

UNIVERSIDAD NACIONAL TECNOLÓGICA DE LIMA SUR

FACULTAD DE INGENIERÍA Y GESTIÓN

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



**“PROPUESTA DE UN MODELO DE DATOS EN BIG DATA
PARA LA ESTANDARIZACIÓN DEL DATA LAKE PARA UNA
ENTIDAD FINANCIERA DEL PERÚ.”**

TRABAJO DE SUFICIENCIA PROFESIONAL

Para optar el Título Profesional de

INGENIERO DE SISTEMAS

PRESENTADO POR EL BACHILLER

MARTINEZ CAYHUALLA, RENZO OMAR

ASESOR

OCHOA CARBAJAL, HERNAN

Villa El Salvador

2021

DEDICATORIA

Primeramente, a Dios porque a él le debo todo lo que tengo y todo lo que soy. A mis padres, Justina y Julio, y a mis hermanas, por sus ejemplos y apoyo incondicional para seguir superándome día a día, ellos son un pilar fundamental para seguir adelante

AGRADECIMIENTO

- A mi alma máter, Universidad Nacional Tecnológica de Lima Sur (UNTELS), quien me abrió las puertas para desarrollarme profesionalmente.
- Al Ing. Hernán Ochoa Carbajal, mi asesor, por su guía y apoyo en brindarme las pautas necesarias para la elaboración de este trabajo de investigación.
- A mis docentes universitarios, quienes han contribuido a través de sus conocimientos en mi formación profesional.
- A mi Enamorada, quien me motivo a no darme por vencido y sacar adelante este proyecto.
- A todas las personas que de alguna u otra manera han o están en mi vida, de los cuales guardo sus consejos, enseñanzas y ánimos.

ÍNDICE

RESUMEN	VIII
INTRODUCCIÓN.....	X
CAPÍTULO I. ASPECTOS GENERALES	1
1.1 CONTEXTO	1
1.2 DELIMITACIÓN TEMPORAL Y ESPACIAL DEL TRABAJO	2
1.2.1 <i>Temporal</i>	2
1.2.2 <i>Espacial</i>	2
1.3 OBJETIVOS	2
1.3.1 <i>Objetivo General</i>	2
1.3.2 <i>Objetivos Específicos</i>	2
1.4 ALCANCE Y LIMITACIONES	3
1.4.1 <i>Alcance</i>	3
1.4.2 <i>Limitaciones</i>	3
CAPÍTULO II. MARCO TEÓRICO	5
2.1 ANTECEDENTES	5
2.1.1 <i>Antecedentes Internacionales</i>	5
2.1.2 <i>Antecedentes Nacionales</i>	6
2.2 BASES TEÓRICAS	7
2.2.1 <i>Modelo de Datos en Big Data</i>	7
2.2.2 <i>Base de datos para el Modelado de datos y sus requisitos</i>	8
2.2.3 <i>Fundamentos del Modelado de Datos</i>	8
2.2.4 <i>Modelo Conceptual</i>	9
2.2.5 <i>Modelo Lógico</i>	13
2.2.6 <i>Modelo Físico</i>	22
2.3 DEFINICIÓN DE TÉRMINOS BÁSICOS	24
CAPÍTULO III: DESARROLLO DEL TRABAJO PROFESIONAL	25
3.1 DETERMINACIÓN Y ANÁLISIS DEL PROBLEMA.....	25
3.2 MODELO DE LA SOLUCIÓN PROPUESTA	26
3.2.1 <i>Requisitos</i>	27
3.2.2 <i>Diagrama ER</i>	28
3.2.3 <i>Normalización</i>	35
3.2.4 <i>Modelo Lógico</i>	40
3.3 RESULTADOS.....	41
3.3.1 <i>Modelo Físico</i>	41
CONCLUSIONES	48
RECOMENDACIONES	49
REFERENCIAS	50
ANEXOS	52

LISTA DE FIGURAS

Figura 1. Arquitectura Referencial del Data Lake	4
Figura 2. Entidad Regular	11
Figura 3. Entidad Débil	11
Figura 4. Relación entre Entidades.....	11
Figura 5. Cardinalidad.....	12
Figura 6. Identificador	13
Figura 7. Ejemplo de clave externa.....	19
Figura 8. Modelo Lógico	21
Figura 9. Modelo Físico	23
Figura 10. Archivos del Aplicativo Móvil.....	27
Figura 11. Archivo de Afiliados	27
Figura 12. Archivo de Transacciones de la Moneda Digital	28
Figura 13. Modelo Entidad Relación.....	34
Figura 14. Modelo Lógico	41
Figura 15. Barra de herramientas del erwin.....	42
Figura 16. Barra de Herramientas Opción Tools	42
Figura 17. Ventana para la Configuración de Atributos	43
Figura 18. Creación de Etiqueta en la Ventana para la Configuración de Atributos	44
Figura 19. Estandarización de Atributos	45
Figura 20. Modelo Físico de Datos	46
Figura 21. Modelo Físico Final	47

LISTA DE TABLAS

Tabla 1. Tabla sin Normalizar	15
Tabla 2. Tabla con la 1ª Normalización	15
Tabla 3. Clave Compuesta	17
Tabla 4. Tabla de Usuarios.....	17
Tabla 5. Tabla de películas.....	18
Tabla 6. Tabla Normalizada de Clientes.....	20
Tabla 7. Tabla de Alquileres de Video	20
Tabla 8. Tabla de Saludos.....	21
Tabla 9. Descripción de Atributos de los 2 Archivos Planos.....	29
Tabla 10. Entidades Existentes en la Capa Processing.....	30
Tabla 11. Entidades en la Capa Processing del Data Lake.....	31
Tabla 12. Tabla Cliente	35
Tabla 13. Tabla AfiliadoAplicativo.....	36
Tabla 14. CuentaFinanciera	36
Tabla 15. Tabla Transacciones.....	37
Tabla 16. Tabla Estado	38
Tabla 17. Tipo Transaccion	38
Tabla 18. Tabla Transacciones Normalizada.....	39

LISTA DE ANEXOS

Anexo A. Cantidad de Suscriptores a la Aplicación Móvil de la Moneda Digital	52
Anexo B. Representación Gráfica de los 2 Tipos de Clientes	52
Anexo C. Gráfico de Cantidad de Usuarios Nuevos en promedio por mes en el año 2020.	53
Anexo D. <i>Gráfico del Promedio de Transacciones al Mes en el Año 2020</i>	53
Anexo E. Gráfico del Volumen de Dinero en Promedio en el Año	54

RESUMEN

Los datos permiten extraer información útil para las empresas apoyándose de diferentes modelos predictivos basados en Machine Learning para la mejor toma de decisiones. Actualmente tenemos que los datos van en aumento cada día, los diferentes tipos que se requiere almacenar, el tiempo en extraer, procesar y analizar los datos con las diferentes herramientas tradicionales que utilizan las entidades financieras, no están preparados para las exigencias actuales. Por ende, las empresas sean de cualquier rubro optan por migrar sus datos e información al Data Lake basándose en Big Data. A fin de que el siguiente trabajo de Investigación se realizó con la finalidad de proponer un modelo de datos para la carga en la capa Data Processing del Data Lake basándose en los principios del modelado de datos y apoyándose de la herramienta Erwin. Buscando de tal modo tener los datos estandarizados, para su futura explotación por parte de la entidad financiera.

Palabras clave: Machine Learning, Big Data, Data Lake.

Abstract

Data allows to extract useful information for companies supported by different predictive models based on Machine Learning for better decision making. Currently we have that data are increasing every day, the different types that are required to store, the time to extract, process and analyze data with the different traditional tools used by financial institutions, are not prepared for the current demands. As a result, companies in any industry are opting to migrate their data and information to the Data Lake based on Big Data. Therefore, the following research work was carried out with the purpose of proposing a data model for the Data Processing layer of the Data Lake based on the principles of data modeling and supported by the Erwin tool. Seeking to have standardized data for its future exploitation by the financial institution.

Keywords: Machine learning, Big Data, Data Lake.

INTRODUCCIÓN

Con el transcurso de los años y debido a los avances tecnológicos en el sector bancario, así como la creciente competencia, se vieron en la necesidad de implementar el Data Lake basándose en Big Data para poder manejar grandes volúmenes de datos y anticiparnos al comportamiento del cliente con la ayuda de modelos predictivos que influyan en la toma de nuevos caminos y estrategias.

El trabajo de investigación tiene como principal propósito proponer un modelo de datos, generados por la Aplicación Móvil de la Moneda Digital, que fueron migrados al Data Lake. Como primer paso, identificar los diferentes Data Source que fueron almacenados en la capa Data Ingestion del Data Lake. Para luego ingestarlos a la siguiente capa, Data Processing. De este modo se buscará tener los datos estandarizados en el Data Lake. Para posteriormente ser explotada por parte de la entidad financiera.

Durante el desarrollo del presente proyecto de investigación, se tomará en cuenta los siguientes capítulos:

En el capítulo II, se elaborará el Marco Teórico, en el que se plantea las bases teóricas relacionadas al modelado de datos en la capa Data Processing del Data Lake. Bajo los principios del Modelo de datos. El cual nos ayudara en tener una base para el desarrollo de la investigación.

En el Capítulo III, comprende el desarrollo del proyecto de investigación, la metodología y herramienta utilizada, así como la consolidación en la solución del problema y los resultados esperados.

CAPÍTULO I. ASPECTOS GENERALES

1.1 Contexto

El contexto en donde se realizará el trabajo de investigación es en una empresa peruana del sector financiero. Al ser una empresa que cuenta con los datos personales de sus clientes y transacciones monetarias deben ser custodiadas.

A raíz de la pandemia por el Coronavirus (COVID-19) y a mediados de marzo del 2020 cuando se inició la cuarentena. Debido a esto las suscripciones a la Aplicación Móvil de la Moneda Digital empezó a crecer más cada año (**Anexo A**) en los clientes bancarizados y los no bancarizados que representan alrededor del 75% del total del mercado (**Anexo B**) desde su lanzamiento en el año 2016. Lo que llevo a que el volumen de datos creciera considerablemente y que el Data Warehouse como modelo tradicional de almacenamiento de datos no estaba preparado para las exigencias actuales. Por lo que se migro los datos de la Aplicación Móvil de la Moneda Digital al Data Lake. Lo que conlleva a realizar este trabajo de investigación, proponer un modelo de datos en la capa Data Processing del Data Lake para su explotación final por parte de la entidad financiera.

1.2 Delimitación Temporal y Espacial del Trabajo

1.2.1 Temporal

- Fecha Inicio: Enero del 2021.
- Fecha Término: Diciembre del 2021.

1.2.2 Espacial

El presente trabajo de investigación está limitado a una empresa del sector financiero. Pero puede ser aplicado en otras entidades del mismo sector.

1.3 Objetivos

1.3.1 Objetivo General

- Proponer un Modelo de Datos en el Data Lake con el fin de lograr un estándar en la capa Data Processing.

1.3.2 Objetivos Específicos

- Identificar los datos relacionados con la Aplicación Móvil de la Moneda Digital, su fuente de origen dentro de la empresa.
- Consolidar y validar los datos de la Moneda Digital en un único ambiente estandarizado que permita su consulta y/o explotación por parte de cualquier miembro de la entidad financiera.
- Determinar la influencia y aportes que genera el Data Lake con respecto a los datos estandarizados que estarán almacenados en la Capa Processing.

1.4 Alcance y limitaciones

1.4.1 Alcance

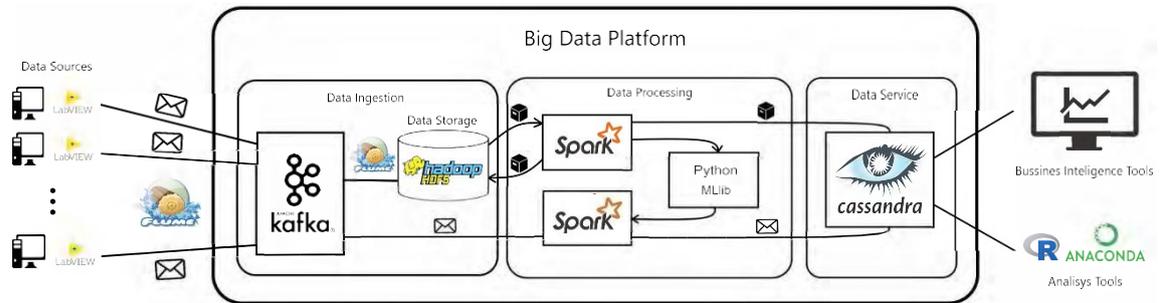
- El alcance del proyecto abarcará el modelamiento de datos entre las capas del Data Lake, Fuente: Capa Ingestion y Destino: Capa Processing.
- Los datos que se modelarán serán los generados por la Aplicación Móvil de la Moneda Digital.

1.4.2 Limitaciones

Una de las limitaciones que presenta este trabajo de investigación será de carácter bibliográfico ya que debido al contexto de esta pandemia la gran mayoría de bibliotecas se encuentran cerradas, por lo que no podemos acercarnos a consultar ningún libro, no obstante, la base de esta investigación será construida acudiendo a diversos repositorios y base de datos que están en la web.

Además, por políticas de privacidad de datos como: La arquitectura actual, las diferentes herramientas y/o aplicaciones que utiliza la entidad financiera. Esta en la obligación de utilizar una arquitectura Data Lake como referencia para el desarrollo del trabajo de investigación. Adicional el modelo conceptual, lógico y físico que se implementará será uno referencial.

Figura 1. *Arquitectura Referencial del Data Lake*



Nota: La Figura 1 representa una arquitectura Data Lake que será tomada como referencia. Adaptado de "A Big Data Architecture for Fault Prognostics of Electronic Devices: Application to Power MOSFETs" (p. 102163), por Alonso C. , Pulido, Cartón, & Bregon, 2019, *IEEE Access*, 7

CAPÍTULO II. MARCO TEÓRICO

2.1 Antecedentes

2.1.1 Antecedentes Internacionales

- Rodríguez (Colombia, 2020), en su tesis *“Metadatos Estadísticos para el aseguramiento de la Calidad de los Datos”* quien propone un modelo de metadatos estadísticos en base a la ISO/IEC 11179 para el aseguramiento de la calidad de datos. Esto nace como una necesidad debido a que las organizaciones producen y comparten Datasets de diferentes estructuras. Rodríguez llega a la conclusión que se debe tener un modelo genérico que compartan una visión y conocimiento común del Dataset, sea externo o interno en la organización.
- Soche (Colombia, 2016), en su tesis *“Metodología para el Modelamiento de Datos basado en Big Data, enfocados al Consumo de Tráfico (voz-datos) generados por los Clientes”* quien propone el modelamiento de datos basado en big data, como una tecnología que ayuda a las organizaciones a tomar mejores decisiones y a cuidar a sus clientes, esto se realizará a través del análisis del consumo (Voz-Datos) que estos generan al llamar por su teléfono móvil o por el consumo de internet móvil. Soche llega a la conclusión que el software Hadoop como herramienta útil al modelamiento de datos basándose en Big Data, le permitirá realizar el almacenamiento y procesamiento de datos. Esta herramienta logrará tener la información en línea para utilizarla en cualquier momento por parte de la entidad del sector de telecomunicaciones.

- Lorenti (Ecuador, 2019), en su tesis *“El Big Data aplicado en la Redes Sociales para una Empresa de Telefonía Móvil Celular en Ecuador”* quien propone aplicar la tecnología Big Data en las redes sociales y sus beneficios en el país, y planteó como objetivo general, establecer la influencia en sus suscriptores. Como técnica de recolección de datos se utilizará una encuesta de base estructurada y como muestra el número de suscriptores activos de la operadora de telefonía móvil. Lorenti espera obtener en base a los resultados de su investigación poder anticiparse a las necesidades de sus clientes y satisfacerlas eficazmente, prevenir que la competencia gane mercado por no escucharlos.

2.1.2 Antecedentes Nacionales

- Almora (Ica, 2018), en su tesis *“Análisis y usos del Big Data aplicado en la Universidad Nacional San Luis Gonzaga de Ica: caso facultad de ingeniería de sistemas”* en base a su investigación quiere entender con mayor amplitud el Big Data como una nueva tecnología para grandes volúmenes de datos, que herramientas están relacionadas con él, así como el alcance de los datos para que sean aprovechados correctamente en los estudiantes de la facultad de Ingeniería De Sistemas. Nellie, en base al estudio realizado a los estudiantes, son conscientes que está en tendencia el aprovechamiento de la gran cantidad de datos que se están produciendo en el mundo y en las organizaciones el cual abre un nuevo camino para aprovecharlo profesionalmente.
- Linares (Lima, 2019), en su tesis *“Implementación de un Sistema de Big Data aplicado a la Migración de Datos bajo la distribución Cloudera apache Hadoop, en el Banco Interbank”* quien desea implementar un sistema de migración de datos que permita la optimización de los tiempos en el procesamiento de datos. También, gestionar el crecimiento exponencial de datos con el fin de garantizar la

escalabilidad y alta disponibilidad del sistema. Carlos, llega a la conclusión que la ingesta de datos con Apache Sqoop a Hadoop, se pudo centralizar todos los datos provenientes de la fuente de datos ORACLE depositados en el Data Lake del banco, lo cual permitirá no solo trabajar con datos estructurados si no también semiestructurados y no estructurada con el fin de potenciar la toma de decisiones de la compañía.

- Huapaya (Lima , 2021), en su tesis “*Integración de entorno Big Data en la gestión financiera de un Sistema Bancario para optimizar productos, servicios internos y toma de decisiones*” quien propone integrar los datos en un entorno Big Data y asegurar la accesibilidad de los usuarios del negocio. El entorno Big Data comprende la capa de ingesta de datos en un Data Lake mediante el procesamiento distribuido en spark; y otra capa de consumo de información a través de sandboxes, donde los usuarios podrán realizar analítica avanzada y reportería con herramientas de inteligencia de negocios. Huapaya, llega a la conclusión que el Data Lake gobernado nos brindan mucha volumetría para tomar decisiones, además la información es veraz por el gobierno del dato implantado, tiene el consenso de diferentes áreas del negocio por el valor y variedad predominan en el Data Lake.

2.2 Bases teóricas

2.2.1 Modelo de Datos en Big Data

Silberschatz, Korth y Sudarshan (2002) definen el modelado de datos como una manera de tener organizados y estructurados los datos, de tal modo se puedan manipular en las bases de datos. Actualmente, las empresas manejan una gran cantidad de datos o también llamados Big Data con el propósito de brindar a sus clientes excelentes ideas de comercio.

Además, el CEUPE señala que existen sistemas de base de datos relacionales, que se han manejado durante 30 años y que actualmente no cubren las expectativas de las empresas. Por consiguiente, se ha dado a conocer varias opciones de base de datos para procesar datos masivos y superar las debilidades de las bases de datos tradicionales y ese es el motivo por el cual surgen las bases de datos Big Data.

2.2.2 Base de datos para el Modelado de datos y sus requisitos

Debido a la necesidad de implementar base de datos que puedan procesar grandes cantidades de datos. Teniendo en cuenta las principales características del big data como: volumen, variedad y velocidad.

- a. Volumen.** Según Lopez (2012), podemos llamarlo como Big Data. Debido a que las empresas almacenan un mayor de tipos de datos como terabytes, inclusive tera bytes de información.
- b. Variedad.** Por otro lado, Ivanov, Korfiatis y Zicari (2013) señalan que debido a que tenemos múltiples fuentes de datos, almacenamos diferentes tipos de datos como: estructurados, semi estructurados y no estructurados.
- c. Velocidad.** Almora (2018) define que en diversas aplicaciones basándose en Big Data, los datos se generan de manera continua y se procesan en tiempo real llamado también “Streaming Data”.

2.2.3 Fundamentos del Modelado de Datos

Por otro lado, CEUPE (2021), señala que el profesional a desarrollar el Modelado de Datos debe conocer a profundidad los fundamentos, para poder desempeñarse de una manera eficaz con los responsables del proyecto.

a. **Tareas aplicadas en el Modelado de datos.** Además, CEUPE (2021) señala las tareas que debe seguir para el Modelado de Datos se detallaran a continuación:

- Reconocer las entidades.
- Determinar los atributos de cada entidad.
- Asignar los estándares de nomenclatura.
- Reconocer relaciones entre entidades.
- Fijar claves primarias y foráneas que existe en las entidades.
- Normalizar para evitar la reiteración de los Datos.

b. **Tipos de modelados de datos.** CEUPE (2021) define a continuación los tipos de Modelado de Datos:

- Modelo de Datos conceptual: Constatar las relaciones que existen entre las entidades así mismo detallar los atributos de cada una de ellas.
- Modelo de Datos lógico: Se emplea como base para el desarrollo del Modelo de Datos Físico.
- Modelo de Datos Físico: Es el modelo resultado final del Modelo de Datos Lógico.

En conclusión, durante las primeras fases del proyecto se debe definir el Modelado de datos para validar los requerimientos y lograr integrarlo con los Modelo de datos existentes (Centro Europeo de Postgrado, 2021).

2.2.4 Modelo Conceptual

Se emplea principalmente durante la etapa de Análisis del problema debido a que están orientados a interpretar los elementos que influyen en el problema y sus relaciones. El ejemplo más tradicional es el modelo Entidad Relación, se basa en la representación del mundo real mediante los elementos básicos llamados entidades y la relación que existen entre ellos (SDi, 2017).

Modelo Entidad Relación.

Es un modelo que nos ayuda a establecer esquemas conceptuales de bases de datos. Es decir, es un estándar que nos ayuda en establecer el modelo mencionado.

Se le nombra modelo E/R también como EI (Entidad / Interrelación). Sus siglas tradicionales son las E/R que nos ayuda entender tanto en el idioma del inglés y el español.

Estableció la noción de entidad, atributos y relación (Sánchez, 2004, pág. 17).

- Entidad

Es la representación de un elemento (abstracto o real) u objeto el cual nos ayuda con el almacenamiento de los datos en las bases de datos. Ejemplo de entidades son Renzo, el número de cuenta 198768955757, Numero de Tarjeta 455778945123.

Las entidades se representan mediante un rectángulo el cual nombre va dentro de ello. A continuación, se mostrará los 2 tipos de entidades (Sánchez, 2004, pp. 17-18).

- Regulares: Son entidades normales cuya existencia no depende de otras si no por sí mismas. Su representación es tal como se ve en la Figura 2:

Figura 2. Entidad Regular



Nota: Representación de la Entidad Regular para el diseño de Base de Datos. Adaptado de *Diseño Conceptual de Base de datos* (p. 18), por J. Sánchez, 2004, Creative Commons

- **Débiles:** Para este caso la entidad depende de otras, como por ejemplo la entidad Tarjeta de Credito depende de una Cuenta Financiera. Su representación es tal como se ve en la Figura 3.

Figura 3. Entidad Débil



Nota: Representación de la Entidad Débil para el diseño de Base de Datos. Adaptado de *Diseño Conceptual de Base de datos* (p. 18), por J. Sánchez, 2004, Creative Commons

- **Relación.**

Según Sánchez (2004, pág. 18) es la representación de las relaciones entre las entidades. Su grafica es tal como se muestra en la Figura 4, se une mediante líneas con las entidades y el nombre de las entidades es mediante el uso del verbo.

Figura 4. Relación entre Entidades



Nota: Representación de la relación entre 2 entidades. Adaptado de *Diseño Conceptual de Base de datos* (p. 19), por J. Sánchez, 2004, Creative Commons

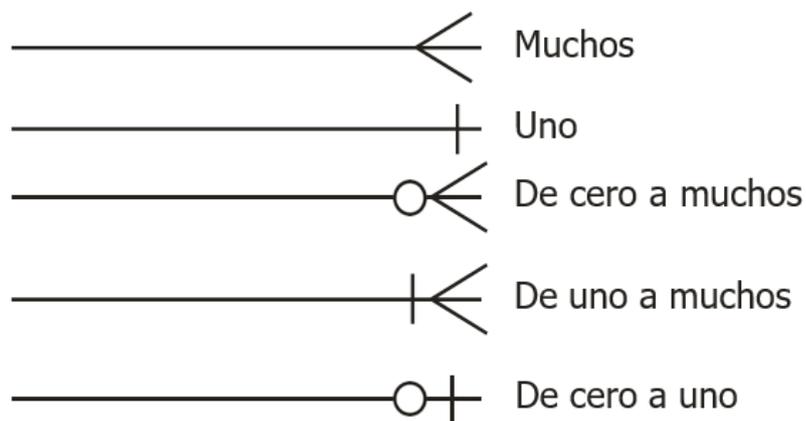
- **Cardinalidad.**

Se denota el número de relaciones que una entidad pueda tener. La cual a continuación se indicará los 2 tipo de Cardinalidad que existe:

- Cardinalidad mínima. Muestra el número mínimo de asociaciones que tendrá la entidad. El valor que puede obtener es cero o uno.
- Cardinalidad máxima. Muestra el número máximo de asociaciones que tendrá la entidad. El valor que puede obtener es uno o muchos.

En los esquemas entidad / relación podemos representar la cardinalidad de diferentes formas. A continuación, la más usada es tal como se muestra en la Figura 5.

Figura 5. *Cardinalidad*



Nota: Representación de la cardinalidad que existen entre las entidades. Adaptado de *Diseño Conceptual de Base de datos* (p. 20), por J. Sánchez, 2004, Creative Commons

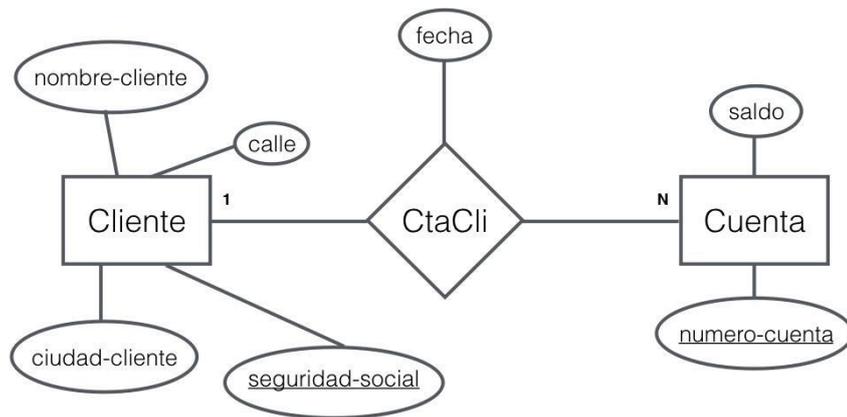
- **Atributo.**

Es la representación de las propiedades de las entidades. Se representa mediante un círculo, el cual va el nombre del atributo. Ejemplo: Tenemos la entidad Cliente por lo que sus atributos serán Nombres del cliente, numero de documento, número de tarjeta, número de cuenta, fecha afiliación, fecha bloqueo, etc.

- **Identificador.**

Indica si uno o más campos pueden tomar valores únicos en una entidad. Se representa subrayado el nombre del atributo. Tal como se muestra en la Figura 6 (Sánchez, 2004, pp. 17-21).

Figura 6. *Identificador*



Nota: Adaptado de EL MODELO ENTIDAD-RELACION Y EL MODELO RELACIONAL EN BASES DE DATOS[Fotografía], por D. Giordana, 2017. (<https://bandageek.com/2017/02/el-modelo-entidad-relacion-y-el-modelo-relacional-en-bases-de-datos/>). CC BY 2.0

2.2.5 Modelo Lógico

Este modelo es estructurado en comparación al Modelo de Datos Conceptual. El modelo determina la estructura y las relaciones que existen entre las entidades. La ventaja del modelo lógico es que nos facilita una base para el modelo de datos físico.

A continuación, se detallan las características del Modelo Lógico (Fatima, 2019):

- El modelo lógico detalla los requerimientos de los proyectos. Así mismo se puede unificar a otros Modelos de Datos según tenga definido su alcance.
- Su estructura y elaboración es independiente de los Sistemas de Gestión de Base de Datos.
- Se establece una longitud precisa para los diferentes tipos de datos.

- **Normalización**

Según Guru99 (2021) define a la normalización de datos como un proceso de diseño que estructura las tablas con el fin de evitar redundancia de los datos.

Existe una controversia si hay más de la 6° Forma normal. Pese a ello la mayoría de los Modelo de Datos, llegan a su mejor rendimiento a la 3° Forma Normal. Para el desarrollo del presente trabajo de investigación se realizará la normalización hasta el nivel 3^a.

a. **1ª Normalización.** Además, Guru99 (2021) nos explica la 1° Normalización en base al siguiente ejemplo. El cual se tiene una tabla de una biblioteca de videos el cual se alquilas películas. Sin realizar ninguna normalización, todos los datos se almacenan en una sola tabla tal como se muestra en la Tabla 1.

Tabla 1. *Tabla sin Normalizar*

Full Names	Physical Address	Movies rented	Salutation	Category
Janet Jones	First Street Plot No 4	Pirates of the Caribbean, Clash of the Titans	Ms.	Action, Action
Robert Phil	3 rd Street 34	Forgetting Sarah Marshal, Daddy's Little Girls	Mr.	Romance, Romance
Robert Phil	5 th Avenue	Clash of the Titans	Mr.	Action

Nota: La Tabla 1 es un ejemplo de una tabla de una biblioteca de videos. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía], por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

Siguiendo, pasaremos a la 1º Normalización teniendo en cuenta lo siguiente a continuación.

- Cada casilla de la tabla debe tener un solo valor.
- Cada registro debe ser distinto.

Tabla 2. *Tabla con la 1ª Normalización*

FULL NAMES	PHYSICAL ADDRESS	MOVIES RENTED	SALUTATION
Janet Jones	First Street Plot No 4	Pirates of the Caribbean	Ms.
Janet Jones	First Street Plot No 4	Clash of the Titans	Ms.
Robert Phil	3 rd Street 34	Forgetting Sarah Marshal	Mr.
Robert Phil	3 rd Street 34	Daddy's Little Girls	Mr.
Robert Phil	5 th Avenue	Clash of the Titans	Mr.

Nota: La Tabla 2 es una tabla aplicando la 1ª Normalización del ejemplo que se tomó de la figura 9. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía], por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

Antes de seguir con la siguiente normalización, tengamos en cuenta las siguientes definiciones:

b. Clave Primaria. Guru99 (2021) afirma que es un valor único en la columna de una tabla de Base de Datos. El cual nos ayuda a identificar los registros de manera única.

Tiene los siguientes atributos:

- Una clave primaria no puede valor NULL.
- El valor del dato de la llave primaria en la tabla debe ser único.
- Los datos de la clave primaria no se pueden editar.
- La columna que sea la clave primaria deben contener un valor cuando se ingeste un nuevo registro.

c. Clave Compuesta. La clave compuesta son las claves primarias de varias columnas utilizadas para reconocer un registro único en una tabla. En el ejemplo de la Tabla 3 tenemos a 2 personas con el mismo nombre, pero con diferentes direcciones.

Tabla 3. Clave Compuesta

Composite Key			
Robert Phil	3 rd Street 34	Daddy's Little Girls	Mr.
Robert Phil	5 th Avenue	Clash of the Titans	Mr.

Nota: La Tabla 3 es un ejemplo de una clave compuesta en una tabla. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía], por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

Por consiguiente, necesitamos el nombre completo y la dirección como la llave compuesta.

Seguiremos ahora con forma normal 2NF.

d. 2ª Normalización. Tal como se ve, no se puede avanzar con la pequeña base de datos de ejemplo. Debemos seguir con las siguientes reglas presentadas a continuación para conseguir la 2º NF.

- Regla 1- Estar en 1NF.
- Regla 2- Clave principal debe estar en una sola columna de la tabla.

Tabla 4. Tabla de Usuarios

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION
1	Janet Jones	First Street Plot No 4	Ms.
2	Robert Phil	3 rd Street 34	Mr.
3	Robert Phil	5 th Avenue	Mr.

Nota: La Tabla 4 es de los usuarios de la biblioteca de videos. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía],

por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

Tabla 5. *Tabla de películas*

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

Nota: La Tabla 5 es de las películas alquiladas por los usuarios. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía], por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

Hemos dividido nuestra tabla 1NF en dos tablas. Tabla 4 y Tabla 5. Hemos introducido una nueva columna llamada Membership_id, que es la clave principal para la Tabla 4. Los registros se pueden identificar de forma única en la Tabla 4 con el ID de membresía.

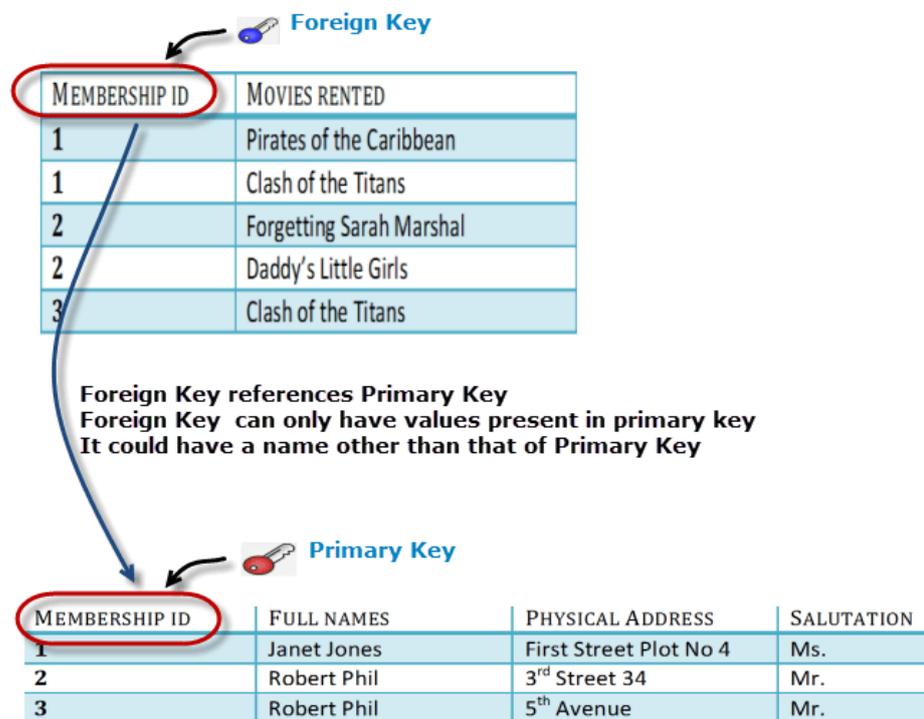
Antes de proceder, comprendamos la siguiente definición para seguir con la 3ª Normalización.

e. Clave Externa. La clave externa hace alusión a la clave principal de otra tabla. Como referencia tenemos la Tabla 5 en donde Membership_ID es la clave externa. A continuación, se describirán sus características.

- La clave externa puede tener diferente nombre que la clave primaria de la otra tabla.
- Asegura que los registros de una tabla tengan relación con la otra tabla.

- A diferencia de la clave primaria, no tienen que ser necesariamente únicos. La mayoría de las veces puede tener muchas claves externas.
- La Clave Externa puede tener datos nulos, aunque las claves primarias no tal como se indicó anteriormente.

Figura 7. Ejemplo de clave externa



Nota: En la Figura 7 se muestra la relación que existen entre las Tabla 4 de películas y Tabla 5 de Usuarios. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía], por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

f. **3ª Normalización.** Para poder pasar de la 2º NF a la 3º NF de nuestra tabla de ejemplo. Procederemos a dividir nuestra Tabla 4 de ejemplo. La cual debe cumplir con las siguientes reglas.

- Regla 1- La tabla debe estar en 2NF.

- Regla 2- No se deben estar los conceptos de 2 entidades en una misma tabla. Se pueden relacionar mediante Claves Externas.

Para continuar con la 3ª FN, se procede a dividir las Tabla 4 de la siguiente manera.

Tabla 6. *Tabla Normalizada de Clientes*

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION ID
1	Janet Jones	First Street Plot No 4	2
2	Robert Phil	3 rd Street 34	1
3	Robert Phil	5 th Avenue	1

Nota: La Tabla 6 muestra los clientes de la tienda de videos. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía], por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

Tabla 7. *Tabla de Alquileres de Video*

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

Nota: En la Tabla 7 muestra los clientes que alquilan videos. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía], por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

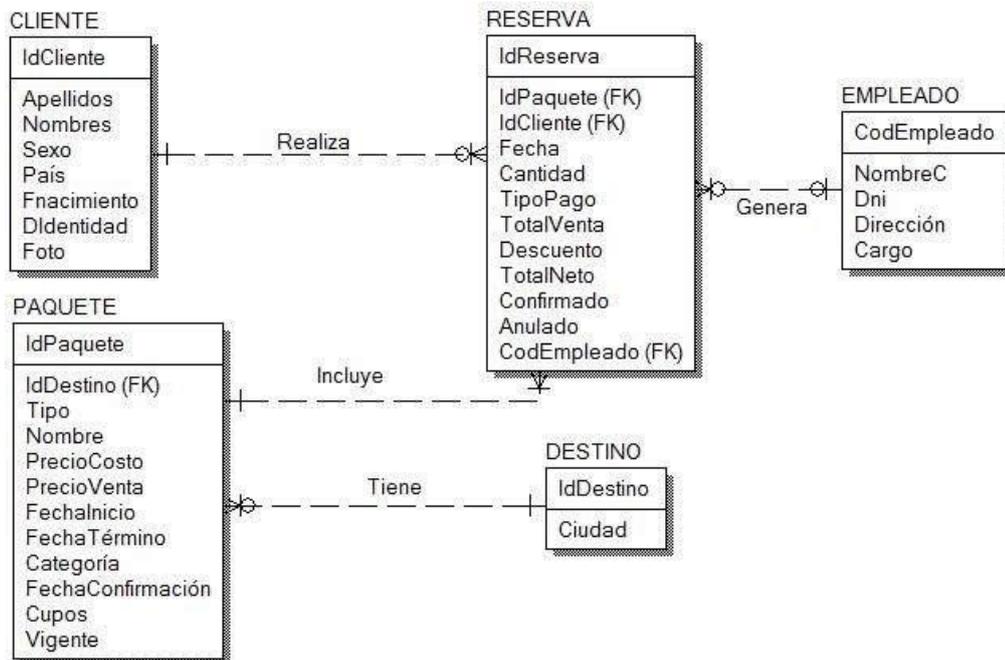
Tabla 8. *Tabla de Saludos*

SALUTATION ID	SALUTATION
1	Mr.
2	Ms.
3	Mrs.
4	Dr.

Nota: Se creo la Tabla 8 para tener la Tabla 6 normalizada. Adaptado de ¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos [Fotografía], por guru99es, 2018, guru99 (<https://guru99.es/database-normalization/>). CC BY 2.0

Finalmente se detalla el Modelo de datos Lógico luego de tomar como base el modelo conceptual y haber seguido las reglas de la normalización.

Figura 8. *Modelo Lógico*



Nota: En la Figura 8 muestra un modelo lógico de una Agencia de Viajes que se utiliza como ejemplo. Adaptado de “Integración de procesos de negocio aplicando la arquitectura orientada a servicios (SOA)” (p. 113), por José Marleny, Marley Peralta, 2017, *Interfases*, 10.

2.2.6 Modelo Físico

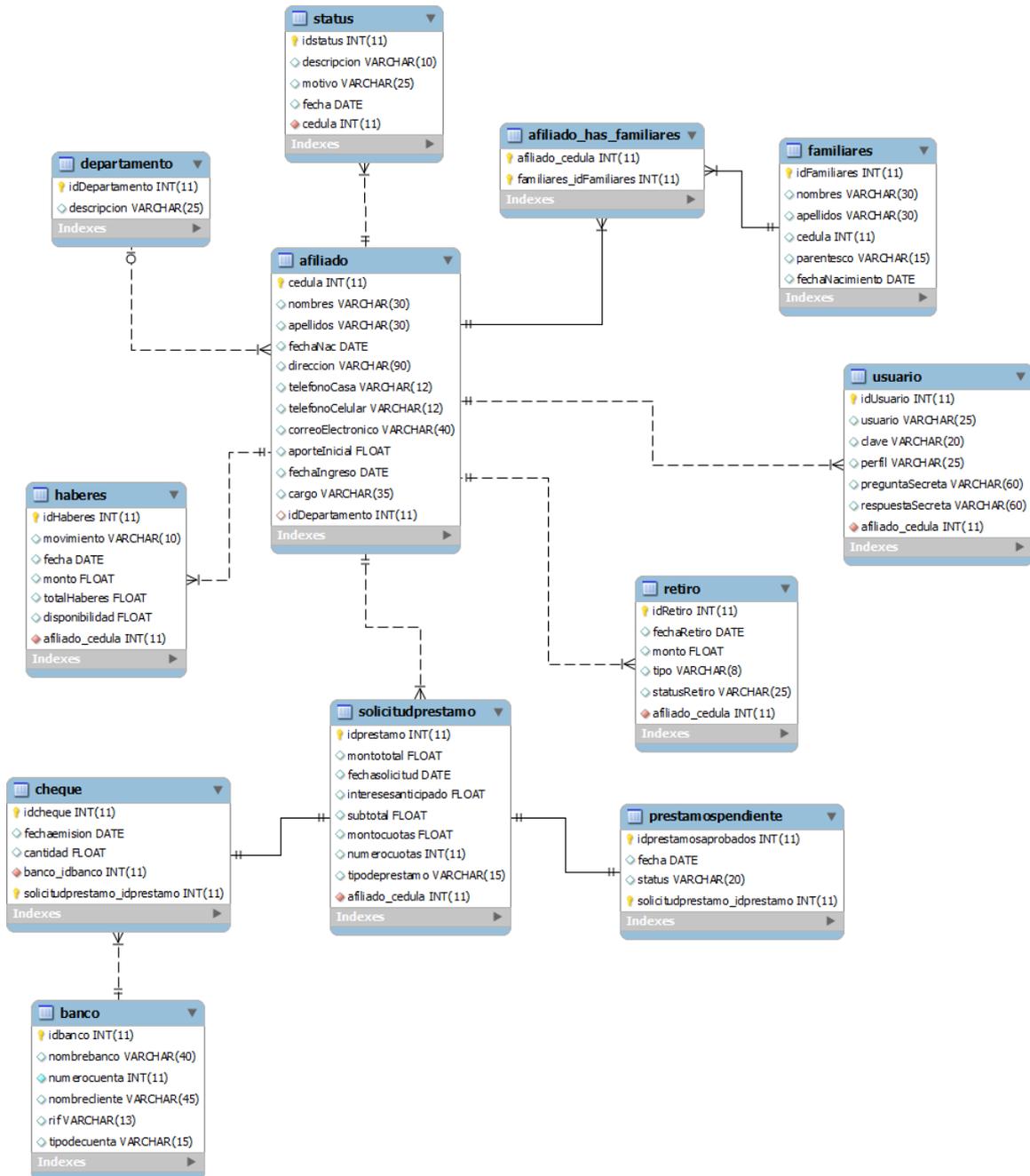
El Modelo de Datos Físico será un modelo para implementar en cualquier Sistema de Gestión de Base de Datos. La cual conlleva la implementación de la forma de las tablas, índices, particiones, etc. El Modelo nos ayuda a tener una vista global de toda la estructura de la base de datos.

A continuación, se detallan las características del Modelo Físico (Fatima, 2019):

- El Modelo de Datos Físicos visualiza las necesidades del proyecto. Dependiendo de su alcance, pueden llegar a integrarse con los demás Modelo de Datos Físicos.
- El Modelo se vincula entre los datos mediante sus tablas. Aborda la cardinalidad de relaciones entre tablas.
- Se debe representar las columnas con su tipo de datos, valores predeterminados y la longitud exacta que tomen en todo el Modelo de Datos. Para así tener la representación de los datos.
- Los índices, los perfiles de acceso y autorización a las tablas están predefinidas.

La siguiente Figura 9 es un ejemplo de modelo físico de datos.

Figura 9. Modelo Físico



Nota: En la Figura 9 muestra un modelo Físico de una aerolínea. Adaptado de *Modelo Físico de la Base de Datos* [Fotografía], por J. Félix, 2012, Caja Ahorros Jfr (<http://caja-ahorros-jfr.blogspot.com/2012/12/modelo-fisico-de-la-base-de-datos.html>). CC BY 2.0

2.3 Definición de Términos Básicos

- **Aplicación Móvil:** Es un programa informático, diseñada para facilitar una tarea en un teléfono móvil. (San Mauro Martín, et al, 2014).
- **Base de datos:** Marqués la define como "un conjunto de archivos, colección de información que están almacenados en una memoria externa y están organizados mediante una estructura de datos" (2011, p. 10).
- **Data Ingestion:** Es la primera capa del Data Lake donde están almacenados todos los tipos de datos que serán procesados en las siguientes capas (Alonso C. , Pulido, Cartón, & Bregon, 2019).
- **Data Processing:** Es la segunda capa del Data Lake en donde los datos se encuentran estandarizados para su futura explotación (Alonso C. , Pulido, Cartón, & Bregon, 2019).
- **Data Service:** Es la ultima capa del Data Lake donde se diseñan los Data Mart y son explotados, por los usuarios de los datos (Alonso C. , Pulido, Cartón, & Bregon, 2019).
- **Data Sources:** Componente del Data Lake donde se encuentran los datos de diferentes fuentes tanto como externa e interna a la organización (Alonso C. , Pulido, Cartón, & Bregon, 2019).
- **Hive:** Según Rodríguez, Hive consiste en "facilitar la consulta y gestión de grandes conjuntos de datos que residen en almacenamiento distribuidos y además proporciona un mecanismo para la ver la estructura de los datos utilizando un lenguaje similar a SQL llamado HiveQL" (2014, p. 79).
- **Moneda Digital:** Medio de pago el cual permite realizar pagos y transacciones únicamente online (Marthin Moran, 2018).

CAPÍTULO III: DESARROLLO DEL TRABAJO PROFESIONAL

3.1 Determinación y Análisis del Problema

La entidad financiera, objeto de estudio, ofrece diferentes tipos de servicios como: préstamos, tarjetas de crédito, seguros, apertura de cuentas bancarias y un sinfín de beneficios por ser parte de la empresa. Al mismo tiempo la entidad financiera brinda el uso de sus cajeros automático, aplicación web y móvil para las transacciones monetarias de sus clientes. Por motivos de confidencialidad no se mencionará el nombre de la entidad financiera ni el nombre de las aplicaciones que brinda a sus clientes. Por lo que para el desarrollo del trabajo de investigación nos enfocaremos en los datos que genera la Aplicación Móvil de la Moneda Digital que maneja la entidad financiera.

Con respecto a la Aplicación Móvil de la Moneda Digital, agruparemos en clientes bancarizados a quienes tengan asociado su número de teléfono móvil y una cuenta ahorro o sueldo. Los clientes no bancarizados a quienes tengan asociado su número de teléfono móvil y su documento de identidad (DNI).

A raíz de la pandemia por el Coronavirus (COVID-19) y a mediados de marzo del 2020 cuando se inició la cuarentena en todo el mundo. La población se tuvo que adaptar vertiginosamente al uso de la Aplicación Móvil de la Moneda Digital ya que, debido al beneficio del uso del aplicativo, no se manipulaban billetes ni monedas y podían establecer el distanciamiento social dictado por el gobierno, evitando la posibilidad de contagio.

Actualmente tenemos que los datos van en aumento cada día, los diferentes tipos de datos que se requiere almacenar, el tiempo en extraer, procesar y analizarlos con las diferentes herramientas tradicionales no están preparados para las exigencias actuales.

Se detalla a continuación el crecimiento que ha tenido el uso de la Aplicación Móvil de la Moneda Digital, la cual se generó una alta volumetría

de datos en el Data Warehouse. Provocando que los datos del aplicativo sean migrados al Data Lake para posteriormente sean modelados antes de la carga en la capa Data Processing, debió a lo cual será objeto de estudio para el trabajo de investigación:

- Las suscripciones al aplicativo de la monea digital empezó a crecer tanto en los clientes bancarizados y los no bancarizados que representan alrededor del 75% del total del mercado (**Anexo B**) desde su lanzamiento en el año 2016.
- Adicional la descarga del aplicativo digital creció vertiginosamente. Antes de la pandemia se tenía 100 mil usuarios nuevos por mes en promedio. Ahora tenemos 500 mil nuevos usuarios por mes en el año 2020 aproximante (**Anexo C**).
- Las transacciones antes eran de 1 millón al mes en promedio en el año antes de la pandemia. Luego en el año 2020, se realiza 11 millones de transacciones al mes (**Anexo D**) en promedio, una cifra récord en el ecosistema financiero.
- Con respecto al volumen de dinero, se ha pasado de un promedio de S/. 100 millones por mes en promedio antes de la pandemia a más de S/. 500 millones en promedio en el año 2020 (**Anexo E**).

3.2 Modelo de la solución propuesta

El primer paso para el diseño de un Modelo de Datos es el desarrollo del Diagrama Entidad Relación.

Se identificaron todos los datos pertenecientes a la Aplicación Móvil de la Moneda Digital. Se estableció el origen e importancia de los datos dentro de la empresa y fueron consolidados en categorías según su relación. Todo

esto, con el fin de definir un conjunto de requisitos y restricciones que permitan generar el Diagrama Entidad Relación.

3.2.1 Requisitos

- Se extraen los 2 archivos planos del Aplicativo Móvil de la Moneda Digital que están almacenados en la Capa Ingestion del Data Lake. Estos archivos son la carga diaria que se tendrán en el repositorio del Data Lake.

Figura 10. Archivos del Aplicativo Móvil

Nombre	Tipo	Tamaño
AFILIADO MONEDA DIGITAL	Archivo de valores separados por co...	1 KB
TRANSACCION MONEDA DIGITAL	Archivo de valores separados por co...	2 KB

Nota: En la Figura 10 se muestran los archivos del Aplicativo Movil de la Moneda Digital. Tomado de *Entidad Financiera*.

- Los archivos planos contienen los datos de los Afiliados y de las transacciones del Aplicativo Móvil de la Moneda Digital.

Figura 11. Archivo de Afiliados

	A	B	C	D	E	F	G	H	I	J
1	Codigo Unico Afiliado Aplicativo Afiliado Numero Telefono Fecha Afiliacion Fecha Bloqueo Saldo Afiliado Numero Cuenta									
2	1	Justina Cayhualla Humani	999999999	20/09/2016		S/300.00	148749446697			
3	2	Dina Maribel Martinez Cayhualla	976543213	6/06/2018		S/65.00	148657464665			
4	3	Geovanna Graciela Martinez Cayhualla	987654321	5/09/2019		S/46.00	163478654746			
5	4	Carmen Rosa Martinez Cayhualla	956789432	5/05/2020		S/65.00	134754847856			
6	5	Renzo Omar Martinez Cayhualla	912345678	14/02/2019		S/23.00	135468745465			
7	6	Yerson Rojas Huamani	999999888	10/11/2018		S/78.00	123456789012			
8	7	Javier Nahui Rojas	999887776	1/08/2020		S/50.00	120123456789			
9	8	Jose Miguel Poma Rojas	999938148	25/08/2019		S/60.00	145432167895			
10	9	Andres Huamani Atequipa	999955543	10/01/2019		S/90.00	159876235876			
11	10	Julio Martinez Quispe	999957654	6/07/2020		S/46.00	189654356876			

Nota: En la Figura 11 se muestran los datos del afiliado del Aplicativo Móvil de la Moneda Digital. Tomado de *Entidad Financiera*.

Figura 12. Archivo de Transacciones de la Moneda Digital

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Transaccion	Operación	Numero Telefono Origen	Numero Telefono Destino	Estado	Monto	Fecha Transaccion	Mensaje Transaccion	tipo Transaccion	Cuenta Origen	Cuenta Destino			
2	1	46874	999999888	999999999	Exitoso	5/2.00	1/12/2020	Pago taxi	Codigo QR	148749446697	189654356876			
3	2	48658	999887776	976543213	Exitoso	5/3.00	2/12/2020	Pago ropa	Codigo QR	148657464665	159876235876			
4	3	35648	999938148	987654321	Exitoso	5/5.00	3/12/2020	Pago medicina	Contacto Telefonico	163478654746	145432167895			
5	4	61754	999955543	956789432	Exitoso	5/0.50	4/12/2020	Pago pollo	Contacto Telefonico	134754847856	120123456789			
6	5	76514	999957654	912345678	Exitoso	5/1.00	5/12/2020	Pago gasolina	Contacto Telefonico	135468745465	123456789012			
7	6	68214	999999999	999999888	Exitoso	5/2.00	6/12/2020	Pago ropa	Contacto Telefonico	189654356876	148749446697			
8	7	31052	976543213	999887776	Exitoso	5/5.00	7/12/2020	Pago medicina	Contacto Telefonico	159876235876	148657464665			
9	8	35143	987654321	999938148	Exitoso	5/6.00	8/12/2020	Para tu caramelo	Contacto Telefonico	145432167895	163478654746			
10	9	32541	956789432	999955543	Exitoso	5/7.00	9/12/2020	Para tu caramelo	Contacto Telefonico	120123456789	134754847856			
11	10	21452	912345678	999957654	Exitoso	5/20.00	10/12/2020	Pago gasolina	Codigo QR	123456789012	135468745465			

Nota: En la Figura 12 se muestran los datos de transacciones del Aplicativo Móvil de la Moneda Digital. Tomado de *Entidad Financiera*.

Con los datos de los 2 archivos planos del afiliado y transacción del Aplicativo Móvil de la Moneda Digital se procederá a modelar el diagrama Entidad Relación indicando las entidades y atributos que tienen. Así mismo, este modelo Entidad Relación se integrará con las entidades ya existentes en la Capa Processing del Data Lake.

3.2.2 Diagrama ER

Este diagrama buscar agrupar la totalidad de todos los datos en forma organizada, mediante la definición de entidades y sus atributos.

Primero se procederá a detallar las entidades o tablas creadas a partir de los archivos planos, el cual se describe a continuación:

- **AfiliadoAplicativo:** Contiene los clientes que están afiliados al Aplicativo Móvil de la Moneda Digital, los cuales contiene tanto a los clientes bancarizados y no bancarizados.
- **Transacciones:** Contiene todos los movimientos y transacciones que genera el Aplicativo Móvil de la Moneda Digital.

A continuación, en la Tabla 9 se procederá a definir los atributos de cada una de las entidades mencionadas anteriormente.

Tabla 9. Descripción de Atributos de los 2 Archivos Planos

Entidad	Atributo	Explicación
AfiliadoAplicativo	Codigo Unico Afiliado Aplicativo	Clave principal de la tabla Afiliado Aplicativo
	Afiliado	Datos personales del afiliado como sus nombres y apellidos
	NumeroTelefono	Numero de teléfono móvil asociado
	Fecha Afiliacion	fecha afiliación del cliente
	Fecha Bloqueo	fecha bloqueo del cliente
	Saldo Afiliado	Saldo disponible del afiliado
	Numero Cuenta	Número de cuenta que tiene asociado el cliente bancarizado
Transacciones	Codigo Unico Transaccion	Clave principal de la tabla Transaccion
	Numero Operación	Numero de operación de la transaccion
	Numero Telefono Origen	Numero de teléfono móvil quien realiza la transferencia de dinero
	Numero Telefono Destino	Numero de teléfono móvil quien recibe la transferencia de dinero
	Estado	Estado de la transaccion
	Monto Transaccion	Monto de la transaccion realizada
	Fecha Transaccion	Fecha de la transaccion realizada
	Comentario	Comentario que se coloca al momento de realizar la transaccion
	Tipo Transaccion	Tipo de transaccion, puede tomar los siguientes valores. 1. Tipo QR, cuando la transferencia es media el scaneo de algun codigo QR 2. Contacto Movil, Cuando la transferencia es realizada en algun contacto del teléfono móvil
	Numero Cuenta Origen	Numero Cuenta cliente quien envia el monto
Numero Cuenta Destino	Numero Cuenta cliente quien recibe el monto	

Nota: En la Tabla 9 se describe los atributos de las entidades que se modelo a partir de los archivos planos. Tomado de *Entidad Financiera*.

Luego, se procederá a detallar las entidades o tablas ya existentes en la Capa Processing del Data Lake, el cual se describe a continuación:

- **CuentaFinanciera:** Identifica a los clientes bancarizados, los cuales son aquellos que tienen afiliado su cuenta bancaria con el Aplicativo Móvil de la Moneda Digital.
- **Cliente:** Contiene los datos básicos del cliente que puede ser usado por cualquier modelo o entidad en el banco.

Tabla 10. Entidades Existentes en la Capa Processing

Entidad	Atributo	Explicación
CuentaFinanciera	Codigo Unico Cuenta	Identificador de la tabla Cuenta Financiera
	Codigo Unico Cliente	Clave principal de la tabla Cliente
	Numero Cuenta	Numero de cuenta que tiene asociado el cliente banzarizado
	Tipo Moneda	Tipo de moneda asociado a la cuenta bancaria
	Saldo Cuenta	Saldo disponible en la cuenta
	Fecha-Afiliacion	Fecha de afiliacion
	Fecha-Bloqueo	fecha de bloqueo de la cuenta
	Nombre Agencia	Nombre de la agencia en donde se realizo la apertura de la cuenta financiera
	Tipo Documento	Tipo Documento cliente asociado
	Numero Documento	Numero documento cliente asociado
Cliente	Codigo Unico Cliente	Codigo Unico de la tabla Cliente
	Nombres	Nombres del cliente
	Apellido Paterno	Apellido Paterno del cliente de la tabla cliente
	Apellido Materno	Apellido Materno del cliente de la tabla cliente
	Fecha Afiliacion	Fecha Afiliación del cliente de la tabla cliente

Nota: En la Tabla 10 se describe los atributos de las entidades existentes en la Capa Processing del Data Lake. Tomado de *Entidad Financiera*.

Por último, se procederá a integrar las entidades, con sus respectivos atributos, de los archivos planos y existentes en la Capa Processing del Data Lake. En la Figura 10, se define que contemplara con un total de 4 entidades, cada uno con sus respectivos atributos. Este diagrama define la relación que existen entre las entidades.

Tabla 11. Entidades en la Capa Processing del Data Lake

Entidad	Atributo	Explicación
AfiliadoAplicativo	Codigo Unico Afiliado Aplicativo	Clave principal de la tabla Afiliado Aplicativo
	Codigo Unico Cliente	Clave principal de la tabla Cliente
	NumeroTelefono	Numero de telefono movil asociado
	Fecha Afiliacion	fecha afiliacion del cliente
	Fecha Bloqueo	fecha bloqueo del cliente
	Saldo Afiliado	Saldo disponible del afiliado
	Codigo Unico Cuenta	Identificador de la tabla cuenta financiera
Transacciones	Codigo Unico Transaccion	Clave principal de la tabla Transaccion
	Numero Operación	Numero de operación de la transaccion
	Numero Telefono Origen	Numero de telefono movil quien realiza la transferencia de dinero
	Numero Telefono Destino	Numero de telefono movil quien recibe la transferencia de dinero
	Estado	Estado de la transaccion
	Monto Transaccion	Monto de la transaccion realizada
	Fecha Transaccion	Fecha de la transaccion realizada
	Comentario	Comentario que se coloca al momento de realizar la transaccion
	Tipo Transaccion	Tipo de transaccion, puede tomar los siguientes valores. 1. Tipo QR, cuando la transferencia es media el scaneo de algun codigo QR 2. Contacto Movil, Cuando la transferencia es realizada en algun contacto del telefono movil
	Codigo Unico Cuenta Origen	Codigo Unico cliente quien envia el monto
	Codigo Unico Cuenta Destino	Codigo Unico cliente quien recibe el monto

CuentaFinanciera	Codigo Unico Cuenta	Identificador de la tabla Cuenta Financiera
	Codigo Unico Cliente	Clave principal de la tabla Cliente
	Numero Cuenta	Numero de cuenta que tiene asociado el cliente banzarizado
	Tipo Moneda	Tipo de moneda asociado a la cuenta bancaria
	Saldo Cuenta	Saldo disponible en la cuenta
	Fecha-Afiliacion	Fecha de afiliacion
	Fecha-Bloqueo	fecha de bloqueo de la cuenta
	Nombre Agencia	Nombre de la agencia en donde se realizo la apertura de la cuenta financiera
	Tipo Documento	Tipo Documento cliente asociado
	Numero Documento	Numero documento cliente asociado
Cliente	Codigo Unico Cliente	Codigo Unico de la tabla Cliente
	Nombres	Nombres del cliente
	Apellido Paterno	Apellido Paterno del cliente de la tabla cliente
	Apellido Materno	Apellido Materno del cliente de la tabla cliente
	Fecha Afiliacion	Fecha Afiliacion del cliente de la tabla cliente

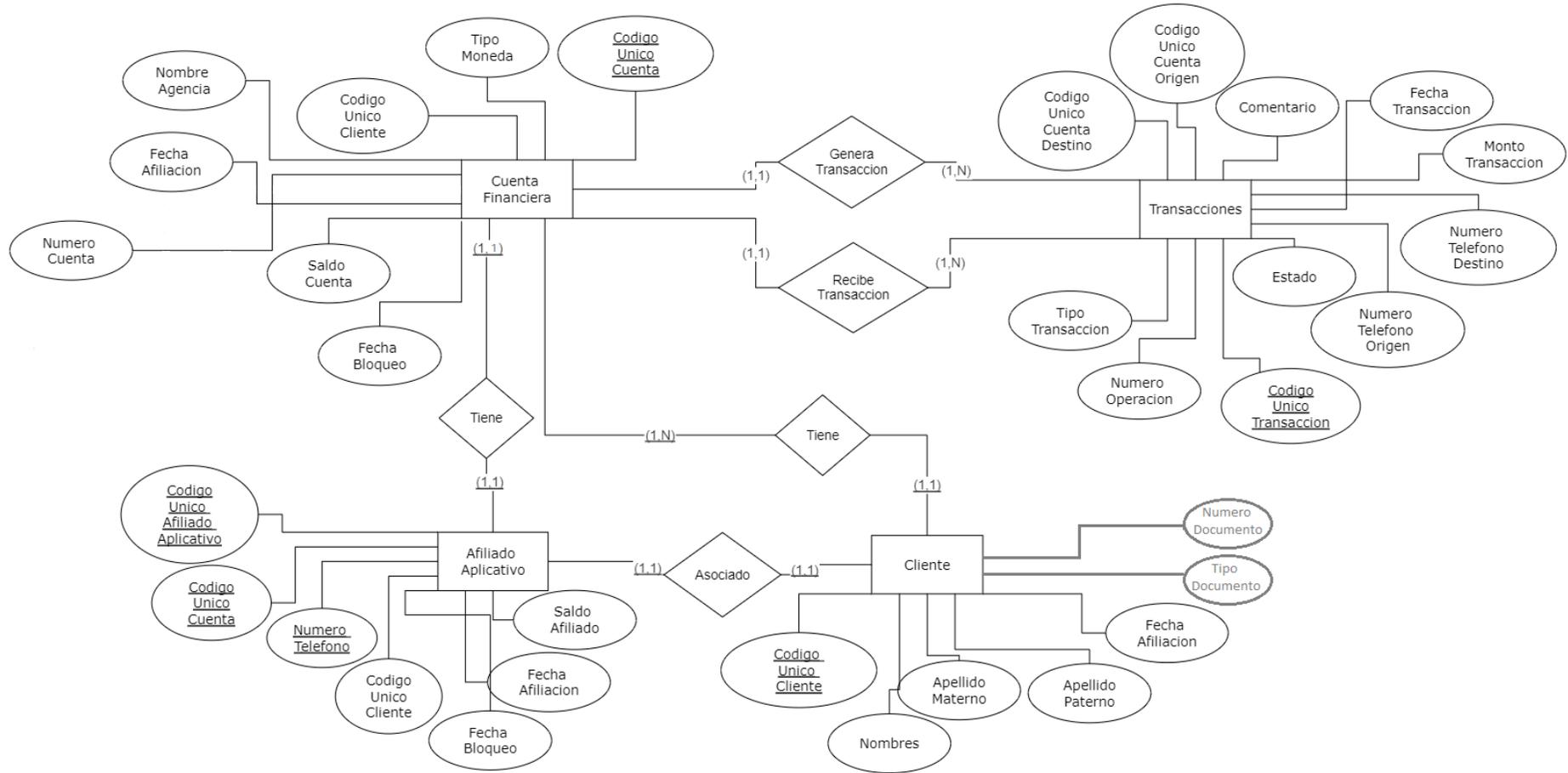
Nota: En la Tabla 11 se describe las entidades y sus atributos de la Capa Processing del Data Lake. Tomado de *Entidad Financiera*.

Las relaciones entre las entidades se establecen de la siguiente manera:

- Cada CuentaFinanciera puede tener muchas Transacciones, la relación es de 1:N entre CuentaFinanciera y Transacciones.
- Cada AfiliadoAplicativo puede estar asociado a cada CuentaFinanciera, la relación es de 1:1 entre AfiliadoAplicativo y CuentaFinanciera.
- Cada Cliente puede estar asociado a cada AfiliadoAplicativo, la relación es de 1:1 entre Cliente y AfiliadoAplicativo.

- Cada CuentaFinanciera puede estar asociado a cada Cliente, la relación es de 1:1 entre CuentaFinanciera y Cliente.

Figura 13. Modelo Entidad Relación



Nota: En la Figura 13 se muestra la propuesta de un Modelo Entidad Relación para una entidad Financiera.

3.2.3 Normalización

Para pasar al diagrama relación lógico, se hace la normalización, buscando encontrar errores y evitar la redundancia de los datos. Esta permite evaluar si el diagrama Entidad Relación cumple con las reglas de normalización para convertirlas en tablas. Estas reglas pueden llegar en forma normales, que pueden llegar hasta la sexta forma normal(6FN), estableciendo que para el presente trabajo se trabajara hasta la tercera forma normal(3FN).

Debido a que el Diagrama Entidad Relación se modelo cumpliendo la 1FN; teniendo los atributos atómicos en las 4 entidades tal como se muestra en la Figura 10. Y cumpliendo con la 2 FN indicando la clave primaria en cada entidad tal como se mostrarán a continuación:

Tabla 12. *Tabla Cliente*

Codigo Unico Cliente	Nombres	Apellido Paterno	Apellido Materno	Tipo Documento	Numero Documento	Fecha Afiliacion
1	Renzo Omar	Martinez	Cayhualla	DNI	48597898	17/01/2018
2	Carmen Rosa	Martinez	Cayhualla	DNI	47263641	5/03/2020
3	Geovanna Graciela	Martinez	Cayhualla	DNI	42612584	1/07/2019
4	Dina Maribel	Martinez	Cayhualla	DNI	46895656	6/06/2017
5	Justina	Cayhualla	Humani	DNI	49563588	14/08/2014
6	Julio	Martinez	Quispe	DNI	49887989	5/05/2019
7	Andres	Huamani	Atequipa	DNI	49589787	22/06/2018
8	Jose Miguel	Poma	Rojas	DNI	49878749	25/07/2016
9	Javier	Ñahui	Rojas	DNI	42156878	10/09/2020
10	Yerson	Rojas	Huamani	DNI	43887947	23/09/2018

Nota: En la Tabla 12 se muestra la entidad Cliente con la 1FN y la 2FN. Tomado de *Entidad Financiera*.

Tabla 13. Tabla AfiliadoAplicativo

Codigo Unico Afiliado Aplicativo	Codigo Unico Cliente	Numero Telefono	Fecha Afiliacion	Fecha Bloqueo	Saldo Afiliado	Codigo Unico Cuenta
1	5	999999999	20/09/2016		S/300.00	10
2	4	976543213	6/06/2018		S/65.00	9
3	3	987654321	5/09/2019		S/46.00	8
4	2	956789432	5/05/2020		S/65.00	7
5	1	912345678	14/02/2019		S/23.00	6
6	10	999999888	10/11/2018		S/78.00	1
7	9	999887776	1/08/2020		S/50.00	2
8	8	999938148	25/08/2019		S/60.00	3
9	7	999955543	10/01/2019		S/90.00	4
10	6	999957654	6/07/2020		S/46.00	5

Nota: En la Tabla 13 se muestra la entidad AfiliadoAplicativo con la 1FN y la 2FN.
Tomado de *Entidad Financiera*.

Tabla 14. CuentaFinanciera

Codigo Unico Cuenta	Codigo Unico Cliente	Numero Cuenta	Tipo Moneda	Saldo Cuenta	Fecha Afiliacion	Fecha Bloqueo	Nombre Agencia
1	5	123456789012	Soles	S/300.00	14/08/2014		Chorrillos
2	4	120123456789	Soles	S/65.00	6/06/2017		Barranco
3	3	145432167895	Soles	S/46.00	1/07/2019		Surco
4	2	159876235876	Soles	S/65.00	5/03/2020		Miraflores
5	1	189654356876	Soles	S/23.00	17/01/2018		Villa Salvador
6	10	135468745465	Soles	S/60.00	23/09/2018		Chorrillos
7	9	134754847856	Soles	S/200.00	10/09/2020		San Isidro
8	8	163478654746	Soles	S/163.00	25/07/2016		El agustino
9	7	148657464665	Soles	S/746.00	22/06/2018		La Molina
10	6	148749446697	Soles	S/58.00	5/05/2019		San borja

Nota: En la Tabla 14 se muestra la entidad CuentaFinanciera con la 1FN y la 2FN.
Tomado de *Entidad Financiera*.

Tabla 15. Tabla Transacciones

Codigo Unico Transaccion	Numero Operación	Numero Telefono Origen	Numero Telefono Destino	Estado	Monto Transaccion	Fecha Transaccion	Comentario	Tipo Transaccion	Codigo Unico Cuenta Origen	Codigo Unico Cuenta Destino
1	46874	999999888	999999999	Exitoso	S/2.00	1/12/2020	Pago taxi	Contacto Telefonico	10	5
2	48658	999887776	976543213	Exitoso	S/3.00	2/12/2020	Pago ropa	Contacto Telefonico	9	4
3	35648	999938148	987654321	Exitoso	S/5.00	3/12/2020	Pago medicina	Contacto Telefonico	8	3
4	61754	999955543	956789432	Exitoso	S/0.50	4/12/2020	Pago pollo	Contacto Telefonico	7	2
5	76514	999957654	912345678	Exitoso	S/1.00	5/12/2020	Pago gasolina	Contacto Telefonico	6	1
6	68214	999999999	999999888	Exitoso	S/2.00	6/12/2020	Pago ropa	Contacto Telefonico	5	10
7	31052	976543213	999887776	Exitoso	S/5.00	7/12/2020	Pago medicina	Contacto Telefonico	4	9
8	35143	987654321	999938148	Exitoso	S/6.00	8/12/2020	Para tu caramelo	Contacto Telefonico	3	8
9	32541	956789432	999955543	Exitoso	S/7.00	9/12/2020	Para tu caramelo	Contacto Telefonico	2	7
10	21452	912345678	999957654	Exitoso	S/20.00	10/12/2020	Pago gasolina	Contacto Telefonico	1	6

Nota: En la Tabla 15 se muestra la entidad Transaccion con la 1FN y la 2FN. Tomado de *Entidad Financiera*.

Con la evaluación realizada se garantiza que el Modelo de Datos se encuentra en 2FN, y no será necesario realizar más normalizaciones salvo para la tabla Transacciones. Debido a que el concepto Transacciones se encuentra relacionado con el concepto Tipo Transacciones y Estado. Por lo que se crearan 2 nuevas tablas a partir de la Tabla Transacciones.

Tabla 16. *Tabla Estado*

Codigo Tipo Estado	Descripcion Tipo Estado
1	Exitoso
2	Fallido
3	Abortado

Nota: En la Tabla 16 se muestra la entidad Estado la cual contiene 3 Estados para las Transacciones. Tomado de *Entidad Financiera*.

Tabla 17. *Tipo Transaccion*

Codigo Tipico Transaccion	Descripcion Tipo Transaccion
1	Contacto Telefonico
2	Codigo QR

Nota: En la Tabla 17 se muestra la entidad Tipo Transaccion la cual contiene 2 tipos de transacciones que puede realizar el Aplicativo Móvil de la Moneda Digital. Tomado de *Entidad Financiera*.

Tabla 18. Tabla Transacciones Normalizada

Codigo Unico Transaccion	Numero Operación	Numero Telefono Origen	Numero Telefono Destino	Estado	Monto Transaccion	Fecha Transaccion	Comentario	Tipo Transaccion	Codigo Unico Cuenta Origen	Codigo Unico Cuenta Destino
1	46874	999999888	999999999	1	S/2.00	1/12/2020	Pago taxi	2	10	5
2	48658	999887776	976543213	1	S/3.00	2/12/2020	Pago ropa	2	9	4
3	35648	999938148	987654321	1	S/5.00	3/12/2020	Pago medicina	1	8	3
4	61754	999955543	956789432	1	S/0.50	4/12/2020	Pago pollo	1	7	2
5	76514	999957654	912345678	1	S/1.00	5/12/2020	Pago gasolina	1	6	1
6	68214	999999999	999999888	1	S/2.00	6/12/2020	Pago ropa	1	5	10
7	31052	976543213	999887776	2	S/5.00	7/12/2020	Pago medicina	1	4	9
8	35143	987654321	999938148	2	S/6.00	8/12/2020	Para tu caramelo	1	3	8
9	32541	956789432	999955543	2	S/7.00	9/12/2020	Para tu caramelo	1	2	7
10	21452	912345678	999957654	2	S/20.00	10/12/2020	Pago gasolina	2	1	6

Nota: En la Tabla 18 muestra la entidad Transacciones la cual ya no contiene otros conceptos. Tomado de *Entidad Financiera*.

Con las evaluaciones realizadas se garantiza que el Modelo de Datos en Big Data se encuentra en 3FN, y no será necesario realizar más normalizaciones. Como se ha demostrado, todos los atributos son atómicos, evitan duplicidad de datos, cada llave principal solo contiene una sola columna y los conceptos entre las entidades están separadas.

3.2.4 Modelo Lógico

Este modelo consiste en traducir el Modelo Entidad Relación en un lenguaje entendible. El objetivo de este modelo es organizar las modificaciones que se realizaron al momento de la normalización, incluyendo por lo tanto en el modelo las nuevas tablas creadas. Este modelo permite establecer las tablas definitivas que satisfacen una correcta organización de los datos.

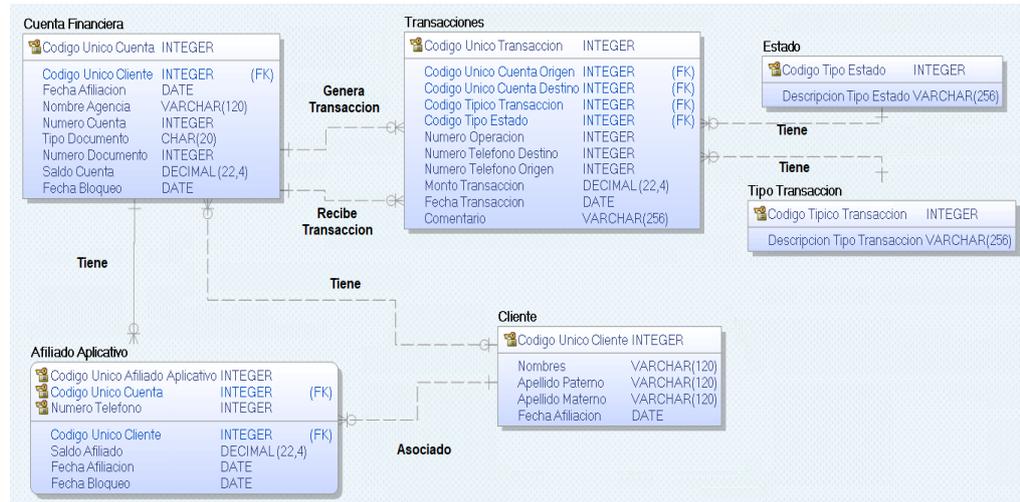
Para construir el Modelo Lógico definitivo, se analizan las relaciones entre las tablas. Una llave foránea consiste en relacionar la llave primaria de una tabla con la llave primaria de la otra.

Para construir el Modelo Lógico definitivo, se analizan las relaciones de uno a muchos (1..N) que existe entre las tablas TipoTransaccion y Transacciones. La relación entre TipoTransaccion y Transacciones, se realiza mediante la llave primaria correspondiente al Código Tipo Transaccion, es decir, en la tabla Transacciones no se almacenan datos del TipoTransaccion, si no por medio de la llave foránea, con el Código Tipo Transaccion, se obtiene la información que se desee de la tabla TipoTransaccion.

Lo mismo sucede con la relación entre las tablas Estado y Transacciones de uno a muchos (1..N). Se realiza mediante la llave primaria correspondiente al Código Tipo Estado, es decir, en la tabla Transacciones no se almacenan datos del estado de las transacciones, si no por medio de la llave foránea, con el Código Tipo Estado, se obtiene la información que se desee de la tabla Estado.

De este modo se realiza el modelo Lógico como se muestra en la Figura 14, la cual permite identificar la totalidad de tablas, sus atributos y llaves primarias y foráneas, necesarias para relacionarlas, completando el Modelo de Datos Lógico propuesto.

Figura 14. Modelo Lógico



Nota: En la Figura 14 muestra las 4 entidades propuestas en el modelo Conceptual y las 2 entidades nuevas a partir de la entidad Transacciones.

3.3 Resultados

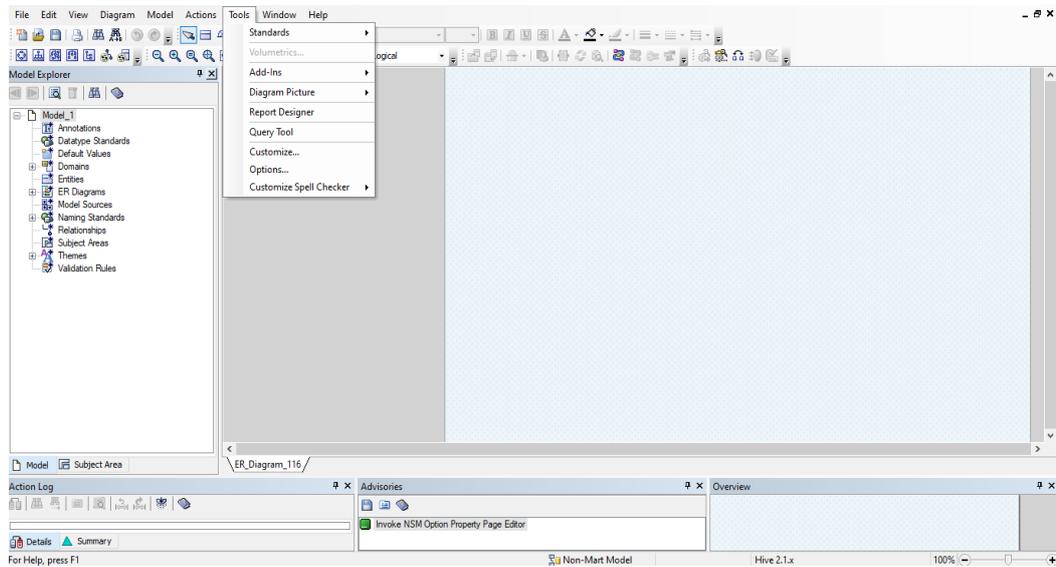
3.3.1 Modelo Físico

Con base al Modelo Lógico y con el apoyo de la herramienta Erwin, se define los nombres de las entidades y los atributos estándar que se tendrá en la capa Data Processing del Data Lake.

A continuación, se mostrará la configuración que se debe realizar para estandarizar todos los atributos el cual seguiremos los siguientes pasos de forma secuencial:

- Nos dirigimos a la barra de herramientas y seleccionamos la opción Tools tal como se muestra en la Figura 15.

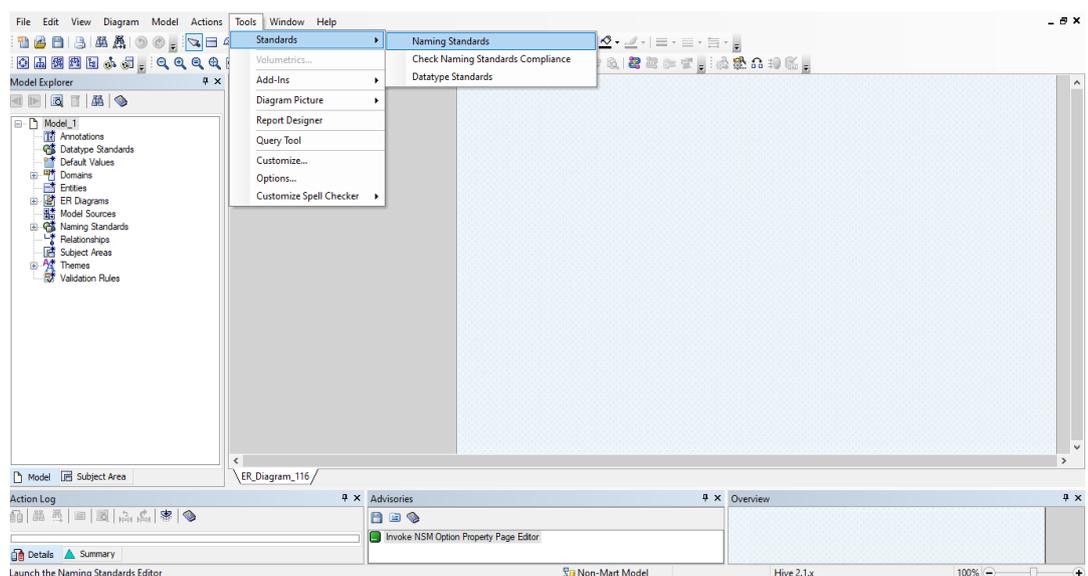
Figura 15. Barra de herramientas del erwin



Nota: En la Figura 15 se muestra en primer lugar los pasos a seguir para la configuración de los atributos. Tomado de *Entidad Financiera*.

- Nos dirigimos a la lista desplegable Standards y seleccionaremos la opción Naming Standards, tal como se muestra en la Figura 16.

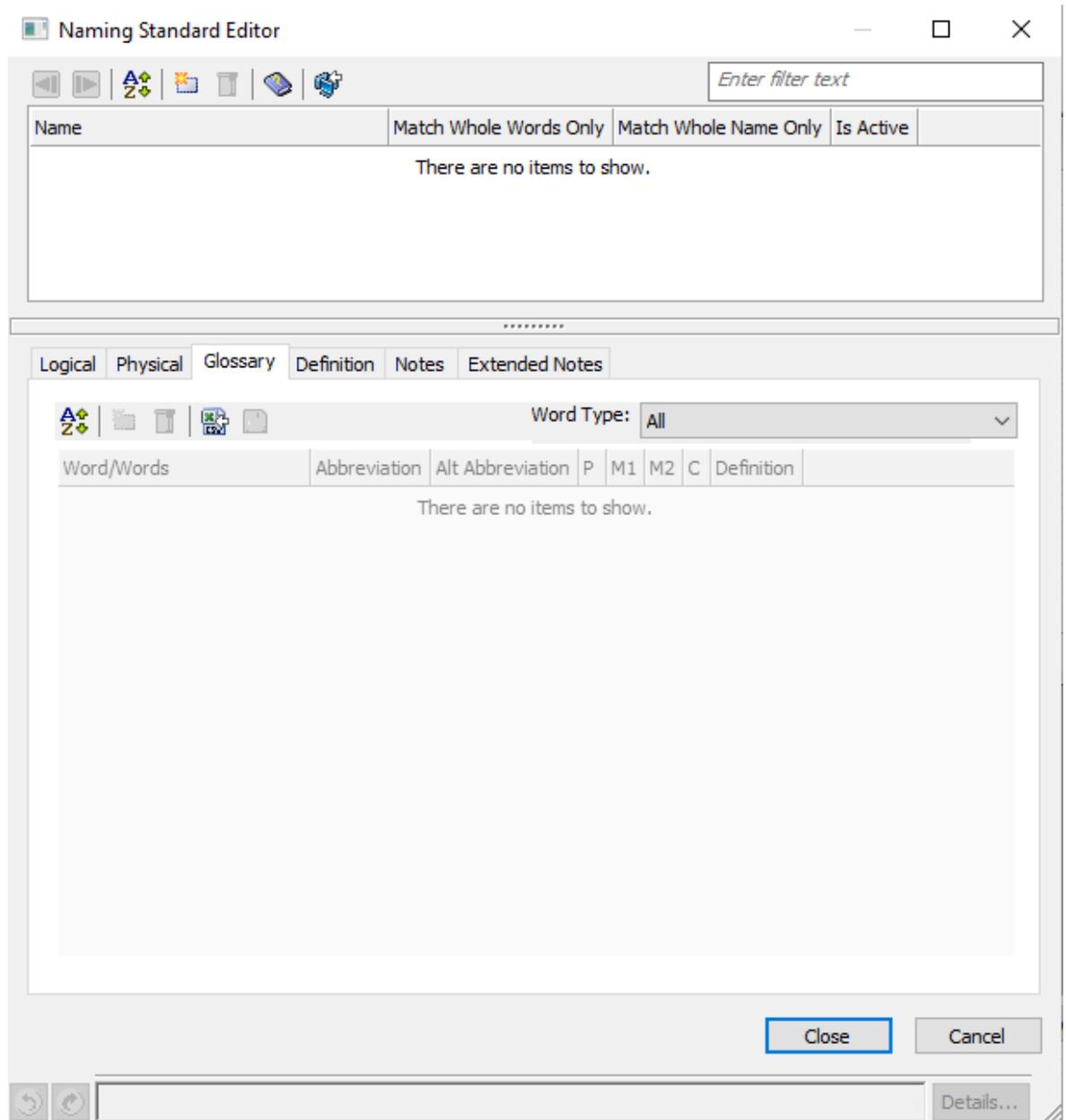
Figura 16. Barra de Herramientas Opción Tools



Nota: En la Figura 16 muestra el siguiente paso que se debe seguir para la configuración de los atributos. Tomado de *Entidad Financiera*.

- Prosiguiendo, nos mostrará la siguiente ventana en donde se procederá a configurar, según el estándar de los atributos que se seguirá por lineamientos, en la Capa Data Processing. Tal como se muestra en la Figura 17.

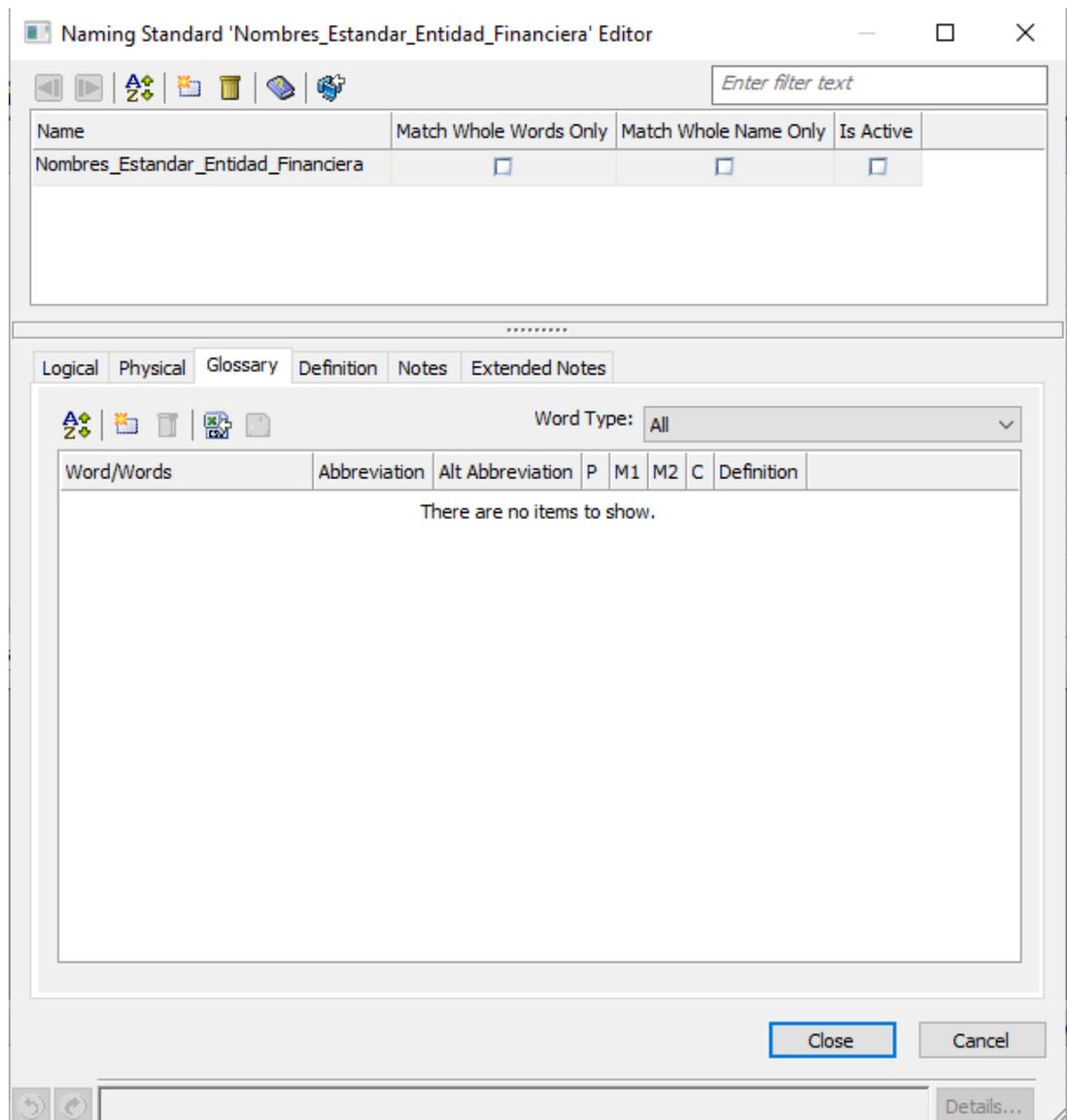
Figura 17. Ventana para la Configuración de Atributos



Nota: En la Figura 16 se muestra la ventana para la configuración de los atributos. Tomado de *Entidad Financiera*.

- A partir de la ventana de la configuración de atributos, tal como se muestra en la Figura 17, se procederá a crear una etiqueta en donde se guardará todos los atributos el cual llamaremos Nombres_Estandar_Entidad_Financiera; como se muestra en la Figura 18.

Figura 18. Creación de Etiqueta en la Ventana para la Configuración de Atributos

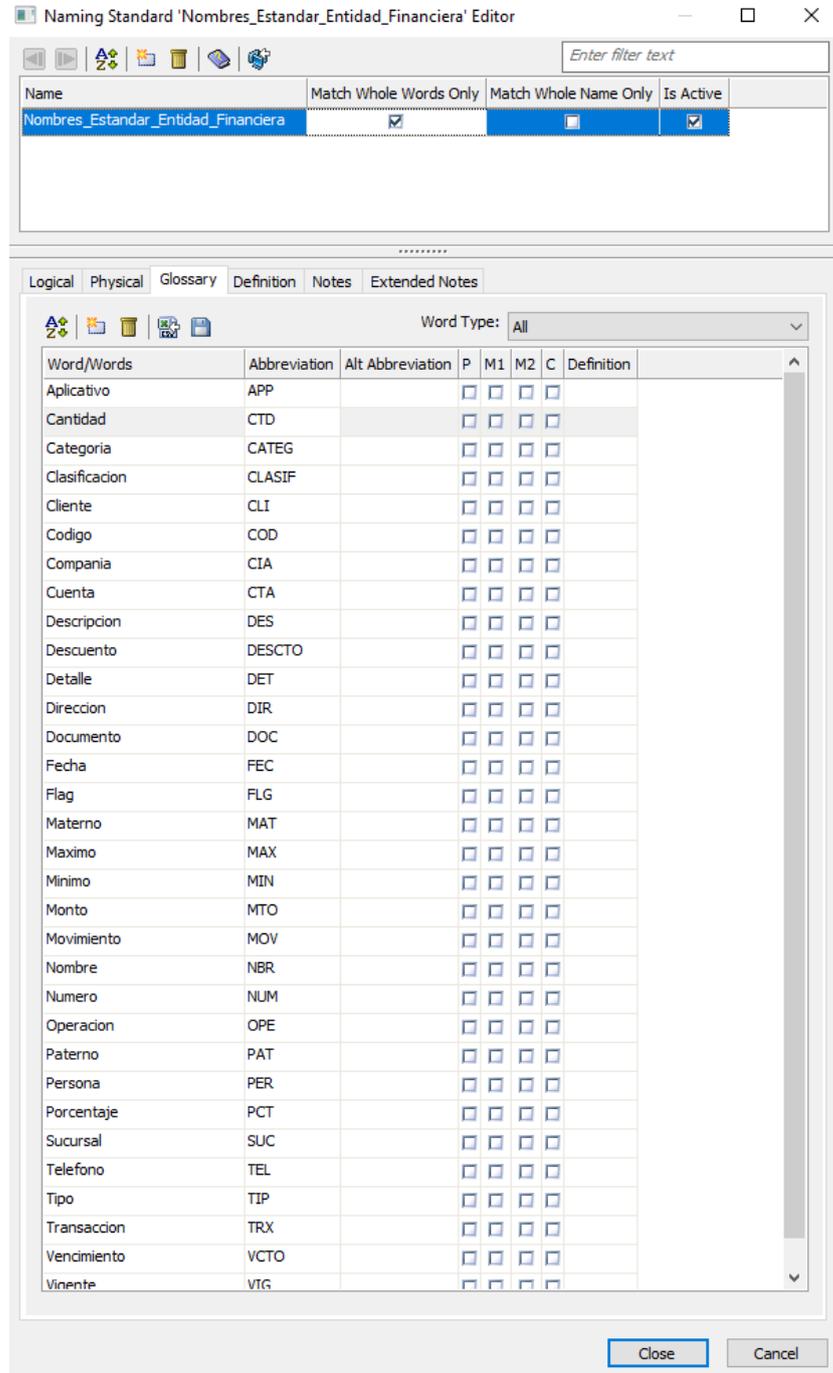


Nota: En la Figura 18 se creará la etiqueta Nombres_Estandar_Entidad_Financiera para guardar los atributos en esta. Tomado de *Entidad Financiera*.

- Seguidamente se procederá a modelar los atributos según se establezca en la columna Abbreviation tal como se muestra en la

Figura19. El cual seguirá este estándar en todos los atributos de la Capa Data Processing del Data Lake.

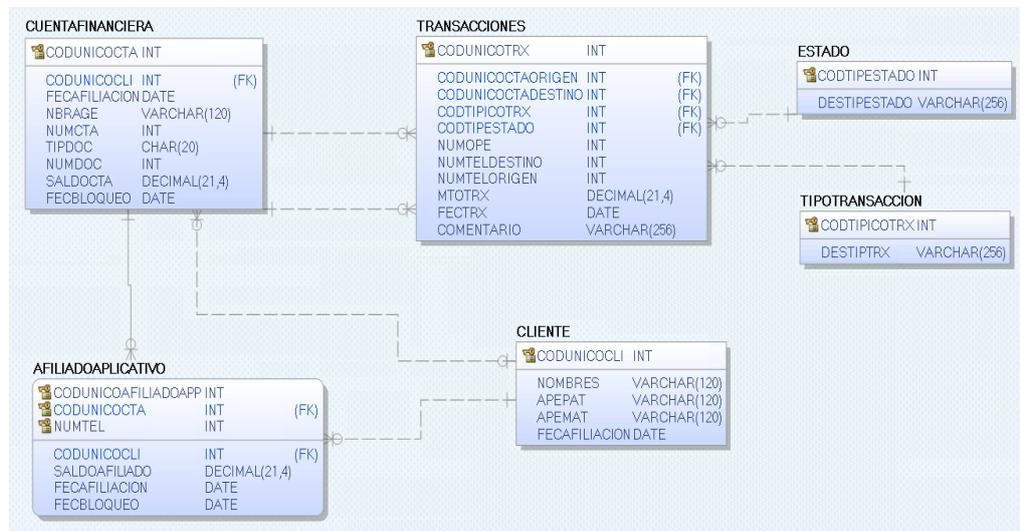
Figura 19. Estandarización de Atributos



Nota: En la Figura 19 se estandarizo los atributos que se utilizara en la Capa Data Processing del Data Lake. Tomado de *Entidad Financiera*.

Inmediatamente se procederá a construir el Modelo Físico de Datos, en base a los lineamientos indicados para la estandarización de los atributos y del Modelo Lógico de Datos tal como se muestra en la Figura 20. Cabe mencionar que pueda haber atributos que no estén estandarizados debido a que aún no estén contemplados en los lineamientos por parte de la Entidad Financiera, por lo que cuando se genere el Modelo Físico de Datos se mantendrá su nombre sin ningún cambio.

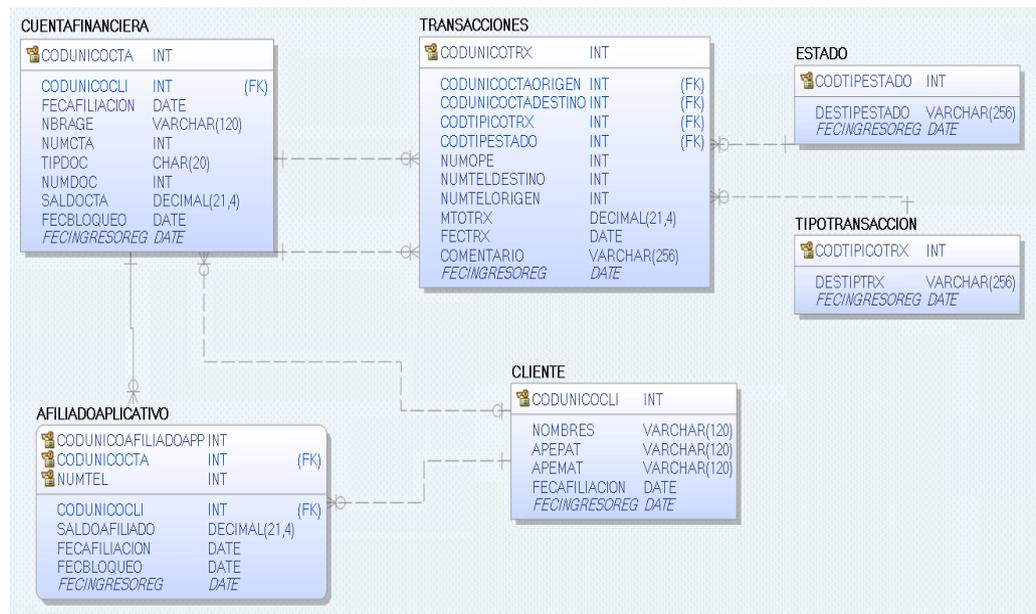
Figura 20. Modelo Físico de Datos



Nota: En la Figura 20 se logra ver el Modelo Físico de Datos propuesto para la Entidad Financiera.

Por último, a cada tabla se le añadirá el atributo *Fecha Ingreso Registro* la cual estará en *cursiva* e indicará la fecha de ingreso del registro al Data Lake para un mayor manejo de los datos. Este Modelo es la propuesta final del trabajo de investigación, la cual se contemplará en la Figura 21.

Figura 21. Modelo Físico Final



Nota: En la Figura 21 muestra el Modelo Físico Final propuesto para una Entidad Financiera.

CONCLUSIONES

- La propuesta del Modelo de Datos en la capa Data Processing del Data Lake nos ayudara a poder manejar un estándar de datos como las tablas, atributos y la relación que existe. Así mismo integrara los nuevos modelos que se implementara para evitar la redundancia de datos en el Modelo Fisico del Data Lake
- La identificación de los datos y sus relaciones permitirán que el Modelo Físico simplificara a un 50% el número de datos almacenados en las 6 tablas. Es decir, debido al modelo propuesto solo almacenaran datos que no existan en la capa Data Processing del Data Lake.
- La consolidación de datos en un único ambiente como el Data Processing, la cual permitirá a cualquier miembro de la entidad financiera poder visualizar los datos, para su futura explotación en la capa Data Service del Data Lake.
- Los aportes que genera el Data Lake son muy importantes ya que al tener los datos separados en varias capas nos permitirá manejar una capa estándar (Data Processing), la cual ayudara al ordenamiento de datos en el Data Lake.

RECOMENDACIONES

- Revisar el Modelo de Datos propuesto cuando se requiera hacer nuevos modelos para evitar redundancia e integrarlo en un solo modelo en la Capa Data Processing.
- Tener los datos claros y completos en la capa Data Ingestion para poder realizar un correcto modelado en la capa Data Processing.
- Seguir la secuencialidad de la arquitectura referencial del Data Lake. Ya que antes de llevar los datos a la capa Data Service, en donde será explotada, debe pasar primeramente por la capa Data Processing.
- El modelo podría ser implementado en cualquier motor de base de datos.

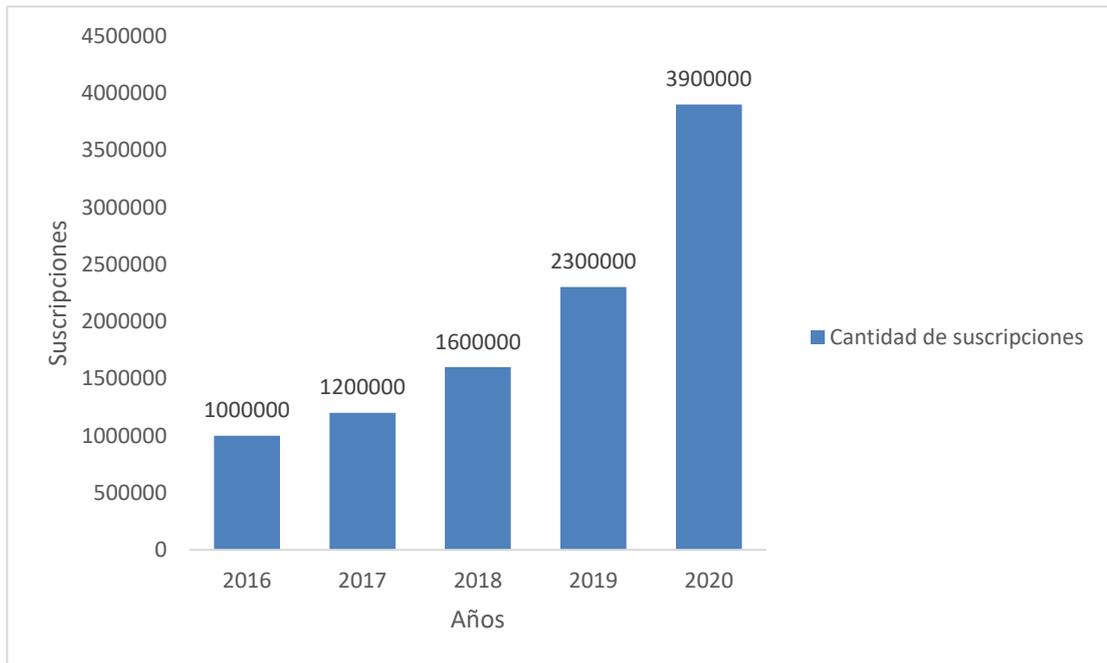
Referencias

- Abraham Silberschatz, H. F. (2002). *Fundamentos de Base de datos*. Madrid: McGraw-Hill Interamericana de España S.A.U.
- Almora, N. (2018). *Análisis y usos del Big data aplicado en la universidad nacional "San Luis Gonzaga" de Ica: caso Facultad de Ingeniería de Sistemas [Tesis de Licenciatura, Universidad Nacional San Luis Gonzaga]*. Repositorio Institucional. Obtenido de <https://repositorio.unica.edu.pe/handle/UNICA/3095>
- Alonso, C. J., Pulido, B., & Carton, M. (16 de 07 de 2019). *IEEE*. Obtenido de IEEE ORG: <https://ieeexplore.ieee.org/document/8764332>
- Alonso, C., Pulido, B., Cartón, M., & Bregon, A. (2019). A Big Data Architecture for Fault Prognostics of Electronic Devices: Application to Power MOSFETs. *IEEE Access*, 7, 102163. doi:10.1109/ACCESS.2019.2929111
- Centro Europeo de Postgrado. (26 de Junio de 2021). *¿Qué es el Modelado de Datos en Big data?* Obtenido de mbaonlineceupe: https://mbaonlineceupe.com/que-es-el-modelado-de-datos-en-big-data/#El_Modelado_de_Datos_y_el_Modelado_de_clases_diferencia
- Daniel Gudiel, J. C. (2014). *CORE*. Obtenido de CORE : <https://core.ac.uk/download/pdf/45359197.pdf>
- Fatima, N. (29 de Agosto de 2019). *Qué, por qué y cómo de las herramientas de modelado de datos*. Obtenido de Astera: <https://www.astera.com/es/type/blog/understanding-data-modeling-tools/>
- Giordana, D. (15 de Noviembre de 2017). *EL MODELO ENTIDAD-RELACION Y EL MODELO RELACIONAL EN BASES DE DATOS*. Obtenido de bandageek: <https://bandageek.com/2017/02/el-modelo-entidad-relacion-y-el-modelo-relacional-en-bases-de-datos/>
- Gomez, M. (2013). *libroselec*. Publidisa Mexicana S.A de C.V. Obtenido de http://www.cua.uam.mx/pdfs/conoce/libroselec/Notas_del_curso_Bases_de_Datos.pdf
- guru99es. (17 de Abril de 2018). *¿Qué es la normalización? 1NF, 2NF, 3NF y BCNF con ejemplos*. Obtenido de Guru99: <https://guru99.es/database-normalization/>
- Honores, J., Rivas, W., & Zea, M. (2016). *Fundamentos de bases de datos*. Machala. Obtenido de <http://repositorio.utmachala.edu.ec/handle/48000/6925>
- Huapaya, J. (2021). *Integración de entorno big data en la gestión financiera de un sistema bancario para optimizar productos, servicios internos y toma de decisiones[Tesis de Licenciatura, Universidad Nacional Mayor de San Marcos]*. Obtenido de <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/17040>
- José, F. (2012). *Modelo Físico de la base de datos*. Obtenido de Caja Ahorros jfr: <http://caja-ahorros-jfr.blogspot.com/2012/12/modelo-fisico-de-la-base-de-datos.html>

- Linares, C. (2019). *Implementación de un sistema de Big Data aplicado a la migración de datos bajo la distribución cloudera con Apache Hadoop, en el Banco Interbank [Tesis de Licenciatura, Universidad Tecnológica del Peru]*. Repositorio Institucional. Obtenido de <https://repositorio.utp.edu.pe/handle/20.500.12867/1944>
- Lorenti, R. (2019). *El Big Data aplicado en las redes sociales para una empresa de telefonía móvil celular en Ecuador [Tesis de Licenciatura, Universidad Tecnológica Empresarial de Guayaquil]*. Obtenido de <http://biblioteca.uteg.edu.ec/xmlui/handle/123456789/983>
- Marthin Moran, F. P. (Marzo de 2018). *BCRP*. Obtenido de Monedas Digitales: <https://www.bcrp.gob.pe/docs/Publicaciones/Revista-Moneda/moneda-173/moneda-173-05.pdf>
- Merma, J., & Marleny, P. (2017). Integración de procesos de negocio aplicando la arquitectura orientada a servicios (SOA). *Interfases(10)*, 113. Obtenido de <https://revistas.ulima.edu.pe/index.php/Interfases/article/view/1771>
- Nuño, P. (19 de Febrero de 2018). *¿Para qué sirven las Normas ISO?* Obtenido de *Emprende Pyme*: <https://www.emprendepyme.net/para-que-sirven-las-normas-iso.html>
- Rodriguez, F. (2014). *repositorio uptc*. Obtenido de repositorio uptc: <https://repositorio.upct.es/bitstream/handle/10317/4402/tfg482.pdf?sequence=1&%3BisAllowed=y>
- Rodríguez, I. (2020). *Metadatos estadísticos para el aseguramiento de la calidad de los datos [Tesis de doctorado, Universidad Nacional de Colombia]*. Repositorio Institucional. Obtenido de <https://repositorio.unal.edu.co/handle/unal/78418>
- Sánchez, J. (2004). *Diseño de base de datos*. https://www.emtelco.com.co/sites/default/files/2016-12/disenobd_0.pdf. Obtenido de emtelco: https://www.emtelco.com.co/sites/default/files/2016-12/disenobd_0.pdf
- SDi. (20 de Abril de 2017). *RESUMEN DE LAS UNIDADES I A LA IV*. Obtenido de sdioppolo: <https://sdioppolo.wordpress.com/2017/04/20/resumen-de-las-unidades-i-a-la-iv/>
- Soche, S. (2016). *Metodología para el modelamiento de datos basado en big data, enfocados al consumo de tráfico (voz-datos) generado por los clientes [Tesis de Especialización, Universidad Militar de Nueva Granada]*. Repositorio Institucional. Obtenido de <https://repository.unimilitar.edu.co/handle/10654/14950>

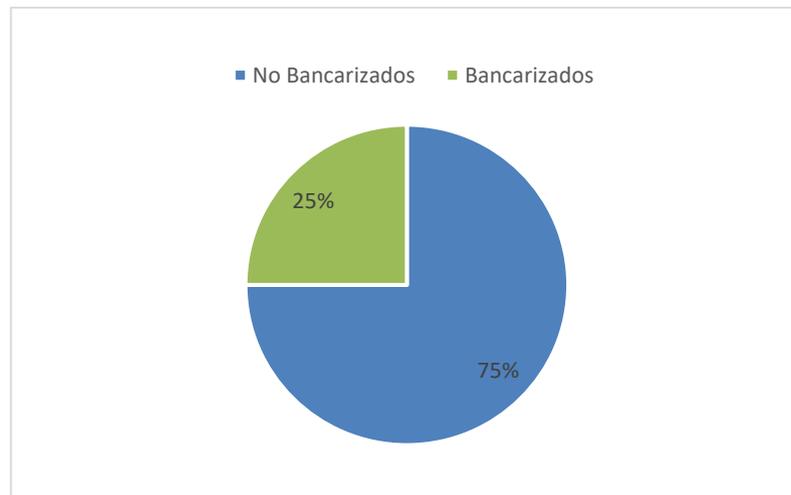
ANEXOS

Anexo A. Cantidad de Suscriptores a la Aplicación Móvil de la Moneda Digital



Nota: Se puede apreciar la cantidad de suscripciones que hubo en el transcurso de los años desde su lanzamiento en el año 2016. Tomado de *Entidad Financiera*.

Anexo B. Representación Gráfica de los 2 Tipos de Clientes



Nota: Como se ve puede apreciar se tiene que el 75% de los usuarios del Aplicativo Móvil de la Moneda Digital no tienen una cuenta Financiera. Tomado de *Entidad Financiera*.

Anexo C. Gráfico de Cantidad de Usuarios Nuevos en promedio por mes en el año 2020.



Nota: Se aprecia la suscripción en promedio por mes en el año 2020 a la Aplicación Móvil de la Moneda Digital. Tomado de *Entidad Financiera*.

Anexo D. Gráfico del Promedio de Transacciones al Mes en el Año 2020



Nota: Se aprecia que las transacciones al mes en promedio en el año 2020 fueron las más altas que tuvo el Aplicativo Móvil de la Moneda Digital. Tomado de *Entidad Financiera*.

Anexo E. Gráfico del Volumen de Dinero en Promedio en el Año

