organizaciones en el mercado.

> *"Business Intelligence es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios."*

La inteligencia de negocios tiene un ciclo donde las organizaciones establecen los objetivos, analizan sus progresos, reflexionan, actúan, miden su éxito y vuelven a empezar una nueva fase. Según (Peña, 2006) las fases son:

- El análisis de los datos a recopilar.
- Reflexión, que supone el estudio minucioso de los hechos y de la situación.
- La conexión de la acción al ciclo del BI es a través del proceso de toma de decisiones, en donde las acciones se suceden como resultado de las decisiones.
- La medición evalúa los resultados al compararlos con otros estándares cuantitativos y las expectativas que se plantean originalmente.

Figura 1.1: muestra las fases de BI según Peña.

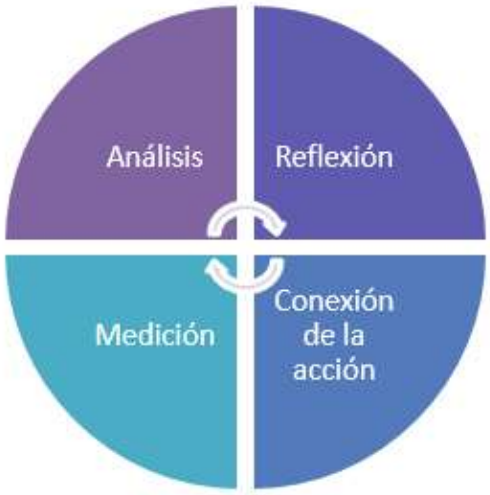


Figura 1. 3 - Fases de BI según (Peña, 2006).

(Fuente: Elaboración propia)

El BI tiene varios beneficios que nos permiten:

- Manejar el crecimiento de la organización, esto nos permite evolucionar, crecer, y poder crear estrategias competitivas para la toma de decisiones.
- Muchas empresas consideran una solución de BI con el objetivo de tener la capacidad de medir gastos, por medio de indicadores de gestión.
- Al tener almacenada mucha información de los clientes podemos transformarla en conocimiento con el objetivo de entender mejor el perfil de cada Cliente.

Por el contrario, también posee varios obstáculos en todas las organizaciones, como, por ejemplo:

- Pocas organizaciones están haciendo lo necesario para darle forma a su información.

- Muchas organizaciones aún no cuentan con la fortaleza tecnológica y analítica propicia.
- Muchas organizaciones no aprovechan al máximo el BI para tomar decisiones.

1.4 Pagos Online

Un sistema de pago electrónico es un sistema de pago que facilita la aceptación de pagos para las transacciones en línea a través de Internet.

Los sistemas de pagos electrónicos(EPS), realizan la transferencia del dinero entre compradores y vendedores en una acción de compra-venta electrónica a través de una entidad financiera autorizada por ambos. Es, por ello, una pieza fundamental en el proceso de compra-venta dentro del comercio electrónico.

El comercio electrónico por Internet se ofrece como un nuevo canal de distribución sencillo, económico y con alcance mundial las 24 horas del día todos los días del año, y esto sin los gastos y limitaciones de una tienda clásica: personal, local, horario, infraestructura, etc. (Wikipedia, 2019)

Otra definición general lo detalla el banco central de Ecuador como:

> *"Mecanismo que permite a las entidades que mantienen una cuenta en el BCE, la ejecución de órdenes de pago en línea y en tiempo real, mediante la transferencia electrónica de fondos, afectando las cuentas corrientes que mantienen". (BCE Ecuador, 2008)*

Dentro del análisis que se realizó de los pagos On-Line se evidencia en los últimos 3 años que la tendencia ha crecido considerablemente, según la Superintendencia de Bancos, el 73% de las transacciones que se realizaron en instituciones financieras el 2017 fueron por canales electrónicos



Total de Transacciones de los últimos 3 años		
2017	2018	2019 – fecha corte
61867664	68746941	35170538

Figura 1. 4 – Total de transacciones por año

(Fuente: Elaboración propia)

Mientras que el Banco Central informa que la tasa de crecimiento del uso de los medios de pago digitales entre 2016 y 2017 aumentó 30%.

Se evidencia que el total de recaudaciones se realicen anualmente en las instituciones financieras del Ecuador, la tendencia va en crecimiento ascendente.

En las siguiente grafico se muestra el total recaudado anualmente de todos los bancos del Ecuador hasta Julio/2017.



Recaudaciones de los últimos 3 años		
2017	2018	2019 – fecha corte
$ 100.050.660,09	$ 108.898.875,00	$ 56.040.870,25

Figura 1. 5 – Total recaudado por año

(Fuente: Elaboración propia)

1.5 Segmentación

La segmentación es una herramienta que busca agrupar datos en diferentes conglomerados, en donde cada uno de los datos pertenecientes a un grupo son similares entre sí y diferentes a los demás clúster. Se puede definir cómo:

> *"El proceso de identificación y caracterización, mediante técnicas estadísticas, de subgrupos de consumidores dentro de un mismo mercado que presentan distintas necesidades, así como la selección de aquellos que podrían ser objeto de una oferta comercial diferente adaptada a los mismos, con el fin de incrementar la eficacia y rentabilidad de la actuación comercial de la empresa". (Agueda, 2006)*

Esta herramienta nos permite analizar el mercado de una manera menos agregada, bajo el supuesto de que los clientes no son iguales, por lo que tratarlos a todos de la misma manera es un error. Es por esto qué en marketing se ha utilizado la segmentación de mercado para detectar oportunidades de negocio, segmentos más rentables y desarrollar estrategias diferenciadas para cada uno de ellos. Una buena segmentación debe tener grupos de tamaño significativo, accesibles e identificables. Lo que implica que los grupos sean lo suficientemente grandes y fácilmente ubicables para poder apoyar una estrategia en cada uno de estos y que sea rentable.

Existen diferentes tipos de segmentación, las cuales dependerán de las diferentes variables a utilizar, que serán elegidas a partir de los objetivos que busca la organización.

1.6 ETL

Extract, Transform and Load («extraer, transformar y cargar», frecuentemente abreviado ETL) es el proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, data mart, o data warehouse para analizar, o en otro sistema operacional para apoyar un proceso de negocio.

Los procesos ETL también se pueden utilizar para la integración con sistemas heredados. (Wikipedia, ETL, 2012)



Figura 1. 6 - Grafico se visualiza muchas fuentes diferentes y cargarlos en varios destinos.

(Fuente: (SAS, 2012))

1.7 Dashboard

Según Stephen (Few, 2006), un Dashboard es una exposición visual de la información más importante requerida para lograr uno o más objetivos; consolidado y ordenado en una única pantalla de tal manera que la información pueda ser monitoreada con una mirada y poder evaluar la situación de la empresa.

Un Dashboard efectivo permite mostrar visualmente a los usuarios información relevante y tareas comunes, costos de administración y cualquier otra información crítica que afecta los objetivos del negocio

Así mismo presenta información altamente resumida con el objetivo tener una visión general del estado de la organización de una manera macro.

1.8 Data Mart.

Un data mart es una versión especial de almacén de datos (data warehouse). R. Los datos existentes en este contexto pueden ser agrupados, explorados y propagados de múltiples formas para que diversos grupos de usuarios realicen la explotación de los mismos de la forma más conveniente según sus necesidades. (Wikipedia, Datamark, 2016)

El Datamart es un sistema orientado a la consulta, en el que se producen procesos batch de carga de datos (altas) con una frecuencia baja y conocida. Es consultado mediante herramientas **OLAP** (On line Analytical Processing - Procesamiento Analítico en Línea) que ofrecen una visión multidimensional de la información. Sobre estas bases de datos se pueden construir **EIS** (Executive Information Systems, Sistemas de Información para Directivos) y **DSS** (Decision Support Systems, Sistemas de Ayuda a la toma de Decisiones). (Wikipedia, Datamark, 2016)

En síntesis, podemos definir que los datamart son pequeños data warehouse que estarán centradas en un área de negocio de la organización con el objetivo de poder explorar la información de diferentes maneras.

Dentro de las ventajas de aplicar un Datamart a un negocio, se han seleccionado las siguientes:

- Son simples de implementar.
- Conllevan poco tiempo de construcción y puesta en marcha.
- Permiten manejar información confidencial.
- Reflejan rápidamente sus beneficios y cualidades.
- Reducen la demanda del depósito de datos

1.9 Variables

Independiente

Segmentación de clientes mediante los pagos On-Line Estrategia de Ventas Dependiente

Ofrecer productos acordes al perfil de cada grupo de clientes.

Tener una visión que podemos ofertar acorde a la demanda

1.10 RFM

Es una segmentación y también un predictor muy usado en la industria de compras y ventas, que básicamente examina lo que el cliente ha comprado utilizando tres factores; (R) Recientes pagos, (F) Frecuencia de pagos y (M) Monto de los pagos en términos monetarios, debido a que:

- Si un cliente no ha realizado un pago por largo tiempo, la probabilidad que espontáneamente realice un nuevo pago es muy baja.
- Si un cliente ha realizado pagos muchas veces, es altamente probable que se trate de un cliente regular, y que espontáneamente continúe pagando con la misma frecuencia.
- Si un cliente realiza pagos de montos altos, es probable que este cliente esté dispuesto a realizar compromisos con la marca (especialmente si tiene pagos repetitivos de alto monto).

En general, es un método muy efectivo para resumir grandes volúmenes de datos de pagos en una segmentación simple y manejable.

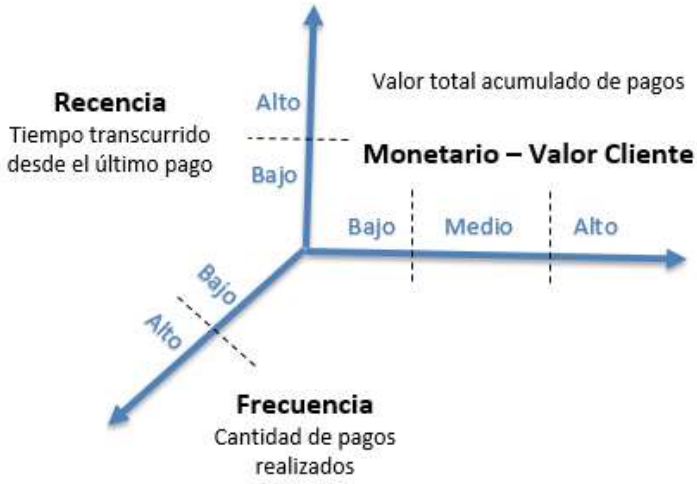


Figura 1. 7 - Grafico modelo RFM.

(Fuente: Elaboración propia)

1.11 Power BI

Power BI es una colección de servicios de software, aplicaciones y conectores que trabajan juntos para convertir sus fuentes de datos no relacionadas en información coherente, visualmente inmersiva e interactiva. Sus datos pueden ser una hoja de cálculo de Excel o una colección de almacenes de datos híbridos basados en la nube y en las instalaciones. Power BI le permite conectarse fácilmente a sus fuentes de datos, visualizar y descubrir lo que es importante, y compartirlo con cualquier persona o persona que desee.

Power BI puede ser simple y rápido, capaz de crear perspectivas rápidas desde una hoja de cálculo de Excel o una base de datos local. Pero Power BI también es robusto y de nivel empresarial, listo para el modelado extenso y el análisis en tiempo real, así como para el desarrollo personalizado. Puede ser su herramienta de visualización e informe

personal, y también servir como motor de análisis y decisión para proyectos de grupo, divisiones o corporaciones completas.

1.12 Sap PowerDesigner

Es una herramienta de modelado empresarial colaborativo que tiene soporte para diferentes tipos de modelos. El más importarte para este proyecto es el Modelo de datos el cual funciona con la mayoría de sistemas de base de datos relacionales y utilizar para crear el modelo conceptual y físico de la base de datos. PowerDesigner permite visualizar, analizar y manipular metadatos para tener una mejor arquitectura empresarial de información brindando un enfoque basado en modelos lo cual permite a la tecnología alinear con el giro del negocio logrando una mejor implementación de arquitecturas de información. Está diseñado con buenas prácticas de análisis, diseño y gestión de metadatos brindando una mejor productividad individual y grupal. (Wang, 2016).

1.13 SQL Server

Microsoft SQL Server es un sistema de gestión de base de datos relacional, desarrollado por la empresa Microsoft.

El lenguaje de desarrollo utilizado (por línea de comandos o mediante la interfaz gráfica de Management Studio) es Transact-SQL (TSQL), una implementación del estándar ANSI

del lenguaje SQL, utilizado para manipular y recuperar datos (DML), crear tablas y definir relaciones entre ellas (DDL).

Dentro de los competidores más destacados de SQL Server están: Oracle, MariaDB, MySQL, PostgreSQL. SQL Server ha estado tradicionalmente disponible solo para sistemas operativos Windows de Microsoft, pero desde 2016 está disponible para GNU/Linux,23 y a partir de 2017 para Docker también.4

Puede ser configurado para utilizar varias instancias en el mismo servidor físico, la primera instalación lleva generalmente el nombre del servidor, y las siguientes - nombres específicos (con un guion invertido entre el nombre del servidor y el nombre de la instalación). (Hotek, 2009)

Sus características principales son:

- Soporte de transacciones.
- Soporta procedimientos almacenados.
- Incluye también un entorno gráfico de administración, que permite el uso de comandos DDL y DML gráficamente.
- Permite trabajar en modo cliente-servidor, donde la información y datos se alojan en el servidor y los terminales o clientes de la red sólo acceden a la información.
- Además permite administrar información de otros servidores de datos.

Capítulo 2

2. Solución de la Propuesta

2.1 Análisis de la situación actual.

La Institución financiera tiene presencia en el Sector bancario desde hace 47 años, es una de las empresas que ofrece a sus clientes productos innovadores de **crédito** a nivel Nacional dentro del Ecuador.

Al momento requiere desarrollar técnicas de inteligencia de negocios para segmentar clientes a partir de sus pagos Online a nivel Nacional que dentro del análisis que realizan para ofertar sus productos innovadores, su campaña de mercadeo no obtiene los resultados esperados al no tener un enfoque del perfil de cada cliente para ofertar el producto que le interesa.

En las últimas campañas de mercadeo se ha evidenciado que los resultados obtenidos no son los esperados ya que el perfil de los clientes que se les ofrece los productos no cumplen con sus necesidades, es por eso que se analizara información de los pagos Online del último mes para poder tener una perspectiva de la demanda de cada cliente.

Por lo que ha planteado objetivos generales para que mediante la inteligencia de negocios cumplir un factor fundamente en poder brindar a sus clientes productos innovadores de crédito acorde a su perfil y poder tomar decisiones adecuadas:

1. Diagnosticar el estado actual del proceso de segmentación de cliente de instituciones bancarias.

2. Aplicar el modelo RFM (Reciente, Frecuencia, Monetario) en el análisis de los datos generados por los pagos online para la segmentación de clientes.
3. Desarrollar un ETL para la estructuración de los datos que serán visualizados en una Dashboard.
4. Utilizar herramientas de Inteligencia de Negocios creando un Dashboard el cual muestre reportes de segmentos de clientes en base a sus pagos Online que le permitan analizar su información.

Este análisis se llevará a cabo en la ciudad de Guayaquil.

2.2 Entendimiento de los objetivos.

Dentro del entendimientos de los objetivos propuestos es utilizar herramientas de Inteligencia de Negocios creando un Dashboard el cual muestre reportes de segmentos de clientes en base a sus pagos Online que le permitan analizar su información.

El propósito del Dashboard es contar con los siguientes reportes para tomar decisiones sobre:

a. Reporte de pagos mensual por hora.
b. Reporte de pagos mensual por Establecimientos.
c. Reporte de pagos mensual por Día.
d. Reporte de pagos por Recencia.
e. Reporte de pagos por Frecuencia.

f. Pagos mensuales por Establecimientos.

g. Clientes Segmentados con menor recencia de pagos.

La información que se analiza es proporcionada por la entidad financiera el cual tienen los siguientes campos definidos:

- CodigoDePersona Number(8)
- FechaDeMovimiento Datetime
- HoraMinutoSegundo Time
- ValorDelMovimiento number(18,2)
- Establecimiento String(500)
- DescripcionAdicional String(500)
- Sucursal. String(2)

Esta información es proporcionada mediante un archivo *.csv el cual posee aproximadamente unos 230000 registro por mes, se evaluará el mes de Enero/2019.

Con el objetivo de que el proceso de toma de decisiones se convierta en una práctica habitual en la entidad financiera se diseñara una Arquitectura de Inteligencia de Negocios que considere 3 (tres) perspectivas fundamentales.

a. Arquitectura de Procesos (Pagos On-Line)

b. Arquitectura de Infraestructura (Hardware)

c. Arquitectura de Software

2.3 Entendimiento de los Datos

Para construir el Dashboard se tendrá que realizar el entendimiento de los siguientes datos:

- CLIENTES

    - CodigoDeCliente
    - NombreDeCliente
    - Direccion
    - Email
- ESTABLECIMIENTOS
    - Descripción
- FECHAS
    - Fecha
    - Anio
    - Mes
    - Semana
    - Dia
- HORA
    - HoraDelDia

2.4 Preparación de los datos

Dentro de la preparación de los datos procedemos a crear la matriz Origen – Destino para tener claro que proceso se realizara en la creación del Dashboard y la segmentación de clientes con los pagos On-Line.

Esto implica seleccionar el subconjunto de datos que se va a utilizar, limpiarlos para mejorar su calidad, añadir nuevos datos a partir de los existentes y darles el formato requerido por la herramienta de modelado.

Dentro de este punto se procederá a la creación de una modelo copo de nieve para poder transformar la información y enviarla por medio de los ETL a su destino que será el Datamark que esta implementado en una base de datos SQL Server 2017.

En la tabla 2.1 describe la Matriz Origen – Destino

ORIGEN		TRANSFORMACION	DESTINO	
ARCHIVO DE TEXTO PAGOS ONLINE	CCONCEPTO NOMBRECONCEPTO ESTABLECIMIENTO CPERSONA FECHAMOV HORAMINUTOSEGUNDO VALORMONEDACUENTA DESCRIPCIONADICIONAL CSUCURSAL_ORIGEN	Transformacion: se procedera insertar todos los registro de los establecimientos del archivo de texto	CodigoEstablecimiento_pk, CodigoEstablecimiento, NombreEstablecimiento	DIM_ESTABLECIMIENTOS
		Transformacion: se procede separa los datos de la fecha, Anio, Mes, Semana, Dia	Fecha_pk, Fecha, Anio, Mes, Semana, Dia	DIM_FECHAS
ARCHIVO DE TEXTO CLIENTES	CODIGOCLIENTE NOMBRE CLIENTE APELLIDO CLIENTE DIRECCION TELEFONO	Transformacion: se une el nombre_cliente + ' ' + el apellido_cliente	CodigoCliente_pk, CodigoCliente, NombreCliente, Direccion, Telefono	DIM_CLIENTES
DIM_CLIENTES DIM_FECHAS DIM_ESTABLECIMIENTOS DIM_HORA		se une por CodigoEstablecimientos_pk se une por CodigoCliente_pk se une por Fecha_pk se une por hora_pk	CodigoEstablecimiento_pk, CodigoCliente_pk, Fecha_pk, Hora_pkIdRegistro, CantidadPagos, TotalPagos, FechaCarga	HEC_PAGOS_ONLINE

Tabla 2. 1 – Descripción de Matriz Origen destino

(Fuente: Elaboración propia)

2.5 Requerimientos Funcionales y no Funcionales.

Dentro de los requerimientos funcionales y no funcionales se describe el proceso general de cómo será el desarrollo de este proyecto, a continuación, se describen cada uno de los requerimientos.

2.5.1 REQ001 - Diagnosticar el estado actual del Proceso

Dentro de esta opción se diagnostica el estado actual del proceso de segmentación de cómo lo está llevando la entidad bancaria, para así poder tener una visión de cómo se deberá realizar el desarrollo del tema propuesto, aquí se procederá con la entrevista al personal de mercadeo con el objetivo de analizar la problemática actual.

Numero: REQ-001	**Requisito: Diagnosticar el estado Actual del Proceso**
Versión	**1.0**
Dependencias	- Entrevista con el área de Mercadeo. - Que la información proporcionada por el Departamento de Mercado sea suficiente para poder levantar la información para el desarrollo del proyecto.
Programador:	Mario Pincay Cedeño
Descripción	Dentro de esta opción se diagnostica el estado actual del proceso de segmentación de cómo lo está llevando la entidad bancaria, para así poder tener una visión como se deberá realizar el desarrollo de tema propuesto aquí se procederá con la entrevista
Importancia	Alta
Estado	Finalizado.

Tabla 2. 2 – Ficha Técnica. REQ-001.

(Fuente: Elaboración propia)

2.5.2 REQ002 – Aplicación del modelo RFM

Dentro de esta opción se deberá solicitar el archivo que es proporcionado por la entidad financiera el cual será procesado para poder crear el modelo RFM con la información generada

Numero: REQ-002	**Requisito: Aplicación del modelo RFM.**
Versión	***1.0***
Dependencias	- *Datos proporcionados por la entidad financiera sean correctos.* - *poder levantar la información para el desarrollo del proyecto.*
Programador:	*Mario Pincay Cedeño*
Descripción	*Dentro de esta opción se deberá solicitar el archivo que es proporcionado por la entidad financiera el cual será procesado para poder crear el modelo RFM con la información generada*
Importancia	*Alta*
Estado	*Finalizado.*

Tabla 2. 3 - Ficha Técnica. REQ-002.

(Fuente: Elaboración propia)

2.5.3 REQ003 – Desarrollo del Modelo de Datos.

Dentro de esta opción se procederá con la creación del modelado de datos desde la herramienta, aquí se diseñará las dimensiones para el Datamart, las cuales se describen a continuación.

- DIM_ESTABLECIMIENOS
- DIM_FECHA
- DIM_HORA
- DIM_CLIENTES
- HEC_PAGOS_ONLINE.

Numero: REQ-003	Requisito: Desarrollo del Modelo de Datos.
Versión	***1.0***
Dependencias	- *Herramienta de modelados Power Design.* - *Herramienta de Bussiness Intelligen Visual Studio.* - *Integration services de Sql Server.*
Programador:	*Mario Pincay Cedeño*
Descripción	*Dentro de esta opción se procederá con la creación del modelado de datos desde la herramienta, se diseñará las dimensiones para el Datamart, las cuales se describen a continuación.* - *DIM_ESTABLECIMIENTOS* - *DIM_FECHA* - *DIM_HORA* - *DIM_CLIENTES* - *HEC_PAGOS_ONLINE.*
Importancia	*Alta*
Estado	*Finalizado.*

Tabla 2. 4 - Ficha Técnica. REQ-003.

(Fuente: Elaboración propia)

2.5.4 REQ004 – Diseño/Desarrollo de ETL.

Dentro de esta opción se procederá con el desarrollo de los ETL el cual insertará por medio la ejecución la información proporcionada por la entidad bancaria el cual llenará los datos de las siguientes dimensiones: DIM_ESTABLECIMIENTOS, DIM_FECHA, DIM_HORA, DIM_CLIENTES, HEC_PAGOS_ONLINE.

Numero: REQ-004	Requisito: Diseño/Desarrollo de ETL.
Versión	***1.0***
Dependencias	- *Modelo Conceptual desarrollado.* - *Herramienta de Bussiness Intelligen Visual Studio.* - *Integration services de Sql Server.*
Programador:	*Mario Pincay Cedeño*
Descripción	*Dentro de esta opción se procederá con el desarrollo de los ETL el cual insertará por medio la ejecución la información proporcionada por la entidad bancaria el cual llenará los datos de las siguientes dimensiones: DIM_ESTABLECIMIENTOS, DIM_FECHA, DIM_HORA, DIM_CLIENTES, HEC_PAGOS_ONLINE, MODELOS_RFM.*
Importancia	*Alta*
Estado	*Finalizado.*

Tabla 2. 5 - Ficha Técnica. REQ-004.

(Fuente: Elaboración propia)

2.5.5 REQ005 – Desarrollo del Dashboard.

Dentro de esta opción se procederá con el desarrollo del Dashboard el cual estarán descritos los siguientes reportes.

- Reporte de Pagos por hora.
- Reporte de Pagos por Establecimientos.
- Reporte de Pagos por Día.
- Reporte de Pagos Mensuales por Establecimientos.
- Reporte de Clientes segmentados con menor recencia de pagos.
- Reporte de Clientes segmentados con mayor frecuencia de pagos.

Numero: REQ-005	Requisito: Desarrollo del Dashboard.
Versión	***1.0***
Dependencias	- *Modelo Conceptual desarrollado.* - *Herramienta Power BI.* - *SQL Server.*
Programador:	*Mario Pincay Cedeño*
Descripción	*Dentro de esta opción se procederá con el desarrollo del Dashboard el cual estarán descritos los siguientes Reportes.* - *Reporte de Pagos por hora.* - *Reporte de Pagos por Establecimientos.* - *Reporte de Pagos por Día.* - *Reporte de Pagos Mensuales por Establecimientos.* - *Reporte de Clientes segmentados con menor recencia de pagos.* - *Reporte de Clientes segmentados con mayor frecuencia de pagos.* - *Clientes por montos de pagos.*
Importancia	*Alta*
Estado	*Finalizado.*

Tabla 2. 6 - Ficha Técnica. REQ-005.

(Fuente: Elaboración propia)

2.6 Roles y Responsabilidades.

Dentro de la definición de los roles del proyecto implementado se definieron los siguientes:

Gerente de Proyectos: es la persona encargada que tiene la responsabilidad total del planeamiento y ejecución acertada del proyecto.

Ingeniero de Sistemas: es la persona encargada del diseño, desarrollo e implementación del proyecto.

Jefe de Mercadeo: persona encargada de realizar la respectiva revisión del módulo a implementar.

A continuación, describimos los roles que serán asignado al proyecto a implementar.

Rol	Responsable	Abreviación
Gerente de Proyecto	Mario Pincay	G
Ingeniero de Sistemas	Mario Pincay	I
Jefe de Mercadeo	Lorena Zambrano	M

Tabla 2. 7 – Tabla de Roles.

(Fuente: Elaboración propia)

Definido los roles del proyecto, se procede con la asignación de cada actividad al rol especifico.

Actividades		Roles y Responsabilidades	
Id Actividad	**Actividad**	**Rol**	**Responsable**
1	**Diagnosticar el estado Actual del Proceso**	G	Mario Pincay
2	**Análisis de datos facilitados por la institución bancaria**	G	Mario Pincay
3	**Desarrollo del modelo de Datos**	I	Mario Pincay
4	**Diseño de base para Datamark**	I	Mario Pincay
5	**Desarrollo de los ETL**	I	Mario Pincay
6	**Desarrollo del módulo de visualización**	I	Mario Pincay
7	**Diseño del Dashboard**	I	Mario Pincay
8	**Visualización del Dashboard**	M	Lorena Zambrano

Tabla 2. 8 – Tabla de Actividades y Roles.

(Fuente: Elaboración propia)

2.7 Prototipos.

Dentro del prototipo que se ha planteado en el diseño del Dashboard, se procederá implementar los siguientes reportes:

- Reporte de pagos por día
- Reporte de pagos por Establecimientos
- Reporte de pagos por hora
- Reporte de pagos por Recencia
- Reporte de pagos por Frecuencia
- Reporte General de Segmentos

Estos reportes serán diseñados en la aplicación de Power BI, así mismo se creará una aplicación web para poder visualizarlos de una manera rápida, el cual será útil para el usuario final, en el siguiente grafico se visualiza el prototipo a desarrollar.

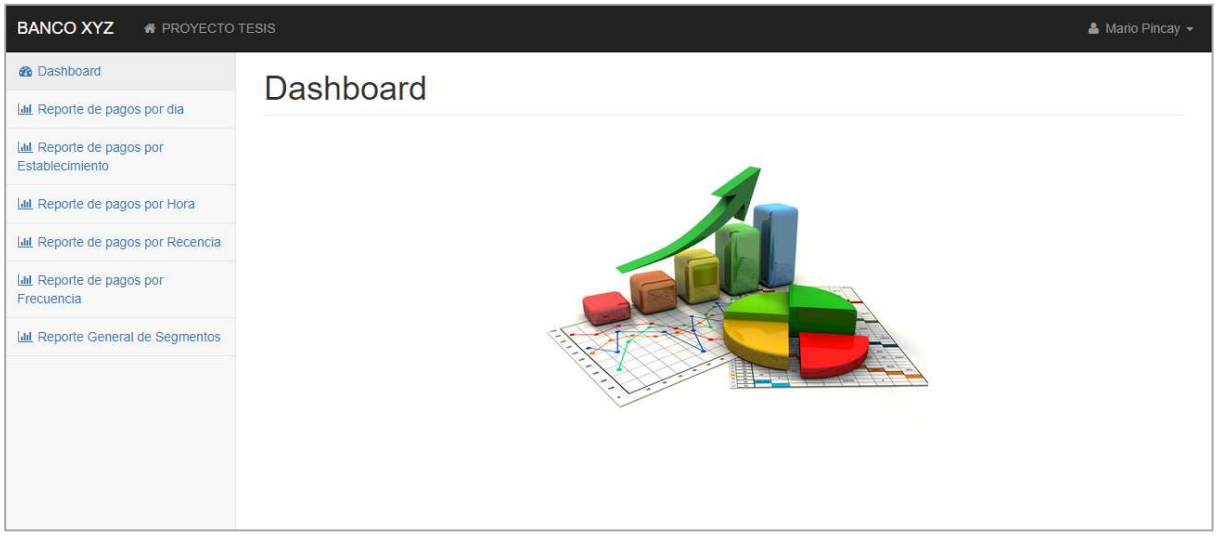


Figura 2. 1 - Prototipo del Dashboard.

(Fuente: Elaboración propia)

2.8 Componentes de la Arquitectura Tecnológica.

Componente que conforman la arquitectura de BI, se describe a continuación en el siguiente gráfico.

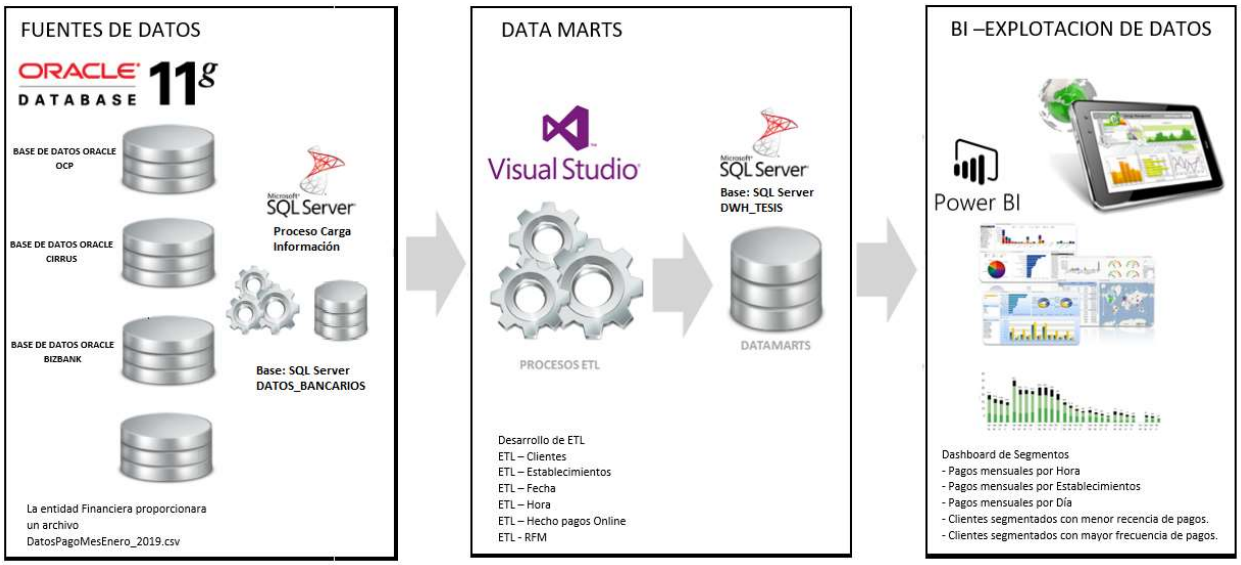


Figura 2. 2 - Arquitectura BI.

(Fuente: Elaboración propia)

2.9 Diseño del modelo de Datos.

La elaboración del diseño del modelo dimensional seguirá el procedimiento basado en cuatro pasos de (Ralph Kimball, 2002).

1. Elegir el proceso de negocio
2. Definir el nivel de granularidad del proceso del negocio
3. Escoger las dimensiones que aplican en cada fila de la tabla de hechos
4. Identificar los hechos numéricos que poblaran la tabla de hechos

2.9.1 Seleccionar el proceso del negocio a modelar

Uno de los objetivos de la entidad financiera es poder ofrecer productos acordes a cada perfil de sus clientes con esto se logrará brindar de una manera eficaz a cada cliente un tipo de producto que le brindará una mejor alternativa en base a su perfil

2.9.2 Definir el nivel de granularidad del proceso del negocio

La granularidad está determinada por cada transacción; esto quiere decir que cada unidad mínima de la carga de los pagos On-Line generara mínimamente una entrada en la tabla de hechos del datamart.

2.9.3 Escoger las dimensiones que aplican en cada fila de la tabla de hechos

Se ha considerado cuatro dimensiones primarias como lo mencionamos a continuación:

DIM_CLIENTES, DIM_ESTABLECIMIENTOS, DIM_FECHAS, DIM_HORA.

En la tabla 2.9 se describe las dimensiones a considerar en el modelo propuesto:

Dimensión	Descripción
DIM_CLIENTES	Tabla dimensión de clientes contiene los datos básicos que posee el cliente, como código, Nombre, dirección y teléfono
DIM_ESTABLECIMIENTOS	Tabla dimensión de establecimientos contiene los datos de cada establecimiento que se realiza el pago Online.
DIM_FECHAS	Tabla dimensión de fecha contiene las fechas de todos los pagos diarios.
DIM_HORA	Dentro de la dimensión de hora contiene las horas de los pagos generados estas están comprendidas desde las 00h00 hasta las 24h00.
HEC_PAGOS_ONLINE	Tabla hechos contiene todo el proceso general de los pagos Online el cual está vinculada con todas las dimensiones del modelo.

Tabla 2. 9 - Descripción de Dimensiones y Hechos

(Fuente: Elaboración propia)

Todas las tablas se definen una clave artificial; no se hará uso de la clave primaria transaccional por temas de estandarización de datos y desempeño. El modelo está pensado para ambos enfoques de minería de datos; tanto la centralizada, así como la distribuida igualmente debería funcionar sin tener que realizar mayores cambios al modelo.

2.9.4 Identificar los hechos que poblaran la tabla de hechos

Adicionalmente, a las dimensiones una peculiaridad está dada en el uso de una dimensión degenerada: CodigoEstablecimieno_pk; esto producto de la granularidad fina definida para el modelo; se ha incorporado una clave artificial, numérica: IdRegistro con el objetivo de mantener un desempeño de consulta aceptable, así como lograr una independencia con respecto a las llaves naturales de los ambientes transaccionales, los cuales serán las fuentes que permitirán poblar el modelo dimensional. La arquitectura según la normalización de

sus dimensiones del modelo dimensional resultante es un esquema copo de nieve que se describe a continuación.

La figura 2.3 presenta el diseño propuesto.

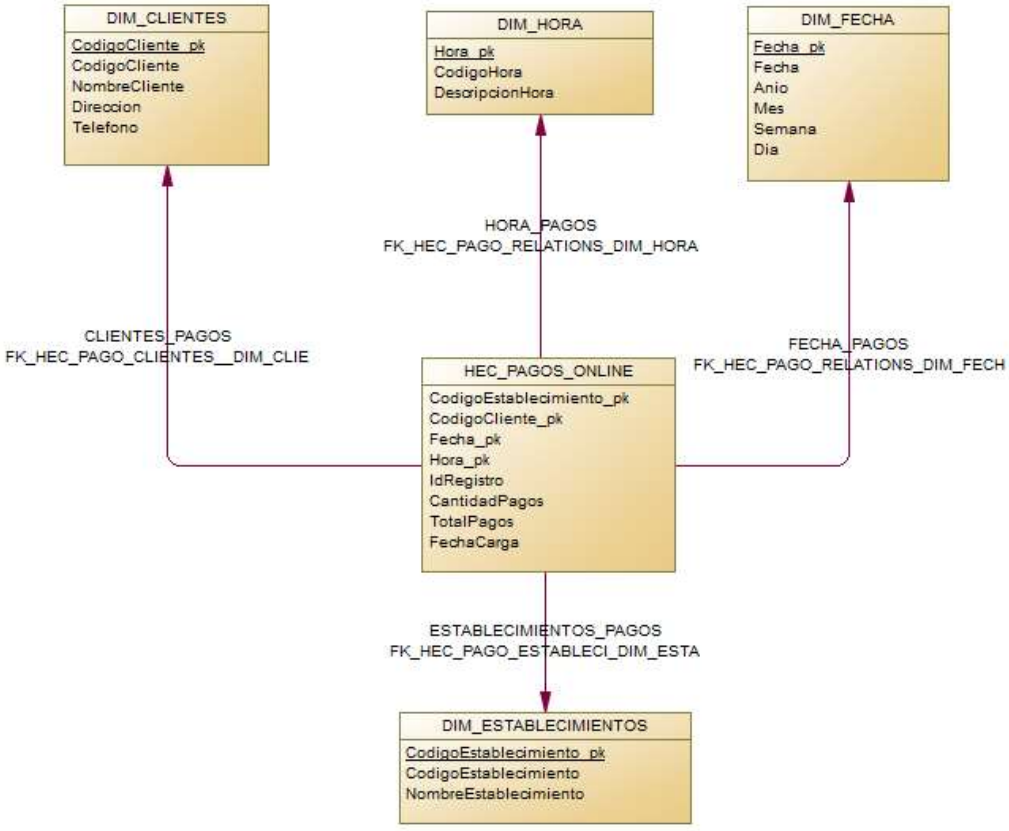


Figura 2. 3 - Modelo Dimensional Copo de Nieve

(Fuente: Elaboración Propia)

Después del diseño del modelo dimensional se procede con la creación de las estructuras de los datos que serán generados desde del Power Design por medio de script que será procesado en el servidor SQL Server ver gráfica.

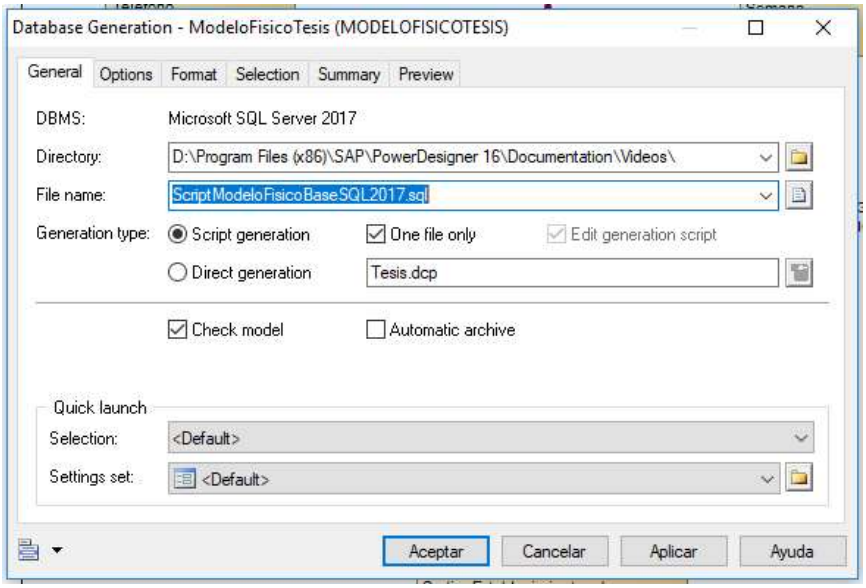
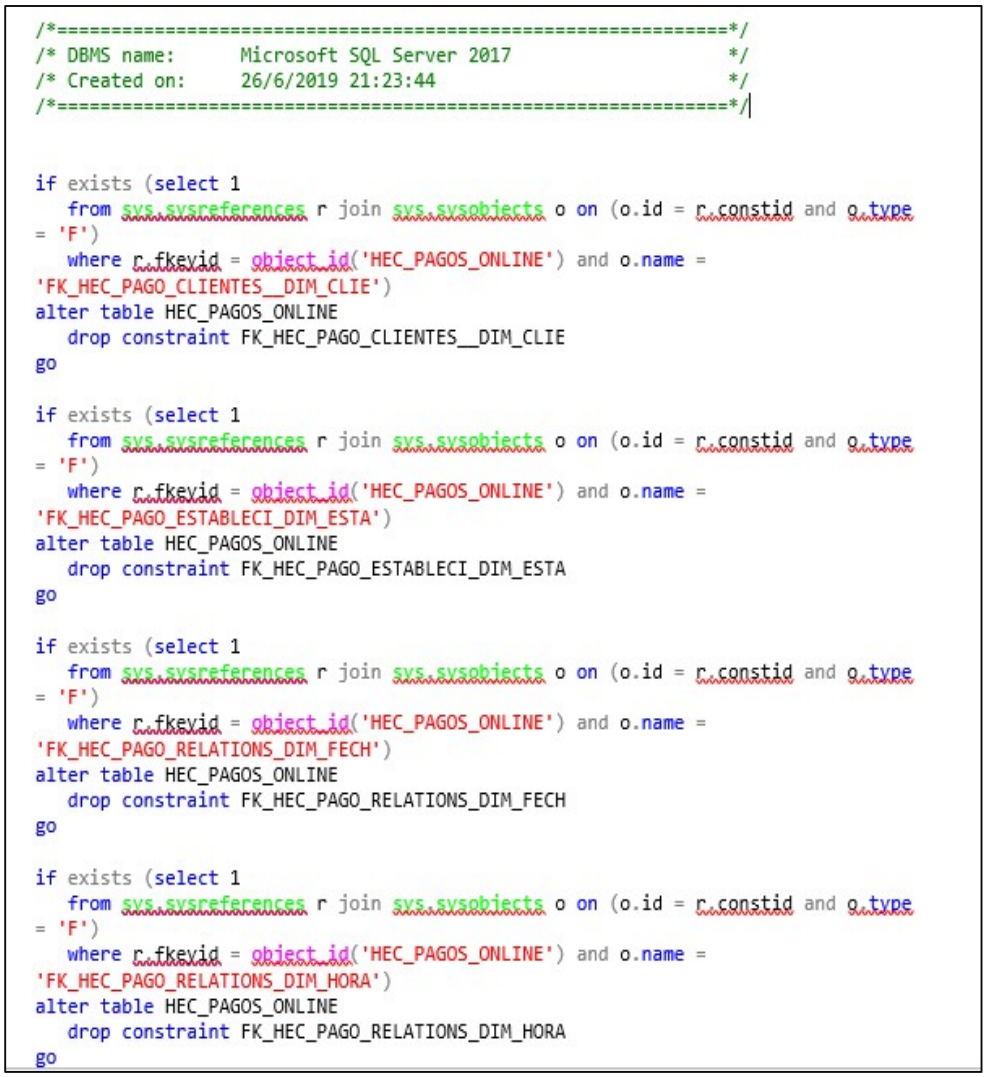


Figura 2. 4 - Grafico de generación de script DDL SQL

(Fuente: Elaboración Propia)

Script generado desde el Power Design el cual creara las estructuras del modelo copo de nieve a la base de datos MS SQL Server, ver la siguiente imagen.



```
/*==============================================================*/
/* DBMS name:      Microsoft SQL Server 2017                    */
/* Created on:     26/6/2019 21:23:44                           */
/*==============================================================*/


if exists (select 1
   from sys.sysreferences r join sys.sysobjects o on (o.id = r.constid and o.type
= 'F')
   where r.fkeyid = object_id('HEC_PAGOS_ONLINE') and o.name =
'FK_HEC_PAGO_CLIENTES__DIM_CLIE')
alter table HEC_PAGOS_ONLINE
   drop constraint FK_HEC_PAGO_CLIENTES__DIM_CLIE
go

if exists (select 1
   from sys.sysreferences r join sys.sysobjects o on (o.id = r.constid and o.type
= 'F')
   where r.fkeyid = object_id('HEC_PAGOS_ONLINE') and o.name =
'FK_HEC_PAGO_ESTABLECI_DIM_ESTA')
alter table HEC_PAGOS_ONLINE
   drop constraint FK_HEC_PAGO_ESTABLECI_DIM_ESTA
go

if exists (select 1
   from sys.sysreferences r join sys.sysobjects o on (o.id = r.constid and o.type
= 'F')
   where r.fkeyid = object_id('HEC_PAGOS_ONLINE') and o.name =
'FK_HEC_PAGO_RELATIONS_DIM_FECH')
alter table HEC_PAGOS_ONLINE
   drop constraint FK_HEC_PAGO_RELATIONS_DIM_FECH
go

if exists (select 1
   from sys.sysreferences r join sys.sysobjects o on (o.id = r.constid and o.type
= 'F')
   where r.fkeyid = object_id('HEC_PAGOS_ONLINE') and o.name =
'FK_HEC_PAGO_RELATIONS_DIM_HORA')
alter table HEC_PAGOS_ONLINE
   drop constraint FK_HEC_PAGO_RELATIONS_DIM_HORA
go
```

Figura 2. 5 – Data Definition Language(DDL) - Script SQL

(Fuente: Elaboración Propia)

2.10 Diccionario de Datos

A continuación, se describe detalladamente el diccionario de datos de cada tabla así mismo una pequeña descripción de cada campo.

Nombre de la tabla: DIM_ESTABLECIMIENTOS
Descripción: en esta tabla se encuentra almacenados todos los establecimientos que se encuentran ingresados en la tabla de los Pagos Online

Campo	Tipo	Tamaño	Descripción
CodigoEstablecimiento_pk	Numeric	18	Clave primaria de la tabla Dim_Establecimientos
CodigoEstablecimiento	Integer		Código único del establecimientos
NombreEstablecimiento	Varchar	200	Descripción del establecimiento

Relaciones:

Campos Clave:
PK_DIM_ESTABLECIMIENTOS

Tabla 2. 10 – Detalle de tabla DIM_ESTABLECIMIENTOS.

(Fuente: Elaboración propia)

Nombre de la tabla: DIM_FECHAS
Descripción: en esta tabla se encuentra almacenados todos las fechas de todos los pagos diarios

Campo	Tipo	Tamaño	Descripción
Fecha_pk	Numeric	18	Clave primaria de la tabla Dim_Fecha
Fecha	Datetime		Fecha única de registro por día
Anio	Integer		Año de la fecha de pago
Mes	Integer		Mes de la fecha de pago
Semana	Integer		Semana del año de la fecha de pago
Dia	Integer		Día de la fecha de pago

Relaciones:

Campos Clave:
PK_DIM_FECHA

Tabla 2. 11 - Detalle de tabla DIM_FECHAS.

(Fuente: Elaboración Propia).

Nombre de la tabla: DIM_CLIENTES
Descripción: en esta tabla se encuentra almacenados todos los clientes que se encuentran en la fechas de todos los pagos diarios

Campo	Tipo	Tamaño	Descripción
CodigoCliente_pk	Numeric	18	Clave primaria de la tabla Dim_Clientes
CodigoCliente	Numeric	18	Código única del cliente
NombreCliente	Varchar	100	Nombre del cliente
Direccion	Varchar	100	Dirección del cliente
Telefono	Varchar	30	Teléfono del cliente

Relaciones:

Campos Clave:
PK_DIM_CLIENTES

Tabla 2. 12 - Detalle de tabla DIM_CLIENTES.

(Fuente: Elaboración Propia)

Nombre de la tabla: DIM_HORA
Descripción: en esta tabla se encuentra almacenados todos los clientes que se encuentran en la fechas de todos los pagos diarios

Campo	Tipo	Tamaño	Descripción
Hora_pk	Numeric	18	Clave primaria de la tabla Dim_Hora
CodigoHora	Integer		Código único de la hora
DescripcionHora	Varchar	50	Descripción de la hora

Relaciones:

Campos Clave:
PK_DIM_HORA

Tabla 2. 13 - Detalle de tabla DIM_HORA.

(Fuente: Elaboración Propia)

Nombre de la tabla: HEC_PAGOS_ONLINE
Descripción: en esta tabla se encuentra almacenados la tabla de hechos el cual se almacena todos los datos que serán procesados de la tabla de pagos Online.

Campo	Tipo	Tamaño	Descripción
CodigoEstablecimiento_pk	Integer		Clave primaria de la tabla Dim_Establecimientos
CodigoCliente_pk	Integer		Clave primaria de la tabla Dim_Clientes
Fecha_pk	Integer		Clave primaria de la tabla Dim_Fecha
Hora_pk	Integer		Clave primaria de la tabla Dim_Hora
IdRegistro	Numeric	18	Codigo único de registros del a tabla hechos
CantidadPagos	Numeric	18	Cantidad de pagos por clientes
TotalPagos	Decimal	(18,2)	Total de pagos por clientes
FechaCarga	Datetime		Fecha de carga de registro

Relaciones:
CodigoEstablecimiento_pk con PK_DIM_ESTABLECIMIENTOS
CodigoCliente_pk con PK_DIM_CLIENTES
Fecha_pk con PK_DIM_FECHA
Hora_pk con PK_DIM_HORA

Campos Clave:
CodigoEstablecimiento_pk
CodigoCliente_pk
Fecha_pk
Hora_pk

Tabla 2. 14 - Detalle de tabla HEC_PAGOS_ONLINE.

(Fuente: Elaboración Propia)

2.11 Diseño de ETL

Dentro de la transformación de los datos se procedió a la creación de los siguientes ETL, para almacenar la información desde la fuente de datos proporcionada por la entidad financiera hacia la base de datos SQL Server 2017.

A continuación, se procederá describir cada proceso ETL diseñado en el Integration Services del Visual Studio .Net

2.11.1 ETL Grabar DIM_ETABLECIMIENTOS

Dentro del diseño se procederá con la configuración del ETL de establecimientos para poder almacenar la información que se necesita, la configuración es la siguiente.

```
SELECT
    ROW_NUMBER() OVER (ORDER BY Establecimiento ASC) AS CodigoEstablecimiento
    , CONVERT(VARCHAR(100), ESTABLECIMIENTO) AS Establecimientos
FROM            DATOS_PAGOS_ONLINE
GROUP BY Establecimiento
```

En la siguiente figura se aprecia la ejecución del ETL

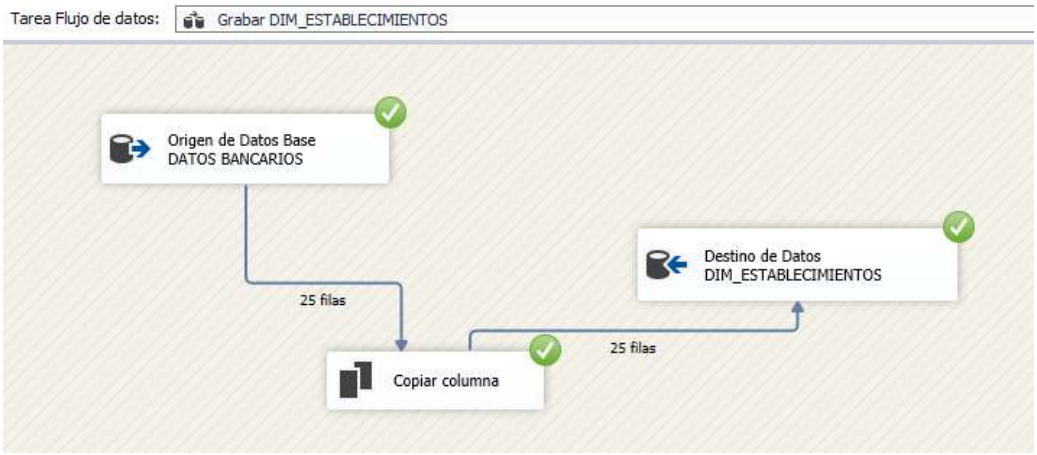
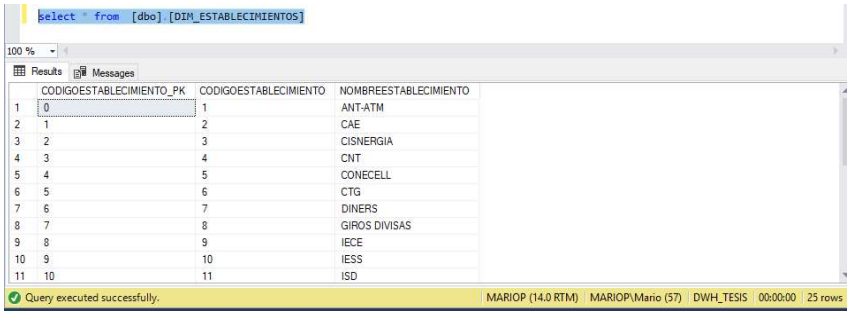


Figura 2. 6 - Ejecución del ETL Establecimientos

(Fuente: Elaboración Propia)

Se confirma dentro de la base del SQL que los datos fueron almacenados en la nueva estructura.



```
select * from [dbo].[DIM_ESTABLECIMIENTOS]
```

100 %

Results | Messages

	CODIGOESTABLECIMIENTO_PK	CODIGOESTABLECIMIENTO	NOMBREESTABLECIMIENTO
1	0	1	ANT-ATM
2	1	2	CAE
3	2	3	CISNERGIA
4	3	4	CNT
5	4	5	CONECELL
6	5	6	CTG
7	6	7	DINERS
8	7	8	GIROS DIVISAS
9	8	9	IECE
10	9	10	IESS
11	10	11	ISD

Query executed successfully. | MARIOP (14.0 RTM) | MARIOP\Mario (57) | DWH_TESIS | 00:00:00 | 25 rows

Figura 2. 7 - Tabla DIM_ESTABLECIMIENTOS.

(Fuente: Elaboración Propia)

2.11.2 ETL Grabar DIM_HORAS

Dentro del diseño se procederá con la configuración del ETL para poder almacenar la información de horas, la configuración es la siguiente.

```
SELECT DATEPART(HH, HoraMinutoSegundo) AS Hora
, 'Hora' AS Descripcion
FROM            DATOS_PAGOS_ONLINE
GROUP BY DATEPART(HH, HoraMinutoSegundo)
ORDER BY Hora
```

En la siguiente figura se aprecia la ejecución del ETL

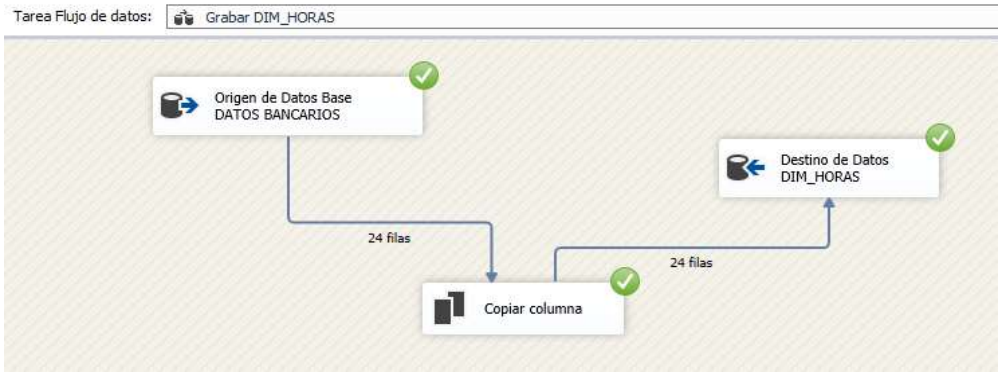


Figura 2. 8 - Ejecución del ETL Horas

(Fuente: Elaboración Propia)

Se confirma dentro de la base del SQL que los datos fueron almacenados en la nueva estructura.



	HORA_PK	CODIGOHORA	DESCRIPCIONHORA
1	0	0	Hora
2	1	1	Hora
3	2	2	Hora
4	3	3	Hora
5	4	4	Hora
6	5	5	Hora
7	6	6	Hora
8	7	7	Hora
9	8	8	Hora
10	9	9	Hora
11	10	10	Hora

Figura 2. 9 - Tabla DIM_HORA.

(Fuente: Elaboración Propia)

2.11.3 ETL Grabar DIM_FECHAS

Dentro del diseño se procederá con la configuración del ETL para poder almacenar la información de fechas, la configuración es la siguiente.

```
SELECT  FechaMovimiento, YEAR(FechaMovimiento) AS Anio
, MONTH(FechaMovimiento) AS Mes
, DATEPART(ISO_WEEK, FechaMovimiento) AS Semana
, DAY(FechaMovimiento) AS Dia
FROM     DATOS_PAGOS_ONLINE
GROUP BY FechaMovimiento
ORDER BY FechaMovimiento
```

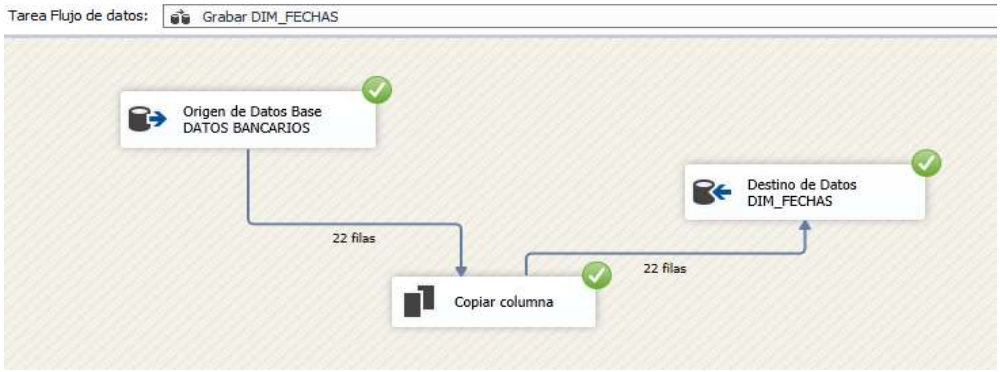
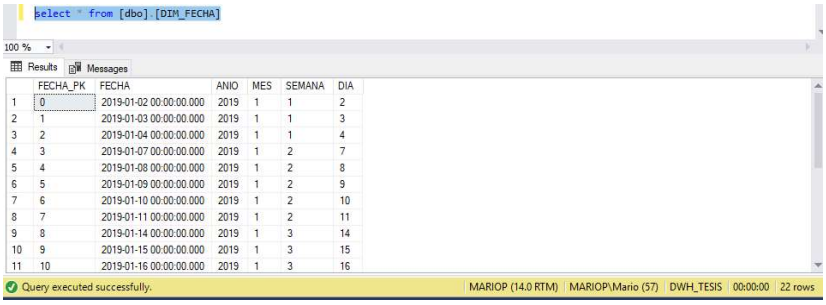


Figura 2. 10 - Ejecución del ETL Fechas.

(Fuente: Elaboración Propia)

Se confirma dentro de la base del SQL que los datos fueron almacenados en la nueva estructura.



	FECHA_PK	FECHA	ANIO	MES	SEMANA	DIA
1	0	2019-01-02 00:00:00.000	2019	1	1	2
2	1	2019-01-03 00:00:00.000	2019	1	1	3
3	2	2019-01-04 00:00:00.000	2019	1	1	4
4	3	2019-01-07 00:00:00.000	2019	1	2	7
5	4	2019-01-08 00:00:00.000	2019	1	2	8
6	5	2019-01-09 00:00:00.000	2019	1	2	9
7	6	2019-01-10 00:00:00.000	2019	1	2	10
8	7	2019-01-11 00:00:00.000	2019	1	2	11
9	8	2019-01-14 00:00:00.000	2019	1	3	14
10	9	2019-01-15 00:00:00.000	2019	1	3	15
11	10	2019-01-16 00:00:00.000	2019	1	3	16

Figura 2. 11 - Tabla DIM_FECHA.

(Fuente: Elaboración Propia)

2.11.4 ETL Grabar DIM_CLIENTES

Dentro del diseño se procederá con la configuración del ETL para poder almacenar la información de Clientes, la configuración es la siguiente.

```
SELECT        CodigoCliente
, 'MARXXXXXXXXXXPIN' AS NombreCliente
, 'G.NORTEXXXXXXXX' AS Direccion
, '043731500' AS Telefono
FROM            DATOS_PAGOS_ONLINE
GROUP BY CodigoCliente
ORDER BY CodigoCliente
```

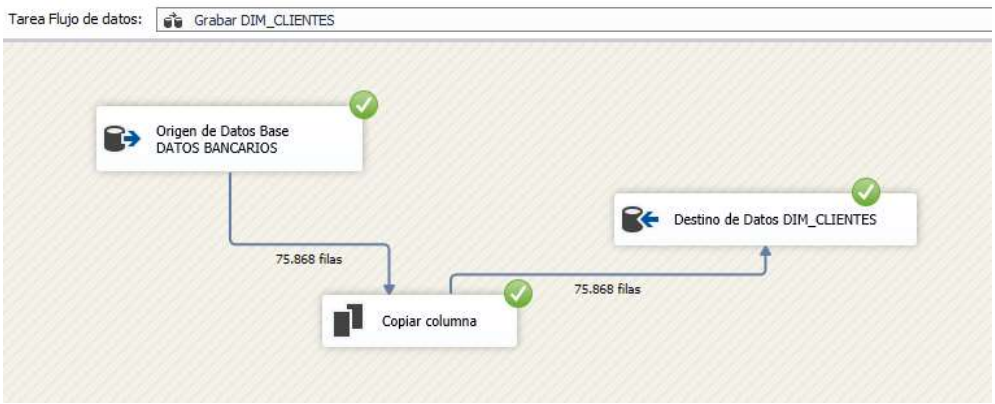
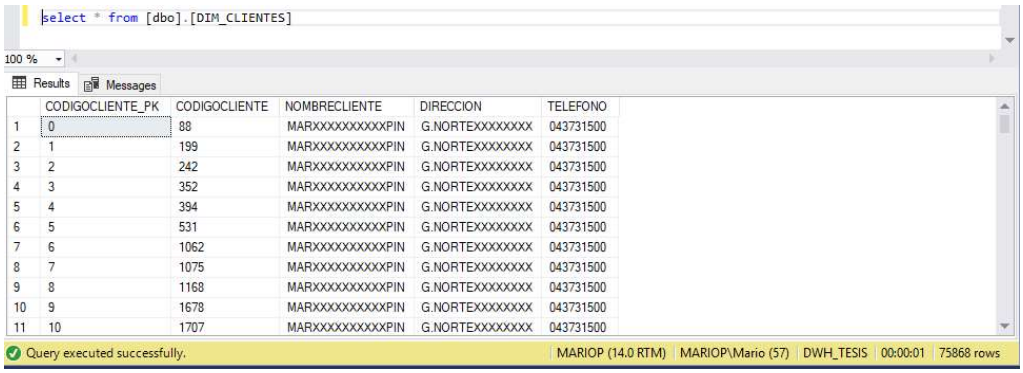


Figura 2. 12 - Ejecución del ETL Clientes.

(Fuente: Elaboración Propia)

Se confirma dentro de la base del SQL que los datos fueron almacenados en las nuevas tablas.



	CODIGOCLIENTE_PK	CODIGOCLIENTE	NOMBRECLIENTE	DIRECCION	TELEFONO
1	0	88	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
2	1	199	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
3	2	242	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
4	3	352	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
5	4	394	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
6	5	531	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
7	6	1062	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
8	7	1075	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
9	8	1168	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
10	9	1678	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500
11	10	1707	MARXXXXXXXXXXPIN	G.NORTEXXXXXXXX	043731500

Figura 2. 13 - Tabla DIM_CLIENTES.

(Fuente: Elaboración Propia)

Finalizado el diseño y la ejecución del proceso de todas las dimensiones se establece la configuración del ETL completo el cual se procederá llenar la tabla de hechos y las tablas del modelo RFM

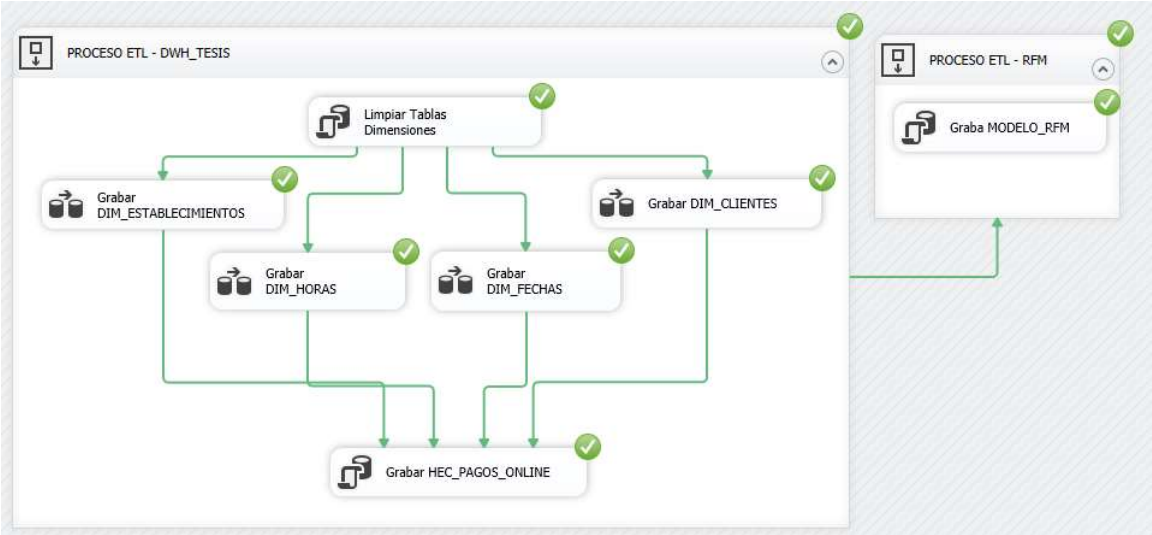


Figura 2. 14 - Ejecución de Procesos ETL.

(Fuente: Elaboración Propia)

Se confirma dentro de la base del SQL que los datos fueron almacenados en la nueva tabla de hechos y la RFM

```
select * from [dbo].[HEC_PAGOS_ONLINE]
```

100 %

Results | Messages

	CODIGOESTABLECIMIENTO_PK	CODIGOCLIENTE_PK	FECHA_PK	HORA_PK	IDREGISTRO	CANTIDADPAGOS	TOTALPAGOS	FECHACARGA
1	3	840610	0	0	1	1	24.04	2019-06-26 21:47:29.423
2	3	2190251	0	0	2	1	71.43	2019-06-26 21:47:29.423
3	3	2442709	0	0	3	1	31.76	2019-06-26 21:47:29.423
4	3	2456620	0	0	4	1	7.03	2019-06-26 21:47:29.423
5	3	2719293	0	0	5	1	66.53	2019-06-26 21:47:29.423
6	3	2721517	0	0	6	1	47.70	2019-06-26 21:47:29.423
7	3	2799439	0	0	7	1	7.35	2019-06-26 21:47:29.423
8	3	2827333	0	0	8	1	20.00	2019-06-26 21:47:29.423
9	3	3681114	0	0	9	1	34.48	2019-06-26 21:47:29.423
10	7	262920	0	0	10	1	3000.00	2019-06-26 21:47:29.423
11	7	372310	0	0	11	1	60.00	2019-06-26 21:47:29.423

Query executed successfully. | MARIOP (14.0 RTM) | MARIOP\Mario (57) | DWH_TESIS | 00:00:03 | 200254 rows

Figura 2. 15 - Tabla HEC_PAGOS_ONLINE.

(Fuente: Elaboración Propia)

```
select * from MODELO_RFM
```

100 %

Results | Messages

	RegistroId	Anio	DescripcionMes	CodigoCliente	rfm_recency	rfm_frequency	rfm_monetary	rfm_score_clasico	rfm_score_new	rfm_score	rfm_avg
1	1	2019	Enero	957389	5	4	2	40	83	542	4
2	2	2019	Enero	3769497	5	1	2	10	63	512	3
3	3	2019	Enero	3509359	4	1	1	4	50	411	2
4	4	2019	Enero	2462161	5	4	1	20	79	541	4
5	5	2019	Enero	3019957	5	5	1	25	86	551	4
6	6	2019	Enero	2158246	5	4	2	40	83	542	4
7	7	2019	Enero	2211413	4	2	2	16	59	422	3
8	8	2019	Enero	2180743	4	2	1	8	56	421	3
9	9	2019	Enero	3784467	5	1	2	10	63	512	3
10	10	2019	Enero	2100391	5	2	2	20	69	522	3
11	11	2019	Enero	2079120	5	1	1	5	59	511	3
12	12	2019	Enero	3510887	5	1	2	10	63	512	3

Query executed successfully. | MARIOP (14.0 RTM) | MARIOP\Mario (55) | DWH_TESIS | 00:00:01 | 75868 rows

Figura 2. 16 - Tabla MODELO_RFM.

(Fuente: Elaboración Propia)

2.11.5 Modelo RFM

Dentro de la implementación del modelo RFM, se procederá tener en claro que variables fueron analizadas para darle valor y calcular los nuevos datos generados, a continuación, se describen la escala de cada variable para ser analizadas y evaluada.

Tabla de Frecuencias.

Aquí se evidencia la escala con el cual se realizó el análisis de la frecuencia de los datos de la entidad bancaria.

Tabla Frecuencia		
Indicador	Transacciones Promedio	
1	1	Menos Frecuente
2	2	Frecuente
3	3	Usuales
4	4 - 6	Medios
5	7	Más Frecuentes

Tabla 2. 15 – Tabla de Frecuencias.

(Fuente: Elaboración Propia)

Tabla de Recencia.

Aquí se evidencia la escala con el cual se realizó el análisis de la frecuencia de los datos de la entidad bancaria.

Tabla Recencia		
Indicador	Ultima Transacción Promedio	
5	1 – 5	Más Frecuente
4	6 – 10	Medios
3	11 – 15	Usuales
2	16 - 20	Frecuente
1	20 < 30	Menos Frecuentes

Tabla 2. 16 – Tabla Recencia.

(Fuente: Elaboración Propia)

Tabla de Monetario.

Aquí se evidencia la escala con el cual se realizó el análisis del valor Monetario de los datos de la entidad bancaria.

Tabla Monetario		
Indicador	Valores Rango de Promedio	
1	< 150	Menos Frecuente
2	150 - < 2000	Frecuente
3	2000 - 3500	Usuales
4	3500 – 4500	Medios
5	> 4500	Más Frecuentes

Tabla 2. 17 – Tabla Monetario.

(Fuente: Elaboración Propia)

2.12 Desarrollo del módulo de visualización.

Dentro del desarrollo del módulo de visualización se procederá realizar un Dashboard el cual los usuarios podrán visualizar de una manera rápida los gráficos estadísticos creados en el módulo.

En el desarrollo del Dashboard se implementó la siguiente aplicación web que mostrará los informes realizados en la herramienta de Power BI.

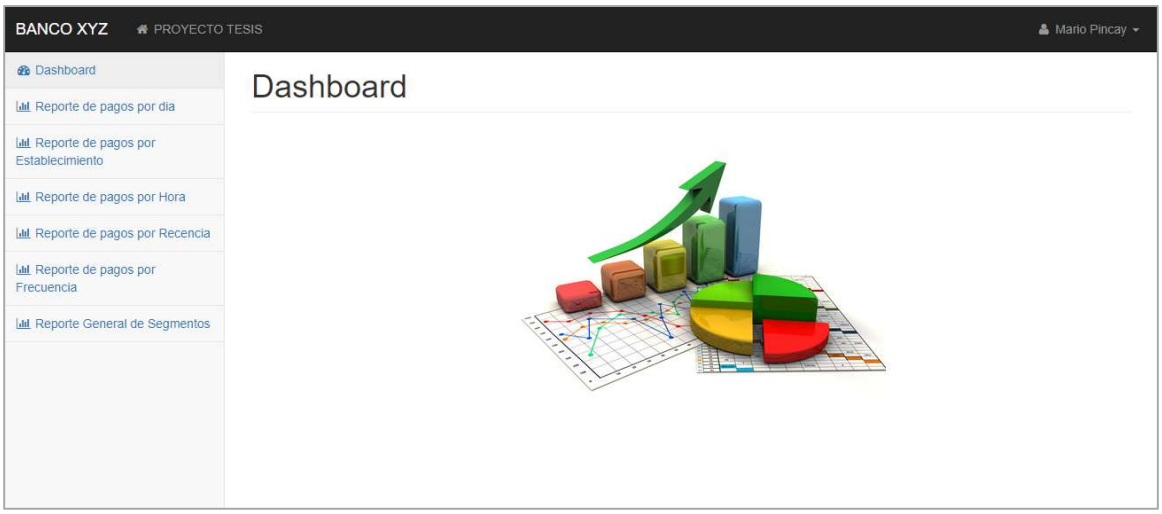


Figura 2. 17 - Dashboard.

(Fuente: Elaboración propia)

A continuación, se describirá de una manera general cada reporte implementado con el objetivo de poder describir la usabilidad en el Dashboard.

2.12.1 Reporte de pagos por horas.

En este reporte se analiza la información de los pagos On-Line generada en todo el mes de Enero/2019, en la gráfica se evidencia la hora que más se realizan pagos.

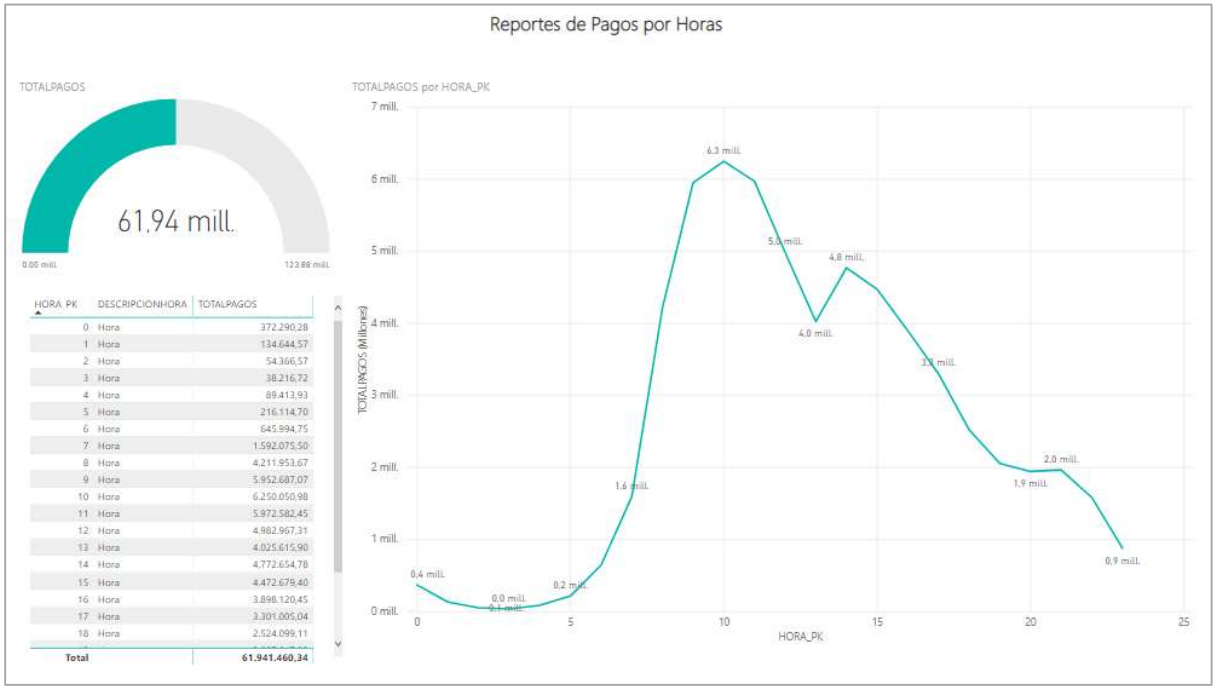


Figura 2. 18 - Grafico de pagos por Horas.

(Fuente: Elaboración propia)

2.12.2 Reporte de pagos por Establecimientos.

En este reporte se analiza la información de los pagos On-Line generada en todo el mes de Enero/2019, en la gráfica se evidencia el establecimiento que tiene más recaudaciones realizadas en el mes.

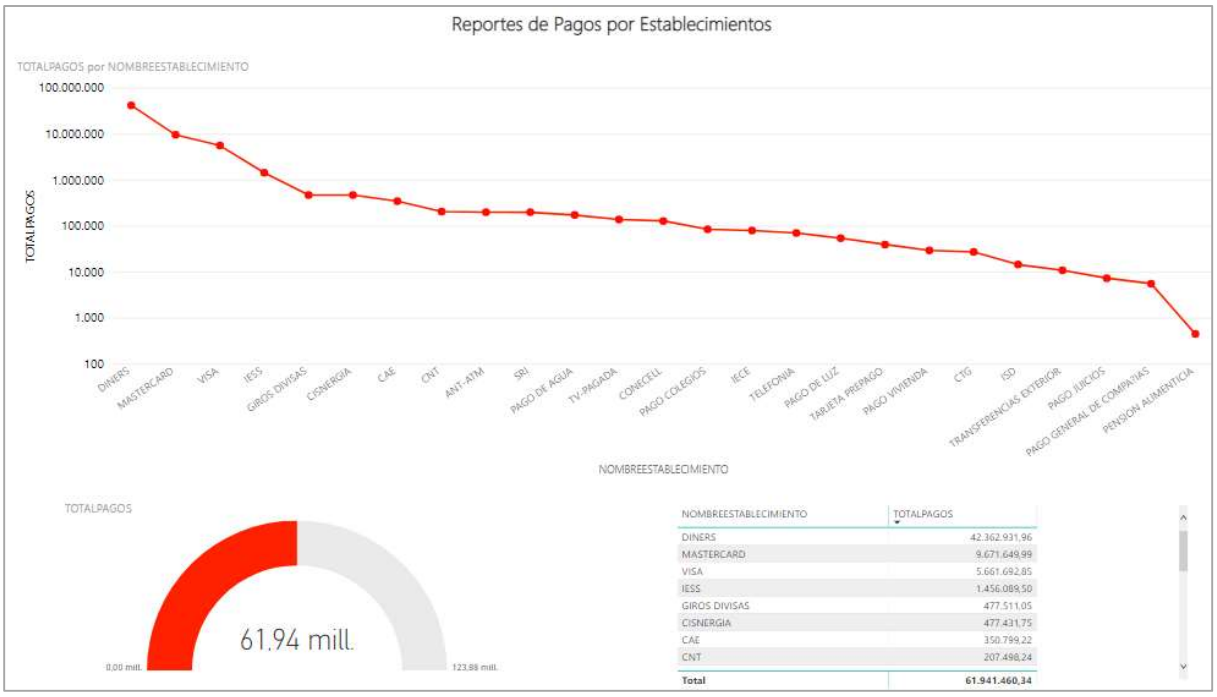
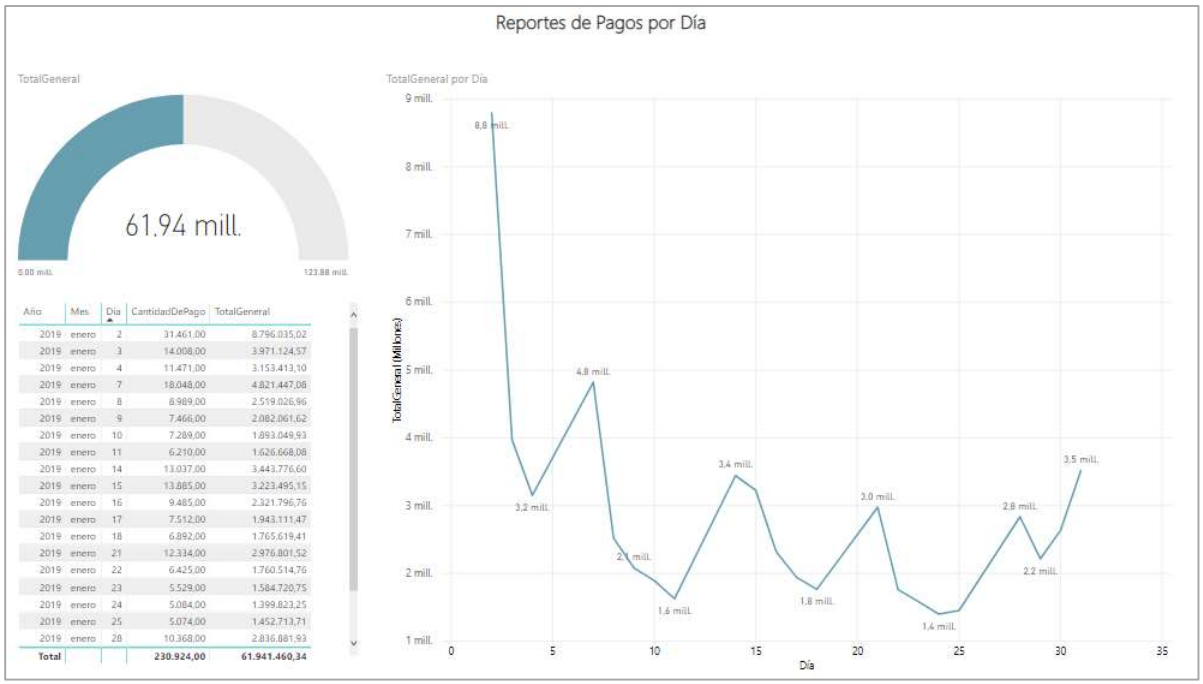


Figura 2. 19 - Grafico de pagos por Establecimientos.

(Fuente: Elaboración propia)

2.12.3 Reporte de pagos por Día.

En este reporte se analiza la información de los pagos Online generada en todo el mes de Enero/2019, en la gráfica se evidencia los días que más se han recaudado en el mes.



Año	Mes	Día	CantidadDePago	TotalGeneral
2019	enero	2	31.461,00	8.796.035,02
2019	enero	3	14.008,00	3.971.124,57
2019	enero	4	11.471,00	3.153.413,10
2019	enero	7	18.048,00	4.821.447,08
2019	enero	8	8.989,00	2.519.026,96
2019	enero	9	7.466,00	2.082.061,62
2019	enero	10	7.289,00	1.893.049,93
2019	enero	11	6.210,00	1.626.668,08
2019	enero	14	13.037,00	3.443.776,60
2019	enero	15	13.885,00	3.223.495,15
2019	enero	16	9.485,00	2.321.796,76
2019	enero	17	7.512,00	1.943.111,47
2019	enero	18	6.892,00	1.765.619,41
2019	enero	21	12.334,00	2.976.801,52
2019	enero	22	6.425,00	1.760.514,76
2019	enero	23	5.529,00	1.584.720,75
2019	enero	24	5.084,00	1.399.823,25
2019	enero	25	5.074,00	1.452.713,71
2019	enero	28	10.368,00	2.836.881,93
Total			**230.924,00**	**61.941.460,34**

Figura 2. 20 - Grafico de pagos por Día.

(Fuente: Elaboración propia)

2.13 Presupuesto

A continuacion detallamos el presupuesto que sera contemplado en el desarrollo del proyecto final.

a. Talento Humano

Dentro del personal que participo en la realización del presente proyecto fue el Gerente de Proyecto y el Analista de Sistemas, en la tabla siguiente se describe el costo del Recurso Humano.

Recursos Humanos			
Rol	**Cantidad**	**Remuneración**	**Valor total**
Gerente de Proyecto	1	3500	3500.00
Analista de Sistemas	1	1000	1000.00
		Total($)	4500.00

Tabla 2. 18 - Recursos Humanos.

(Fuente: Elaboración propia)

b. Servicios

Los servicios de internet, luz y teléfono también fueron requeridos para la realización del proyecto, en la tabla siguiente se describe el costo de estos servicios.

Servicios			
Unidad	**Cantidad de horas**	**Valor por hora($)**	**Valor total($)**
Internet	200 horas	0.50	100.00
Luz	400 horas	0.10	40.00
Teléfono	1 hora	9.00	9.00
		Total($)	149.00

Tabla 2. 19 – Servicios.

(Fuente: Elaboración propia)

c. Recursos Hardware

Los recursos hardware utilizados en el proyecto fue una Laptop, en la tabla siguiente se describe el costo de este recurso.

Recurso de Hardware			
Unidad	**Cantidad**	**Valor unitario($)**	**Valor total($)**
Laptop	1	850.00	850.00
		Total($)	850.00

Tabla 2. 20 – Recurso de hardware.

(Fuente: Elaboración propia)

d. Recursos Software

Para la realización del presente proyecto se utilizaron los siguientes recursos de software: Sql Server, Power BI, Power Design, Visual Studio, en la tabla siguiente se describe el costo de estos recursos.

Recurso de Software			
Unidad	**Cantidad**	**Valor unitario($)**	**Valor total($)**
Licencia SQL Standard	1	931.00	931.00
Licencia Power BI	2 mes	9.99	19.98
Licencia Power Design	1	3750.00	750.00
Licencia de Visual Studio	1	700.00	700.00
		Total($)	2400.98

Tabla 2. 21 – Recurso de Software.

(Fuente: Elaboración propia)

e. Datos

Los datos utilizados dentro del proyecto de investigación fueron proporcionados por la entidad bancaria. A continuación, en la tabla siguiente se describe el costo de los datos.

Datos	
	Valor total($)
Datos de la Empresa	0.00
Total($)	0.00

Tabla 2. 22 – Datos.

(Fuente: Elaboración propia)

f. Presupuesto Total

El presupuesto total corresponde a la suma de los gastos descrito anteriormente, en la tabla siguiente se detalla el costo total para la ejecución del proyecto.

Presupuesto Total	
	Valor total($)
Recursos Humano	4500.00
Servicios	149.00
Recurso de Hardware	850.00
Recurso de Software	2400.98
Datos	0.00
Total($)	7899.98

Tabla 2. 23 – Presupuesto Total.

(Fuente: Elaboración propia)

2.14 Plan del Proyecto.

A continuación detallamos el diagrama de Gantt que se implementar en el desarrollo del proyecto de tesis.

Dentro de la planificación del desarrollo del proyecto se verifico que la duración para el desarrollo e implementación del proyecto será de 35 días, tal como esta evidencia en el cronograma de actividades descrito a continuación.



Figura 2. 21 – Planificación de cronograma de actividades

(Fuente: Elaboración propia)

2.15 Riesgos y Contingencia.

Dentro del proyecto implementado se han tomados los riesgos el cual se pueden presentar en el desarrollo e implementación, lo que deberá tomar las acciones que permitan mitigar los riesgos posibles y tomar acciones para poder canalizarlos.

Riesgos	Contingencia
Problemas con las herramientas de software utilizadas	Manejar las herramientas con datos de pruebas que permitan identificar el alcance y capacidad de las mismas
Datos faltantes en la información	Examinar la información que se va a utilizar y verificar la suficiencia de los datos y parámetros a tomar en cuenta
Retrasos en la planificación de los tiempos.	Establecer un tiempo conveniente para cada una de las actividades que deben realizarse durante el proyecto, considerando la complejidad y posibles inconvenientes en la elaboración de las mismas.

Tabla 2. 24 – Riesgos y Contingencia

(Fuente: Elaboración propia)

Capítulo 3

3. Validación de Resultados

Dentro de la validación de los datos se procedió analizar el entorno del objeto de estudio, el cual se enfocó en los pagos Online de la entidad bancaria, el método de la investigación que se aplico fue la encuesta.

Dentro de la investigación se procedió crear una encuesta a un grupo de personas el cual se plantearon las siguientes preguntas:

Formulario de preguntas Web	
1. Alguna vez ha realizado algún pago por medio de la web?	
	◉ Si ○ No
2. Por qué medio realizas los pagos Web?	
	☑ PayPal ☐ Banca Virtuales del banco de su preferencia ☐ Otros
3. Con que frecuencia mensual realizas los pagos Web?	
	◉ Entre 1 y 5 veces ○ Entre 6 y 15 veces ○ Mayores a 16
4. Qué tipo de pagos realizas?	
	☑ Pago de Tarjetas ☐ Servicios Básicos ☐ Pago de Colegiatura ☑ Otros
5. Seleccione su rango de Edad	
	◉ Entre 18 y 22 años ○ Entre 23 y 30 años ○ Mayores a 31 años

Tabla 3. 1 – Formulario de Encuesta.

(Fuente: Elaboración propia)

Dentro de esta encuesta que se realizó a 111 personas se pudo verificar la tendencia de los pagos Online el cual podemos analizar cada pregunta realizada, el cual se describen a continuación.

- **¿Alguna vez ha realizado algún pago por medio de la web?**

Dentro de la primera pregunta se pudo analizar que el 73,9% que equivalen a 82 personas han realizado por lo menos un pago Online en el último mes tal como se visualiza en el siguiente gráfico.

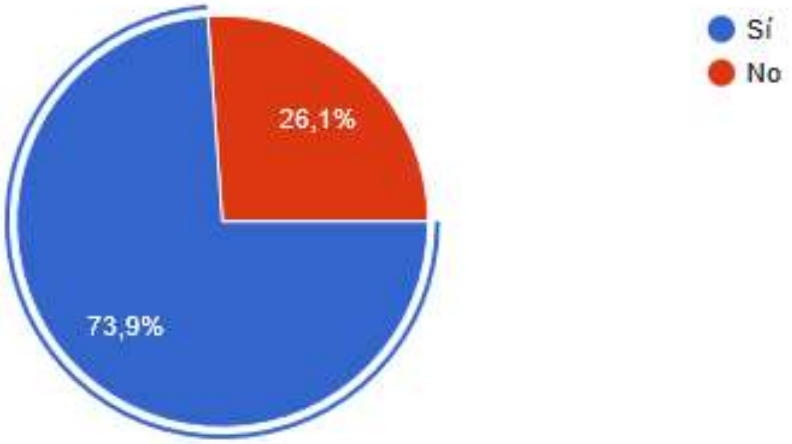


Figura 3. 1 – Grafico de porcentajes de encuestados.
(Fuente: Elaboración propia)

- **¿Por qué medio realizar los pagos Web?**

En la segunda pregunta se pudo analizar que el 92.7% que equivalen a 76 personas han realizado sus pagos web por medio de plataformas digitales de cada institución bancaria tal como se visualiza en el siguiente gráfico.

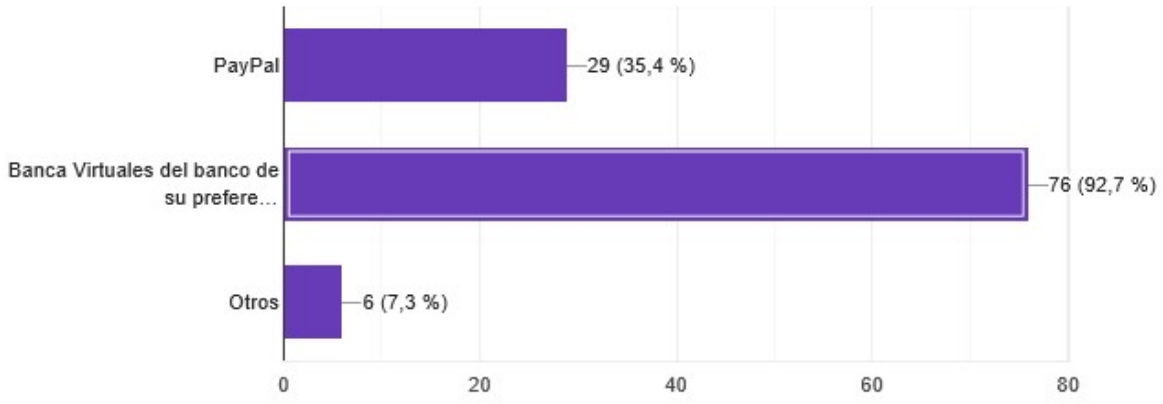


Figura 3. 2 – Grafico de medios de pagos web.

(Fuente: Elaboración propia)

- **¿Con que frecuencia mensual realizas los pagos Web?**

En la tercera pregunta se pudo analizar que el 57.3% que equivalen a 47 personas han realizado por lo menos entre 1 y 5 veces transacciones al menos en el último mes tal como se visualiza en el siguiente gráfico.

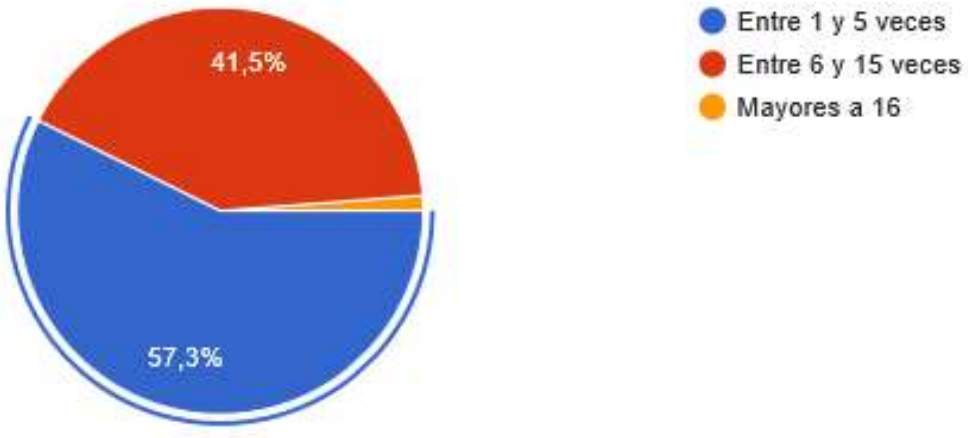


Figura 3. 3 – Grafico de frecuencia de pagos web

(Fuente: Elaboración propia)

- **¿Qué tipo de pagos realizas?**

En la cuarta pregunta se pudo analizar que la tendencia de pagos que se realizan en las plataformas virtuales de cada institución bancaria son de pagos de tarjetas que equivale a un 78% y los servicios básicos que equivales a un 82.9%, tal como se visualiza en el siguiente gráfico.

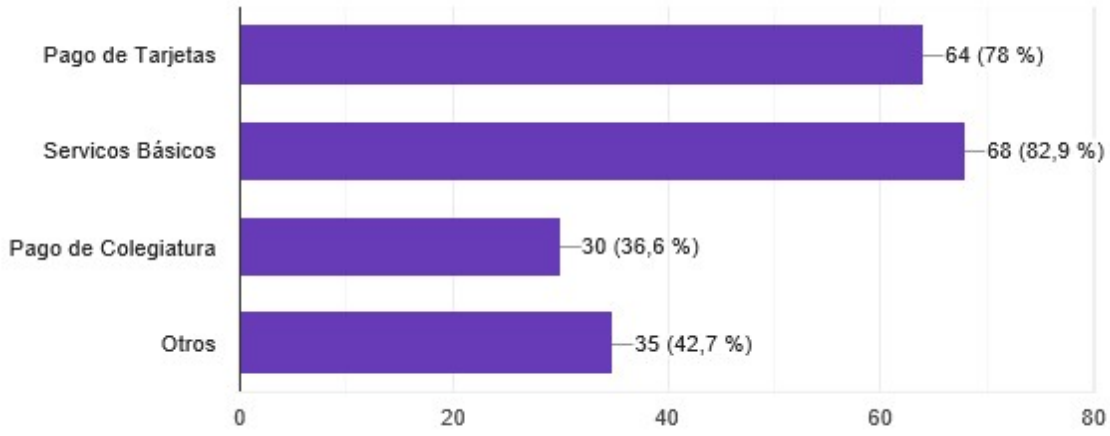


Figura 3. 4 – Grafico de tipo de pagos realizados en plataformas virtuales
(Fuente: Elaboración propia)

- **¿Seleccione su rango de edad?**

En la quita pregunta se pudo analizar que el grupo de personas que realizan estos pagos por medio de plataformas virtuales u otros medios de pagos son personas mayores de 31 años, tal como se visualiza en el siguiente gráfico.

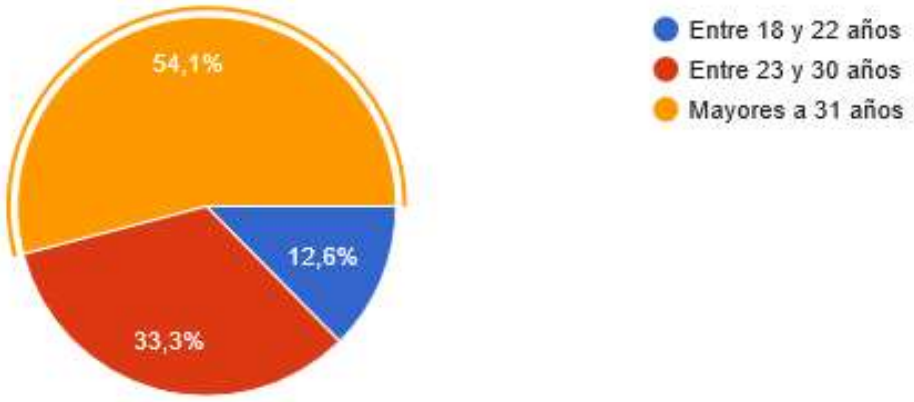


Figura 3. 5 – Grafico tendencia de edades de pagos Web

(Fuente: Elaboración propia)

3.1 Esquema de validación.

Inicialmente, los datos usados para la realización de este estudio, correspondieron a la información presentada en el mes de Enero/2019, de los pagos Online realizados por los clientes de la entidad bancaria.

Dentro del esquema de validación el área de mercadeo de la institución financiera procedió revisar el Dashboard desarrollado el cual indica que cumple con los objetivos planteados así mismo fueron analizados los gráficos generados.

De los datos analizados indicaron:

3.1.1 Análisis del reporte de pagos por hora.

Dentro del reporte de pagos por hora pudieron analizar que la hora que más se realizan los pagos web en la entidad bancaria fue a partir de las 10:00 am, ellos analizaron e informaron

que pueden tomar ventaja con esta información generada en el gráfico y poder ofrecer algún producto para sacar una ventaja competitiva.



Figura 3. 6 – Grafico del reporte de pagos por hora con mayor recaudación
(Fuente: Elaboración propia)

3.1.2 Análisis del reporte de pagos por establecimientos.

Dentro del reporte de establecimiento pudieron analizar que el establecimiento que más recauda durante todo el mes es DINERS, seguido MASTERCARD y otros

establecimientos, ellos analizaron que pueden tomar ventaja competitiva para poder ofrecer productos a cierto grupo de clientes dependiendo de su perfil generado.



Figura 3. 7 – Grafico con mayor recaudación por establecimientos.
(Fuente: Elaboración propia)

3.2 Impacto.

Dentro del proyecto implementado en la entidad bancaria se han tomado los riesgos que se pueden presentar en el desarrollo e implementación, así mismo lo que se deberá evaluar los riesgos posibles y tomar acciones inmediatas.

Desde una perspectiva económica, esta solución tiene un impacto positivo a la entidad bancaria ya que con los reportes generados se podrá establecer una adecuada estrategia de mercadeo y poder ofrecer productos acordes a su perfil.

Conclusiones

Dentro del análisis realizado en el presente trabajo de investigación se han llegado a las siguientes conclusiones:

- La implementación del presente proyecto en la entidad bancaria para el área de mercadeo, apoyaría en el cumplimiento de los objetivos y el fortalecimiento de las campañas de mercadeo, por el cual se lograría tener una mejor visión de los productos o servicios a ofertar.
- Se implementó una arquitectura de BI, para poder cumplir con los objetivos planteados en el proyecto.
- Se aplicó el modelo RFM (Reciente, Frecuencia, Monetario) con la información proporcionada por la entidad bancaria de los pagos Online, el cual se tuvo una perspectiva clara de los clientes, para ofrecer campañas de mercadeo acorde sus preferencias.
- Se utilizó herramientas de Inteligencia de Negocios para la implementación de un Dashboard y brindarle al área de mercadeo una perspectiva más amplia de sus pagos Online.
- Los reportes generados de los pagos Online cumplieron los objetivos planteados, analizar los establecimientos, los días y las horas con mayor recaudación en el mes.

Recomendaciones

Como recomendaciones que serán aplicadas en el proyecto de investigación se deben de considerar las siguientes:

- Se logró implementar una propuesta al área de mercadeo para analizar cada perfil de los clientes, sin embargo, en nuevas versiones se podrá diseñar otros tipos de reportes con el fin de poder sacar provecho a la solución implementada y brindar ventaja competitiva a la entidad bancaria.
- Realizar técnicas de segmentación aplicando diferentes algoritmos de agrupamiento, para ser evaluados y poder tener una mayor precisión para sus campañas de mercadeo.
- La implementación de la propuesta se enfocó al área de Mercadeo en particular, sin embargo, puede ser una referencia para implementar nuevos procesos de la entidad bancaria.

Referencias

ARMSTRONG, Philip KOTLER y Gary, 2007, México, PEARZON Editorial.
Fundamentos de Marketing - Capitulo 7 Segmentación, mercados

Silvia del Rosario Escobar Arias, 2013,
Tesis El Marketing Digital y la Captación de Clientes de la Cooperativa de Ahorro y Crédito Universidad Técnica de Ambato Ltda.url: http://repositorio.uta.edu.ec/bitstream/123456789/5887/1/62%20MKT.pdf

Ralph Kimball, Margy Ross, 2002, New York. Editorial: WILEY
The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling (Second Edition)

Mike Biere, 2003. New Jersy, Editorial: Linda Ramagnano
Business Intelligence for the Enterprise. IBM Press. Indianápolis.

Kimball & Ross, 2002
The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling (Second Edition), New York, Wiley.

Peña, Alejandro. Instituto Politécnico Nacional. México. 2006, ISBN: 970- 94797-1-7
Inteligencia de Negocios: Una Propuesta para su Desarrollo en las organizaciones.

Peralta, Manuel. Sistema de Información. 2000.
http://www.sinnexus.com/business_intelligence http://www.sas.com

IBM, C. (2012). Manual CRISP-DM de IBM SPSS Modeler Copyright IBM Corporation.

ftp://public.dhe.ibm.com/software/analytics/spss/documentation/modeler/15.0/es/CRISP-DM.pdf

IBM SPSS, ISBN 1-56827-287-1.Copyright 2002,
Consulte el manual Data Mining with Confidence,

Calzada, Leticia y José Luis Abreu, 2012,
The impact of business intelligence tools in executive business decisions.

Josep Lluís Cano 2007,
BUSINESS INTELLIGENCE: COMPETIR CON INFORMACIÓN,
http://itemsweb.esade.edu/biblioteca/archivo/Business_Intelligence_competir_con_informacion.pdf

Talaya Agueda, 2006, 2ª Edicion. Esic, Madrid
PRINCIPIOS DE MARKETING. 2006.

Página del banco central – Referencia Normativas
https://contenido.bce.fin.ec/documentos/Administracion/ConceptosSNPV.pdf
https://www.bce.fin.ec/images/SISTEMA_N_PAGOS/GestionLiquidacion/ra_bce-gg-055-2018_SPL.pdf

Conceptos fundamentales ETL
https://es.wikipedia.org/wiki/Extract,_transform_and_load

Stephen Few. O'reilly.2006
Information Dashboard Design.

Power BI
https://docs.microsoft.com/en-us/power-bi/power-bi-overview

https://www.revistalideres.ec/lideres/banca-digital-ecuador-tecnologia-informe.html

https://www.revistalideres.ec/lideres/servicios-banca-electronica-digital-ecuador.html

Arguello, S. (2017). Cantabria: Universidad de Cantabria
La Toma de Decisiones a través del Business Intelligence.

Gartner. (11 de 02 de 2019). Stamford: Gartner Inc.
Magic Quadrant for Analytics and Business Intelligence Platforms.

Anexos

```
/*==============================================================*/
/* DBMS name:      Microsoft SQL Server 2017                    */
/* Created on:     26/6/2019 21:23:44                           */
/*==============================================================*/


if exists (select 1
   from sys.sysreferences r join sys.sysobjects o on (o.id = r.constid and o.type
= 'F')
   where r.fkeyid = object_id('HEC_PAGOS_ONLINE') and o.name =
'FK_HEC_PAGO_CLIENTES__DIM_CLIE')
alter table HEC_PAGOS_ONLINE
   drop constraint FK_HEC_PAGO_CLIENTES__DIM_CLIE
go

if exists (select 1
   from sys.sysreferences r join sys.sysobjects o on (o.id = r.constid and o.type
= 'F')
   where r.fkeyid = object_id('HEC_PAGOS_ONLINE') and o.name =
'FK_HEC_PAGO_ESTABLECI_DIM_ESTA')
alter table HEC_PAGOS_ONLINE
   drop constraint FK_HEC_PAGO_ESTABLECI_DIM_ESTA
go

if exists (select 1
   from sys.sysreferences r join sys.sysobjects o on (o.id = r.constid and o.type
= 'F')
   where r.fkeyid = object_id('HEC_PAGOS_ONLINE') and o.name =
'FK_HEC_PAGO_RELATIONS_DIM_FECH')
alter table HEC_PAGOS_ONLINE
   drop constraint FK_HEC_PAGO_RELATIONS_DIM_FECH
go

if exists (select 1
   from sys.sysreferences r join sys.sysobjects o on (o.id = r.constid and o.type
= 'F')
   where r.fkeyid = object_id('HEC_PAGOS_ONLINE') and o.name =
'FK_HEC_PAGO_RELATIONS_DIM_HORA')
alter table HEC_PAGOS_ONLINE
   drop constraint FK_HEC_PAGO_RELATIONS_DIM_HORA
go

if exists (select 1
            from  sysobjects
           where  id = object_id('DIM_CLIENTES')
            and   type = 'U')
   drop table DIM_CLIENTES
go

if exists (select 1
```

```
            from  sysobjects
           where  id = object_id('DIM_ESTABLECIMIENTOS')
            and   type = 'U')
   drop table DIM_ESTABLECIMIENTOS
go

if exists (select 1
            from  sysobjects
           where  id = object_id('DIM_FECHA')
            and   type = 'U')
   drop table DIM_FECHA
go

if exists (select 1
            from  sysobjects
           where  id = object_id('DIM_HORA')
            and   type = 'U')
   drop table DIM_HORA
go

if exists (select 1
            from  sysindexes
           where  id    = object_id('HEC_PAGOS_ONLINE')
            and   name  = 'RELATIONSHIP_4_FK'
            and   indid > 0
            and   indid < 255)
   drop index HEC_PAGOS_ONLINE.RELATIONSHIP_4_FK
go

if exists (select 1
            from  sysindexes
           where  id    = object_id('HEC_PAGOS_ONLINE')
            and   name  = 'RELATIONSHIP_3_FK'
            and   indid > 0
            and   indid < 255)
   drop index HEC_PAGOS_ONLINE.RELATIONSHIP_3_FK
go

if exists (select 1
            from  sysindexes
           where  id    = object_id('HEC_PAGOS_ONLINE')
            and   name  = 'CLIENTES_PAGOS_FK'
            and   indid > 0
            and   indid < 255)
   drop index HEC_PAGOS_ONLINE.CLIENTES_PAGOS_FK
go

if exists (select 1
            from  sysindexes
           where  id    = object_id('HEC_PAGOS_ONLINE')
            and   name  = 'ESTABLECIMIENTOS_PAGOS_FK'
            and   indid > 0
            and   indid < 255)
   drop index HEC_PAGOS_ONLINE.ESTABLECIMIENTOS_PAGOS_FK
```

```
go

if exists (select 1
            from  sysobjects
           where  id = object_id('HEC_PAGOS_ONLINE')
            and   type = 'U')
   drop table HEC_PAGOS_ONLINE
go

/*==============================================================*/
/* Table: DIM_CLIENTES                                          */
/*==============================================================*/
create table DIM_CLIENTES (
   CODIGOCLIENTE_PK     numeric              identity,
   CODIGOCLIENTE        numeric              null,
   NOMBRECLIENTE        varchar(100)         null,
   DIRECCION            varchar(100)         null,
   TELEFONO             varchar(30)          null,
   constraint PK_DIM_CLIENTES primary key (CODIGOCLIENTE_PK)
)
go

/*==============================================================*/
/* Table: DIM_ESTABLECIMIENTOS                                  */
/*==============================================================*/
create table DIM_ESTABLECIMIENTOS (
   CODIGOESTABLECIMIENTO_PK numeric              identity,
   CODIGOESTABLECIMIENTO int                    null,
   NOMBREESTABLECIMIENTO varchar(200)         null,
   constraint PK_DIM_ESTABLECIMIENTOS primary key (CODIGOESTABLECIMIENTO_PK)
)
go

/*==============================================================*/
/* Table: DIM_FECHA                                             */
/*==============================================================*/
create table DIM_FECHA (
   FECHA_PK             numeric              identity,
   FECHA                datetime             null,
   ANIO                 int                  null,
   MES                  int                  null,
   SEMANA               int                  null,
   DIA                  int                  null,
   constraint PK_DIM_FECHA primary key (FECHA_PK)
)
go

/*==============================================================*/
/* Table: DIM_HORA                                              */
/*==============================================================*/
create table DIM_HORA (
   HORA_PK              numeric              identity,
   CODIGOHORA           int                  null,
   DESCRIPCIONHORA      varchar(50)          null,
```

```
   constraint PK_DIM_HORA primary key (HORA_PK)
)
go

/*==============================================================*/
/* Table: HEC_PAGOS_ONLINE                                      */
/*==============================================================*/
create table HEC_PAGOS_ONLINE (
   CODIGOESTABLECIMIENTO_PK int                 null,
   CODIGOCLIENTE_PK     int                  null,
   FECHA_PK             int                  null,
   HORA_PK              int                  null,
   IDREGISTRO           numeric(18)          identity,
   CANTIDADPAGOS        numeric              null,
   TOTALPAGOS           decimal(18,2)        null,
   FECHACARGA           datetime             null
)
go

/*==============================================================*/
/* Index: ESTABLECIMIENTOS_PAGOS_FK                             */
/*==============================================================*/



create nonclustered index ESTABLECIMIENTOS_PAGOS_FK on HEC_PAGOS_ONLINE
(CODIGOESTABLECIMIENTO_PK ASC)
go

/*==============================================================*/
/* Index: CLIENTES_PAGOS_FK                                     */
/*==============================================================*/



create nonclustered index CLIENTES_PAGOS_FK on HEC_PAGOS_ONLINE (CODIGOCLIENTE_PK
ASC)
go

/*==============================================================*/
/* Index: RELATIONSHIP_3_FK                                     */
/*==============================================================*/



create nonclustered index RELATIONSHIP_3_FK on HEC_PAGOS_ONLINE (FECHA_PK ASC)
go

/*==============================================================*/
/* Index: RELATIONSHIP_4_FK                                     */
/*==============================================================*/
```

```
create nonclustered index RELATIONSHIP_4_FK on HEC_PAGOS_ONLINE (HORA_PK ASC)
go

alter table HEC_PAGOS_ONLINE
   add constraint FK_HEC_PAGO_CLIENTES__DIM_CLIE foreign key (CODIGOCLIENTE_PK)
      references DIM_CLIENTES (CODIGOCLIENTE_PK)
go

alter table HEC_PAGOS_ONLINE
   add constraint FK_HEC_PAGO_ESTABLECI_DIM_ESTA foreign key
(CODIGOESTABLECIMIENTO_PK)
      references DIM_ESTABLECIMIENTOS (CODIGOESTABLECIMIENTO_PK)
go

alter table HEC_PAGOS_ONLINE
   add constraint FK_HEC_PAGO_RELATIONS_DIM_FECH foreign key (FECHA_PK)
      references DIM_FECHA (FECHA_PK)
go

alter table HEC_PAGOS_ONLINE
   add constraint FK_HEC_PAGO_RELATIONS_DIM_HORA foreign key (HORA_PK)
      references DIM_HORA (HORA_PK)
go
```