

**UNIVERSIDAD NACIONAL DE SAN CRISTÓBAL DE HUAMANGA**

**FACULTAD DE INGENIERÍA DE MINAS, GEOLOGÍA Y CIVIL**

**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**



**“DATAMART PARA INDICADORES DE MOROSIDAD DEL SEDA AYACUCHO,  
2018”**

**Tesis presentado por** : Bach. Víctor Alvin Ayvar Risco  
**Para optar el título profesional de** : Ingeniero de Sistemas  
**Tipo de investigación** : Observacional, Retrospectiva y Transversal  
**Área de investigación** : Inteligencia de negocios  
**Asesor** : Dr. Efraín Elías Porras Flores

Ayacucho, Agosto de 2019

## **DEDICATORIA**

A Dios, por haberme permitido llegar hasta este punto y haberme dado salud para lograr mis objetivos, además de su infinita bondad y amor.

A mis padres y abuelos, por haberme apoyado en todo momento, por sus consejos, sus valores, por la motivación constante que me ha permitido ser una persona de bien, pero más que nada, por su amor.

## **AGRADECIMIENTO**

A Dios.

A mi familia.

A la Universidad Nacional San Cristóbal de Huamanga, por medio de ella a la Escuela de Ingeniería de Sistemas y por ende a sus dignos maestros.

A los trabajadores y funcionarios del SEDA Ayacucho, por colaborar en mi desarrollo profesional.

## **CONTENIDO**

DEDICATORIA	ii
AGRADECIMIENTO	iii
CONTENIDO	iv
RESUMEN	vii
INTRODUCCIÓN	viii

## **CAPÍTULO I**

### **PLANTEAMIENTO DE LA INVESTIGACIÓN**

1.1.	DIAGNÓSTICO Y ENUNCIADO DEL PROBLEMA.....	9
1.2.	DEFINICIÓN DEL PROBLEMA DE INVESTIGACIÓN.....	11
1.3.	OBJETIVOS DE LA INVESTIGACIÓN .....	11
1.4.	JUSTIFICACIÓN Y DELIMITACIÓN DE LA INVESTIGACIÓN.....	12
1.4.1.	JUSTIFICACIÓN.....	12
1.4.2.	DELIMITACIÓN .....	12

## **CAPÍTULO II**

### **REVISIÓN DE LA LITERATURA**

2.1.	ANTECEDENTES DE LA INVESTIGACIÓN .....	13
2.2.	MARCO TEÓRICO .....	14
2.2.1.	INFORMACION TÁCTICA .....	14
2.2.2.	INDICADOR .....	16
2.2.3.	DATAMART .....	19
2.2.5.	METODOLOGÍA HEFESTO.....	23
2.2.6.	PROCESAMIENTO ANÁLITICO EN LINEA OLAP (ON-LINE ANALYTICAL PROCESSING) .....	35
2.2.7.	SISTEMA GESTOR DE BASE DE DATOS (SGBD).....	40

2.2.8.	POBLACIÓN .....	43
2.2.9.	CENSO .....	43

### **CAPÍTULO III**

#### **METODOLOGÍA DE LA INVESTIGACIÓN**

3.1.	TIPO Y NIVEL DE LA INVESTIGACIÓN .....	44
3.2.	DISEÑO DE LA INVESTIGACIÓN .....	45
3.3.	POBLACIÓN Y MUESTRA .....	45
3.4.	VARIABLES E INDICADORES .....	45
3.4.1.	DEFINICIÓN CONCEPTUAL DE LAS VARIABLES.....	45
3.4.2.	DEFINICIÓN OPERACIONAL DE LAS VARIABLES .....	46
3.5.	TÉCNICAS E INSTRUMENTOS PARA RECOLECTAR INFORMACIÓN .....	46
3.5.1.	TÉCNICAS .....	46
3.5.2.	INSTRUMENTOS .....	46
3.6.	HERRAMIENTAS PARA EL TRATAMIENTO DE DATOS .....	46
3.7.	TÉCNICAS PARA APLICAR LA METODOLOGÍA HEFESTO .....	47

### **CAPÍTULO IV**

#### **RESULTADOS DE LA INVESTIGACIÓN**

4.1.	RESULTADOS APLICANDO EL MÉTODO HEFESTO .....	50
4.1.1.	ANÁLISIS DE REQUISITOS .....	50
4.1.2.	FASE DEL ANÁLISIS PROCESAMIENTO DE TRANSACCIONES EN LÍNEA (OLTP).....	53
4.1.3.	FASE DE MODELO LÓGICO DEL DATAMART .....	61
4.1.4.	FASES DEL PROCESO ETL.....	70
4.1.5.	IMPLEMENTACIÓN DE CUBOS MULTIDIMENSIONALES .....	79

**CAPÍTULO V**  
**RESULTADOS DE LA INVESTIGACIÓN**

5.1.	CONCLUSIONES.....	102
5.2.	RECOMENDACIONES.....	102
	<b>BIBLIOGRAFÍA</b>	103
	<b>ANEXOS</b>	106

## RESUMEN

El Departamento de Cobranza de la Gerencia Comercial del servicio de agua potable y alcantarillado de Ayacucho S.A cuenta con un software de gestión de usuarios y servicios denominado SIINCO, mediante el cual se realiza el registro y seguimiento de los usuarios que tienen el servicio, dicha información es utilizada por el departamento comercial para el análisis de los indicadores. Así mismo se maneja la información manual por parte del personal del área, el departamento de cobranza del Seda Ayacucho no tiene información táctica para la morosidad de clientes.

El estudio se realizó con los registros digitales del Seda Ayacucho, el tipo de investigación es observacional, retrospectivo y transversal con el nivel de investigación descriptivo.

Se implementó un Datamart usando la metodología HEFESTO; con los registros de la base de datos del SEDA Ayacucho, se construyó las tablas de hechos y dimensiones dispuestas en esquema estrella, se implementó sobre un gestor de base de datos relacional ; el desarrollo y automatización de los procesos ETL mediante Integration Services, la construcción e implementación de los cubos OLAP se realizó mediante Analysis Services, finalmente el acceso a los indicadores por parte de los usuarios finales, se realiza mediante una hoja de cálculo excel.

A partir de los cubos implementados los usuarios del sistema OLAP serán capaces de obtener información táctica y oportuna sobre; la morosidad del Seda Ayacucho. Además, la facilidad de realizar reportes personalizados orientados al análisis de datos y la toma de decisiones respecto a los indicadores definidos.

**Palabras clave:** Morosidad, Información táctica, cobranza, deuda, cliente, Datamart, HEFESTO.

## INTRODUCCIÓN

De acuerdo a Francois (2004), opina que la información táctica es aquella que muestra un aspecto panorámico y central de la realidad y su evolución, sirviendo para proyectar y prever futuras realidades. La morosidad según Castañeda y Tamayo (2013), es el incumplimiento de pagos retrasados que está preestablecido en un tiempo determinado.

La razón para realizar la investigación es para contribuir en la elaboración de estrategias y la disminución de la morosidad del Seda Ayacucho mediante la información táctica que será obtenida, así obtener indicadores de morosidad de los clientes que hacen uso de las prestaciones de servicio del Seda Ayacucho. Esto nos lleva a la pregunta ¿Qué información táctica es necesaria para los indicadores de la morosidad del Seda Ayacucho?. ¿Porque la morosidad de los clientes está considerada como un grupo de riesgo?.

Se utilizó la metodología Hefesto para la construcción del Datamart, en la cual se comienza como plantea Bernabeu (2010), recolectando las necesidades de información de los usuarios y se obtienen las preguntas claves del negocio. Luego, se identificó los indicadores resultantes de las preguntas y sus respectivas perspectivas de análisis, mediante las cuales se construyó el modelo conceptual de datos del Datamart.

Los objetivos específicos son: a) Generar indicadores sobre la cobranza con la finalidad de brindar información sobre los pagos fuera del plazo promedio de cobranza, porcentaje de eficiencia de cobranza y porcentaje de cobranza gestionada. b) Generar indicadores sobre la deuda con la finalidad de brindar información sobre el porcentaje de clientes gestionados por deuda, porcentaje de pagos efectuados y porcentaje de la deuda vencida. c) Generar indicadores sobre el cliente con la finalidad de brindar información sobre porcentaje de reclamos totales y porcentaje de conexiones activas.

# CAPÍTULO I

## PLANTEAMIENTO DE LA INVESTIGACIÓN

### 1.1. DIAGNÓSTICO Y ENUNCIADO DEL PROBLEMA

En la actualidad el Departamento de Cobranza de la Gerencia Comercial del SEDA Ayacucho cuenta con un software de gestión de usuarios y servicios denominado SIINCO, mediante el cual se realiza el registro y seguimiento de los usuarios que tienen el servicio, dicha información es usada en la Gerencia Comercial para la toma de decisiones, bajo la supervisión de la Gerencia General.

El Departamento de Cobranza se encarga de planear y ejecutar las políticas de cobranzas de los servicios de agua potable y alcantarillado, así como brindar una adecuada atención al cliente. Dichas políticas establecen cumplir con las metas de Eficiencia de Cobranza y disminución de la morosidad en forma sostenida.

A continuación, se muestra la eficiencia de cobranza máxima que fue 85.68%, la mínima 63.38%, en promedio se tiene 75.09%. Esta eficiencia no es la mejor, se debe mejorar el promedio.

Tabla 1.1. Escala única de facturación por año y mes

<b>CUADRO N.º 02 FACTURACIÓN, COBRANZA, EFICIENCIA DE COBRANZA</b>					
<b>MESES DEL AÑO 2018</b>	<b>FACTURACIÓN DEL MES (S/.)</b>	<b>DEVOLUCIÓN FACTURACIÓN (S/.)</b>	<b>FACTURACIÓN NETA</b>	<b>COBRANZA DEL MES (S/.)</b>	<b>EFICIENCIA DE COBRANZA (%.)</b>
ENERO	1,927,287.33	27,443.73	1,899,843.60	1,397,508.25	73.56%
FEBRERO	1,858,952.21	127,871.83	1,731,080.38	1,260,782.54	72.83%
MARZO	1,811,143.05	69,815.76	1,741,327.29	1,262,355.36	72.40%
ABRIL	1,674,807.61	31,736.70	1,643,070.91	1,316,240.75	80.11%
MAYO	2,058,663.67	53,208.07	2,005,455.60	1,510,170.15	75.30%
JUNIO	1,773,247.47	171,805.45	1,601,442.02	1,299,085.04	81.12%
JULIO	1,837,890.15	28,228.04	1,809,662.11	1,369,480.08	75.68%
AGOSTO	1,918,910.30	36,878.37	1,882,031.93	1,402,841.28	74.54%
SEPTIEMBRE	1,840,577.68	19,902.20	1,820,675.48	1,273,052.47	69.92%
TOTAL	16,701,479.47	566,890.15	16,134,589.32	12,091,515.92	74.94%

Fuente: SEDA Ayacucho, 2018.

Tabla 1.2. Esca única de facturación por año y mes

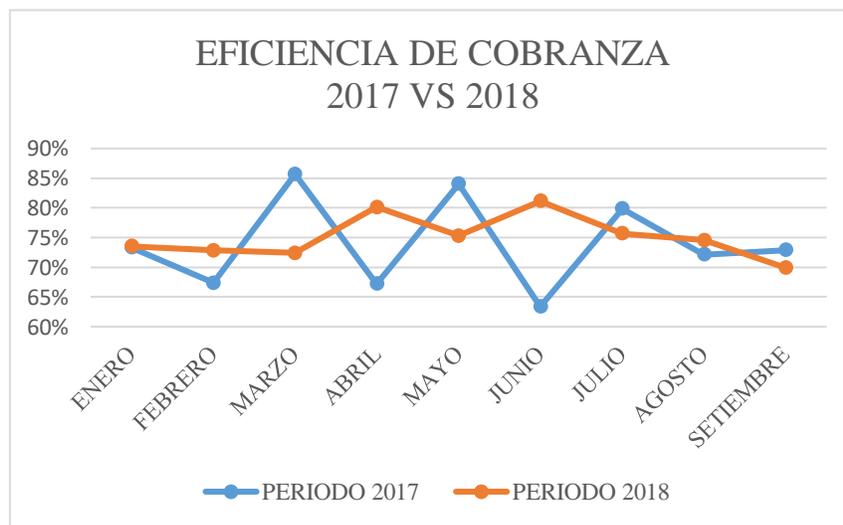
CUADRO N.º 02 FACTURACIÓN, COBRANZA, EFICIENCIA DE COBRANZA					
MESES DEL AÑO 2017	FACTURACIÓN DEL MES (S/.)	DEVOLUCIÓN FACTURACIÓN (S/.)	FACTURACIÓN NETA	COBRANZA DEL MES (S/.)	EFICIENCIA DE COBRANZA (%.)
ENERO	1,763,922.07	67,130.57	1,696,791.50	1,244,617.85	73.35%
FEBRERO	1,723,689.56	25,089.34	1,698,600.22	1,144,464.03	67.38%
MARZO	1,693,421.72	52,842.38	1,640,579.34	1,105,626.15	85.68%
ABRIL	1,573,979.64	47,282.83	1,526,696.81	1,027,636.65	67.21%
MAYO	1,805,136.59	40,791.46	1,764,345.13	1,482,897.01	84.05%
JUNIO	1,725,087.61	26,903.42	1,698,184.19	1,161,261.18	63.38%
JULIO	1,747,760.44	36,872.23	1,710,888.21	1,366,200.03	79.85%
AGOSTO	1,771,571.29	68,769.68	1,702,801.61	1,228,497.35	72.15%
SETIEMBRE	1,782,077.70	25,563.48	1,755,514.22	1,279,889.42	72.91%
OCTUBRE	1,783,073.08	38,009.73	1,745,063.35	1,374,252.65	78.75%
NOVIEMBRE	1,759,276.15	34,066.02	1,725,210.13	1,355,697.15	78.58%
DICIEMBRE	1,868,709.61	25,570.46	1,843,139.15	1,434,281.66	77.82%
TOTAL	20,997,705.46	1,566,891.33	19,430,814.13	15,505,321.13	79.79%

Fuente: SEDA Ayacucho, 2017

Según la tabla N° 1.1, hasta el mes de diciembre del año 2017 se cuenta con una facturación neta de s/. 20,507,813.86 lo cual se llegó a pagar s/. 15,505,321.13, dejando una diferencia de s/. 5,002,492.73 que significa una eficiencia del 75.61% con una morosidad de 24.39%, esto implica hacer una evaluación en las políticas de cobranza ya que se genera perdidas económicas significativas en la empresa.

A continuación se visualiza la eficiencia de cobranza del año 2017 – 2018.

Tabla 1.3. Eficiencia de Cobranza del año 2017 - 2018



Fuente: Seda Ayacucho, 2018.

Según la tabla 1.1 y 1.2, en el mes de setiembre del año 2017 se llegó a cobrar s/.1,279,889.49 de s/.1,755,514.22 y en el año 2018 se cobró s/.1,273,052.47 de s/.1,820,675.48, esto significa que la morosidad llega a un 28.59% promedio lo cual es una pérdida significativa para la empresa que afecta al Departamento Comercial.

## **1.2. DEFINICIÓN DEL PROBLEMA DE INVESTIGACIÓN**

### **PROBLEMA PRINCIPAL**

¿Qué información táctica es necesaria para los indicadores de la morosidad de la empresa de servicio de agua potable y alcantarillado de Ayacucho S.A, 2018?

### **PROBLEMAS SECUNDARIOS**

¿Qué información táctica es necesaria para la cobranza, deuda y cliente?

## **1.3. OBJETIVOS DE LA INVESTIGACIÓN**

### **OBJETIVO GENERAL**

Implementar un Datamart para brindar información táctica mediante técnicas e instrumentos, la tecnología OLAP, la metodología Hefesto, un gestor de base de datos relacional, con la finalidad de brindar indicadores sobre la morosidad de la empresa de servicio de agua potable y alcantarillado de Ayacucho, 2018.

### **OBJETIVOS ESPECÍFICOS**

- a. Generar indicadores sobre la cobranza con la finalidad de brindar información sobre los pagos fuera del plazo promedio de cobranza, porcentaje de eficiencia de cobranza y porcentaje de cobranza gestionada.
- b. Generar indicadores sobre la deuda con la finalidad de brindar información sobre el porcentaje de clientes gestionados por deuda, porcentaje de pagos efectuados y porcentaje de la deuda vencida.
- c. Generar indicadores sobre el cliente con la finalidad de brindar información sobre porcentaje de reclamos totales y porcentaje de conexiones activas.

## **1.4. JUSTIFICACIÓN Y DELIMITACIÓN DE LA INVESTIGACIÓN**

### **1.4.1. JUSTIFICACIÓN**

Es necesario implementar un Datamart que permita integrar los datos de la empresa en el área comercial del Seda Ayacucho. La presente investigación pretende brindar una información táctica en base al análisis de los indicadores y de esta manera tomar estrategias para poder alcanzar las metas.

La utilidad metodológica de la investigación servirá de base para otros estudios que tengan como referencia el presente trabajo y realicen otros estudios para otras empresas servidoras del servicio de agua alcantarillado.

La relevancia social radica en que se identificaran los indicadores de la morosidad, indicadores que involucren el comportamiento de la deuda vencida del Seda Ayacucho. Así como las características que definan el desbalance financiero de la empresa. Estos resultados ayudarán a reducir la morosidad en base a la información táctica obtenida.

El área comercial del Seda Ayacucho, necesita tener información confiable y oportuna, para poder adoptar estrategias en base al análisis de los indicadores y de esta manera alcanzar las metas preestablecidas.

### **1.4.2. DELIMITACIÓN**

La investigación se realizó en la Región Ayacucho, en la gerencia comercial de la empresa de servicio de agua potable y alcantarillado de Ayacucho S.A, abarcando los importes de facturación de todos los clientes del año 2018.

## **CAPÍTULO II**

### **REVISIÓN DE LITERATURA**

#### **2.1. ANTECEDENTES DE LA INVESTIGACIÓN**

Morales y Vargas (2017) en su tesis que plantea identificar los factores externos y su influencia en los índices de morosidad en una empresa comercializadora de productos de electricidad, afirman que existen factores internos y externos que generan problemas financieros en una empresa y con ello retrasos o incumplimientos de pago.

Castañeda y Tamayo (2013) en su investigación sobre la morosidad y su impacto en el cumplimiento de los objetivos estratégicos de la agencia real plaza de la municipal de ahorro y crédito de Trujillo, sostienen que los indicadores de morosidad de los principales tipos de crédito, se han visto sustancialmente incrementados afectando el cumplimiento de los objetivos estratégicos principalmente los relacionados a la calidad de cartera.

Cervera y López (2017) en su tesis análisis y determinantes de la morosidad de los clientes de la caja municipal Sullana, afirman que las variables de la morosidad dependen de la cultura crediticia, volumen de pago y capacidad de pago ya sea por problemas de salud o créditos externos.

Ruiz (2018), en su tesis sobre la morosidad en el pago de arbitrios en la municipalidad distrital de lince, afirma que el desbalance tributario viene a partir de una cultura arraigada en nuestra sociedad que es la de no cumplir con las obligaciones del estado, este hecho catalogado como morosidad tributaria conlleva a una mala gestión tributaria.

Gómez (2008), en su tesis propuesta de estrategias de cobranza para disminuir la morosidad en la empresa electro oriente S.A, afirma que existen empresas que ofrecen algún servicio y no cuentan con políticas adecuadas para atacar la morosidad, este hecho provoca levantamientos masivos y se debe tomar en cuenta la intervención integral de alternativas

viables que permitan afrontar el desbalance de la empresa así como la disminución de la morosidad.

Rojas (2015), en su tesis sobre la gestión de la morosidad y su relación con el apalancamiento financiero de la empresa grupo unión wuanca S.A.C, concluye que la falta de pago de las facturas comerciales es la principal causa de la morosidad lo cual conlleva a una mala gestión en la aplicación de políticas comerciales.

## **2.2. MARCO TEÓRICO**

### **2.2.1. INFORMACION TÁCTICA**

“Abarca periodos de tiempo relativamente cortos; se emplea en los niveles gerenciales medios para instrumentar programas de planeación estratégicas y planes específicos para las áreas funcionales de la organización. Se centran alrededor de los lineamientos de los planes subordinados, necesarios para instrumentar una estrategia en particular y después en el mantenimiento y control del rendimiento real contra los planes definidos. Está orientado principalmente a soportar la toma de decisiones que se asocia a gerencias o subdirecciones para alcanzar la misión empresarial; este tipo de información es extraída específicamente de un área o departamento de la organización, por lo que su alcance es local” (Thierauf, 1994, p.29).

De acuerdo a Francois (2004), la información táctica es aquella que muestra un aspecto panorámico y central de la realidad y su evolución, sirviendo para proyectar y prever futuras realidades, como las estadísticas, los análisis históricos, los informes y documentos de situación política, social, económica y cultural. Con esta información y la base ideológica se contruyeron los modelos para las diversas temáticas.

La gerencia intermedia necesita información operacional que le ayude con la actividad de supervisión, control, toma de decisiones, administración y desempeño actual de la organización y son responsables de encontrar las mejores medidas de ejecutar las decisiones estratégicas de sus superiores (Effy, 2001, p.342).

## **A. COBRANZA**

“La política crediticia así como los reglamentos de crédito y la gestión en las cobranzas inciden en los niveles de la morosidad de las empresas” (Ayuque, 2010).

“La cartera sana debe estar conformada por una gran diversidad de clientes, plazos de cobranza reducidos, bajo índice de morosidad e incobrabilidad y políticas de cobranza explícitas y claras” (Córdoba, 2012, p.342).

Según Zapata (2011), se origina cuando se tiene una cuenta por cobrar por ventas de bienes o servicios a un cliente. Sintetizando, las cuentas por cobrar son el derecho para percibir efectivo en un futuro, a partir de una transacción hecha en la actualidad. Es algo que presenta todo negocio, por tanto es un activo. Las transacciones de cuentas por cobrar involucran dos partes: El acreedor, quien adquiere la cuenta por cobrar registrándose un parámetro expresado en días y es controlado desde cualquier departamento financiero de una empresa, pues en él se basa la gestión de tesorería de la misma, midiendo así la cantidad de días empleados en cada cobranza efectuada.

## **B. DEUDA**

Según el Ministerio de Economía y Finanzas (2018), la deuda es un conjunto de obligaciones de pago que mantiene el usuario, a una determinada fecha, frente a sus acreedores.

Para Hernandez (2009), el departamento comercial cuenta con dos herramientas de suma importancia: El ciclo promedio de cobro y el reporte de la antigüedad de las deudas por cobrar. Estas variables son controlables y cuantitativos cuyo fin es mostrar estados positivos o negativos de la empresa.

Según la Superintendencia De Banca, Seguros y AFP (2018), clasifica a los deudores de acuerdo a los días de atraso en cinco categorías, definida por el flujo de fondos y el grado de cumplimiento de sus obligaciones. Estas son: a) Normal, b) Problemas potenciales, c) Deficiente, d) Dudoso, e) Pérdida.

Tabla 2.1. Clasificación del deudor

CATEGORÍAS DE CLASIFICACIÓN	EN CRÉDITO DE CONSUMO*	EN CRÉDITOS HIPOTECARIOS
0 : Normal	Pago puntual o atraso máximo de 8 días calendario	Pago puntual o atraso máximo de 30 días calendarios
1 : Problemas Potenciales	Atrasos en el pago de entre 9 a 30 días calendarios	Atrasos en el pago de entre 31 a 60 días calendarios
2 : Deficiente	Atrasos en el pago de entre 31 a 60 días calendario	Atrasos en el pago de entre 61 a 120 días calendarios
3 : Dudoso	Atrasos en el pago de entre 61 a 120 días calendarios	Atrasos en el pago de entre 121 a 365 días calendarios
4 : Pérdida	Atrasos en el pago de más de 120 días calendarios	Atrasos en el pago de más de 365 días calendarios

Fuente: Superintendencia de Banca y Seguros, 2018

### C. CLIENTE

Según Kotler (2016), los clientes gestionados o activos dentro de una empresa satisfacen las actividades y servicios que está les brinda dentro de un marco legal y políticas empresariales.

Según la Real Academia Española (2016), los clientes gestionados son consumidores que compran productos o adquieren servicios para satisfacer una necesidad propia.

Según Cruz (2011), la gestión de nuevos clientes puede costar cinco veces más que satisfacer y retener a los clientes actuales. Dicha gestión se basa en los intereses actuales de los clientes, es decir las demandas que estos presenten.

#### 2.2.2. INDICADOR

“Herramientas para clasificar y definir, de forma más precisa, objetivos e impactos (...) son medidas verificables de cambio o resultado (...) diseñadas para contar con un estándar contra el cual evaluar, estimar o demostrar el progreso (...) con respecto a metas establecidas, facilitan el reparto de insumos, produciendo (...) productos y alcanzando objetivos” (ONU, 1999).

“Herramienta cuantitativa o cualitativa que permite mostrar indicios o señales de una situación, actividad o resultado” (CONEVAL, 2014).

**A. INDICADORES DE LA COBRANZA**

**a. N ° DE PAGOS FUERA DEL PLAZO PROMEDIO DE COBRANZAS**

De acuerdo al Seda Ayacucho(2006), identifica el pago fuera del plazo promedio de cobranzas.

$$\frac{\text{pagos fuera de fecha}}{\text{total de pagos}} \times 100$$

Figura 2.1. Formula del indicador N° de pagos fuera del plazo promedio de cobranzas(SEDA, 2006)

**b. PORCENTAJE DE EFICIENCIA DE COBRANZA**

De acuerdo al Seda Ayacucho (2006), Identifica el porcentaje de eficiencia de cobranza.

$$\frac{\text{Cobranza efectuada}}{\text{Facturacion neta}} \times 100$$

Figura 2.2. Formula del indicador porcentaje de eficiencia de cobranza (SEDA, 2006)

**c. PORCENTAJE DE COBRANZA GESTIONADA**

De acuerdo al Seda Ayacucho (2006), Identifica el porcentaje de cobranza gestionada.

$$\frac{\text{Nº de facturas no canceladas}}{\text{Nº de Facturas emitidas}} \times 100$$

Figura 2.3. Formula del indicador porcentaje de cobranza gestionada (SEDA, 2006)

**B. INDICADORES DE LA DEUDA**

**a. PORCENTAJE DE CLIENTES GESTIONADOS POR DEUDA**

De acuerdo al SEDA (2006), Identifica el porcentaje de clientes gestionados por

deuda .

$$\frac{N^{\circ} \text{ de Facturas gestionadas}}{N^{\circ} \text{ de Facturas afecto a deuda}} \times 100$$

Figura 2.4. Formula del indicador porcentaje de clientes gestionados por deuda (SEDA, 2006)

**b. PORCENTAJE DE PAGOS EFECTUADOS**

De acuerdo al SEDA (2006), Identifica el porcentaje de clientes gestionados por deuda .

$$\frac{N^{\circ} \text{ de Facturas canceladas}}{N^{\circ} \text{ de clientes}} \times 100$$

Figura 2.5. Formula del indicador porcentaje de pagos efectuados (SEDA, 2006)

**c. PORCENTAJE DE LA DEUDA VENCIDA**

De acuerdo al SEDA (2006), Identifica el porcentaje de clientes gestionados por deuda .

$$\frac{\text{Total de la deuda vencida}}{\text{Total de la deuda}} \times 100$$

Figura 2.6. Formula del indicador porcentaje de la deuda vencida (SEDA, 2006)

**C. INDICADORES DEL CLIENTE**

**a. PORCENTAJE DE RECLAMOS TOTALES**

De acuerdo a la Superintendencia Nacional de Servicios y Saneamiento (2006), Identifica el porcentaje de clientes gestionados por deuda .

$$\frac{N^{\circ} \text{ de reclamos totales}}{N^{\circ} \text{ de conexiones totales activas}} \times 100$$

Figura 2.7. Formula del indicador porcentaje de reclamos totales (SUNASS, 2006)

**b. PORCENTAJE DE CONEXIONES ACTIVAS**

De acuerdo al SEDA (2006), Identifica el porcentaje de conexiones activas .

$$\frac{N^{\circ} \text{ de conexiones activas}}{N^{\circ} \text{ de conexiones totales}} \times 100$$

Figura 2.8. Formula del indicador porcentaje de conexiones activas (SEDA, 2006)

**2.2.3. DATAMART**

Para Conesa y Curto (2011), un Datamart es un subconjunto de datos del Data Warehouse cuyo objetivo es responder a una necesidad, con una población de usuaria específica. Los datos están estructurados en modelo estrella o copo de nieve. Un Datamart puede ser independiente o dependiente de un Data Warehouse.

Según Gauchet (2015), define que un Datamart es un conjunto de datos aislados de los sistemas de gestión, dedicados a dar soporte a la toma de decisiones, sus datos se centran según diversos ejes, entre ellas un eje temporal cuya herramienta principal es la encapsulación de datos.

“Un Datamart es un subconjunto de datos derivado del Data Warehouse. Está diseñado para soportar requerimientos analíticos específicos de una determinada unidad de negocios. Es un repositorio menos ambicioso que un DWH” (Moss y Atre, 2003).

Vitt y Luckevich (2002) Se denomina Datamart al conjunto de datos estructurados que provienen de las diferentes aplicaciones operacionales. Es un subconjunto de un Datawarehouse con un alcance de contenido limitado, el cual es usado sólo por un área específica o un problema particular de análisis dentro de la organización.

Según Nagabhushana (2006), un Datamart puede ser considerada como una porción del Data Warehouse para un área específica de la empresa cuya implementación se da por etapas y comienza utilizando una estrategia analítica.

Se observa la clasificación de los datamart de acuerdo a Inmon (2002) en:

### **DATAMART DEPENDIENTE**

Los Datamart dependientes son aquellos que reciben los datos desde una Data warehouse. En este tipo de Datamart la fuente de los datos es única.

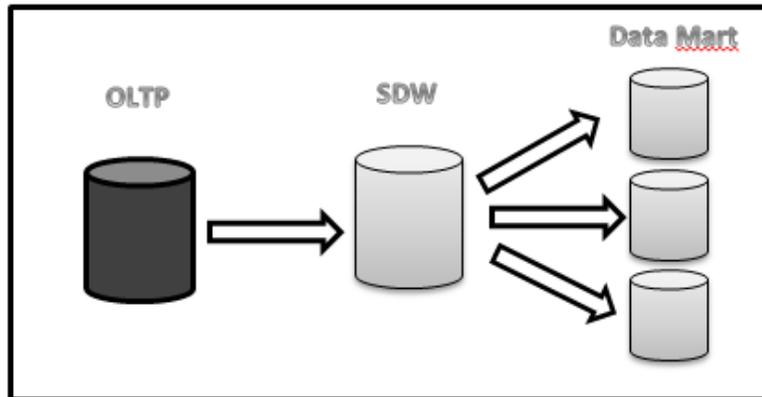


Figura 2.9. Datamart dependiente (Inmon, 2002).

### **DATAMART INDEPENDIENTE**

Son aquellos que toman sus datos directamente desde los sistemas transaccionales y no dependen de otros Data warehouse. Este tipo de Datamart se alimenta generalmente de las organizaciones.

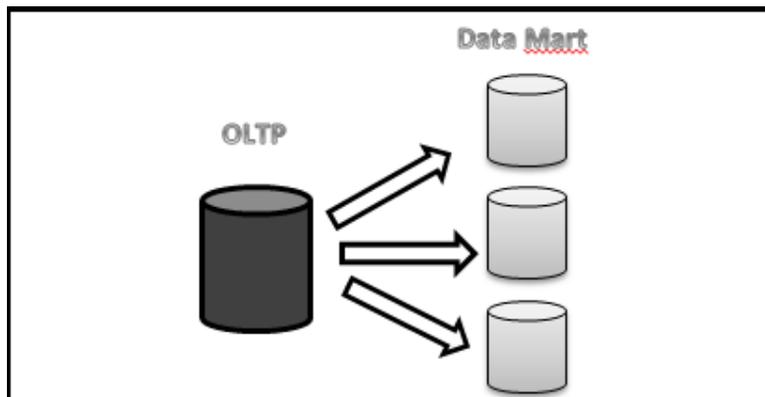


Figura 2.10. Data Mart independiente (Inmon, 2002).

### **DATAMART HÍBRIDO**

Los Datamart híbridos permiten combinar las fuentes de datos de un Data Warehouse corporativo con otras fuentes de datos tales como sistemas transaccionales y/o operacionales.

### 2.2.3.1 VENTAJAS Y DESVENTAJAS

**VENTAJAS.** Para Nader (2002), los principales beneficios de utilizar datamarts son:

- a) Dado que un datamarts soporta menos usuarios que un data Warehouse se puede optimizar para recuperación más rápida los datos que necesitan los usuarios.
- b) Menores cantidades de datos implica que se procesan antes, tanto las cargas de datos como las consultas.
- c) Las peticiones pueden acotarse al área o red sirve esos datos, sin afectar al resto de los usuarios.
- d) Las aplicaciones cliente, pide las consultas es independiente del servidor que la procesa y del servidor de bases de datos que almacenan la información.
- e) Los costos que implica la construcción de un datamarts son mucho menos a los de la implementación de un data Warehouse.

**DESVENTAJAS.** “No permite el manejo de grandes volúmenes de información por lo que muchas veces se debe recurrir a un conjunto de datamarts para cubrir todas las necesidades de información de la empresa” (Vizuete y Yela, 2006).

### 2.2.4. DATA WAREHOUSE

Según IBM (1999), un data warehouse es un repositorio no volátil de datos, transacciones y eventos. El cual incluye datos corporativos, operacionales y externos.

Conesa y Curto (2011), afirman que un data warehouse es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización. Al abarcar un ámbito global y con un amplio alcance histórico, el volumen de datos puede ser muy grande, presenta las siguientes características: a) orientada a un tema, b) integrado, c) variantes en el tiempo y d) no volátil.

Según Efor Internet y Tecnología (s.f), un Data warehouse se trata de un almacenamiento separado y homogeo donde son cargados los datos transformados previamente de distintas bases de datos, es decir cargamos toda la información necesaria a un único sitio para poder

utilizarla.

Según Inmon (2007), afirma que el Data Warehouse es un modelado y análisis de datos para los tomadores de decisiones, que son útiles en el apoyo y soporte, el cual garantiza la confiabilidad y provee operaciones lógicas estructurales de los datos procesados a nivel organizacional.

“Conjunto de tecnologías de soporte a la toma de decisión, cuyo objeto es que quien trabaja con los conocimientos (ejecutivo, director, analista) pueda tomar decisiones de manera más rápida y eficaz” (Chaudhuri y Dayal, 1997).

Según Ralph Kimball (2007), un Data Warehouse es una copia de los datos transaccionales específicamente estructurada para la consulta y el análisis en el procesamiento de la información.

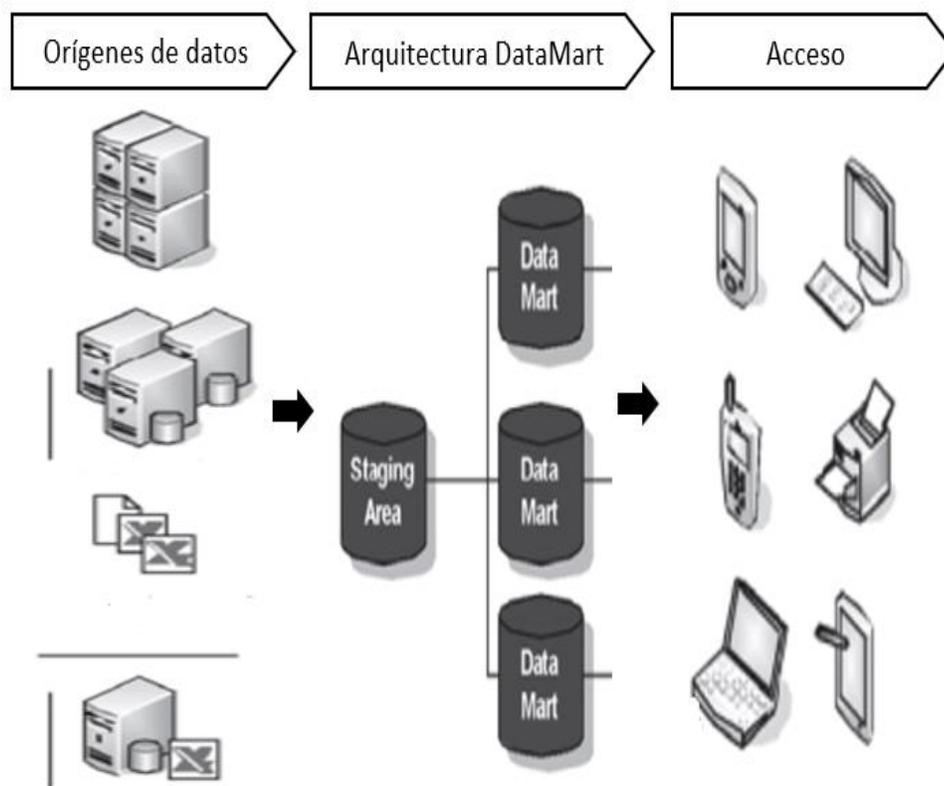


Figura 2.11. Arquitectura multidimensional (Conesa y Curto, 2011)

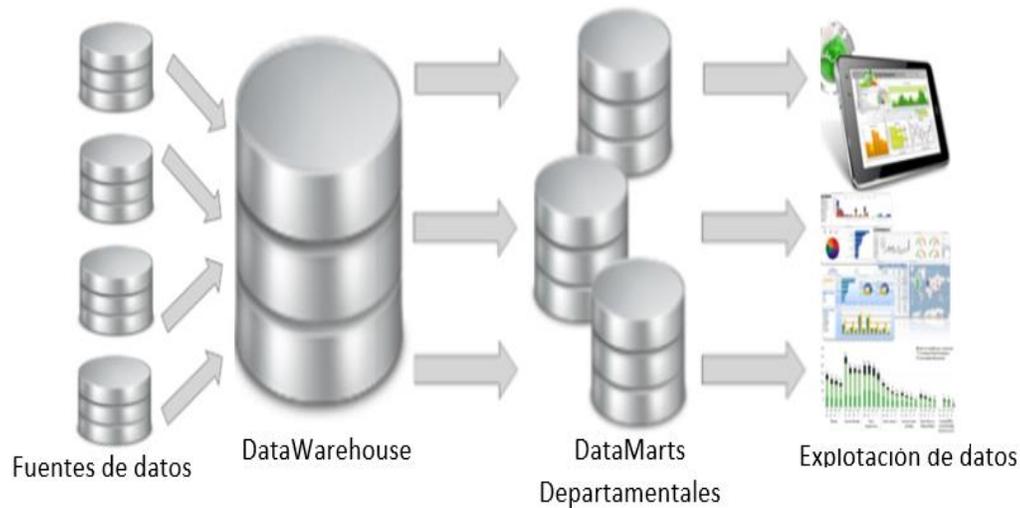


Figura 2.12. Arquitectura del DataWarehouse (Inmon, 2010)

### 2.2.5. METODOLOGÍA HEFESTO

De acuerdo a Bernabeu (2010), Hefesto es una metodología propia, cuya propuesta está fundamentada en una muy amplia investigación, comparación de metodologías existentes, experiencias propias en procesos de confeccion de almacenes de datos. Cabe destacar que HEFESTO está en continua evolución y se han tenido en cuenta, como gran valor agregado, todos los feedbacks que han aportado, quienes han utilizado esta metodología en diversos países y con diversos fines. La idea principal es comprender cada paso que se realizará, para no caer en el tedio de tener que seguir un método al pie de la letra sin saber exactamente qué se está haciendo, ni por qué. La construcción e implementación de un Data Warehouse puede adaptarse muy bien a cualquier ciclo de vida de desarrollo de software, con la salvedad de que para algunas fases en particular, las acciones que se han de realizar serán muy diferentes. Lo que se debe tener muy en cuenta, es no entrar en la utilización de metodologías que requieran fases extensas de reunión de requerimientos y análisis, fases de desarrollo monolítico que conlleve demasiado tiempo y fases de despliegue muy largas. Lo que se busca, es entregar una primera implementación que satisfaga una parte de las necesidades para demostrar las ventajas del Data Warehouse y motivar a los usuarios, la metodología Hefesto sugiere distribuir el proceso en cuatro fases importantes cuya finalidad es brindar la estructura respecto a la construcción del DataMart cuyo fin será brindar información analisada en base a la información procesada, estas son:

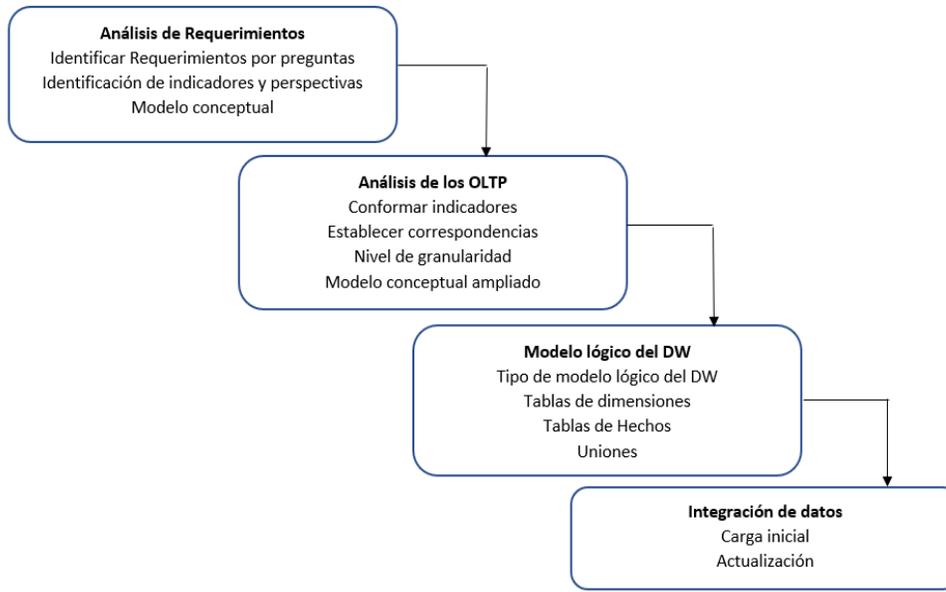


Figura 2.13. Metodología HEFESTO, pasos (Bernabéu, 2010)

## CARACTERÍSTICAS

Esta metodología tiene las siguientes características según (Bernabéu, 2010):

- a) Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.
- b) Se basa en los requerimientos del usuario, por lo cual su estructura es capaz de adaptarse con facilidad y rapidez ante los cambios en el negocio.
- c) Reduce la resistencia al cambio, ya que involucra al usuario final en cada etapa para que tome decisiones respecto al comportamiento y funciones del DW. Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.
- d) Es independiente del tipo de ciclo de vida que se emplee para contener la metodología. Es independiente de las herramientas que se utilicen para su implementación.
- e) Es independiente de las estructuras físicas que contengan el DW y de su respectiva distribución.
- f) Cuando se culmina con una fase, los resultados obtenidos se convierten en el punto de partida para llevar a cabo el paso siguiente.
- g) Se aplica tanto para DM como para DW.
- h) Permite obtener una visión sencilla en el proceso de transformación de la información ya establecida.

### 2.2.5.1 ANÁLISIS DE REQUERIMIENTOS

Según Bernabéu (2010), menciona que el análisis de requisitos se conforma de tres actividades principales: a) La identificación de preguntas, b) La identificación de indicadores y perspectivas y c) Un modelo conceptual. La primera actividad, identificación de preguntas, busca captar las necesidades de información en la organización de acuerdo con sus metas y estrategias haciendo uso de diferentes técnicas como encuestas, cuestionarios, observaciones, etc. El objetivo de esta actividad es identificar las necesidades de información en un alto nivel. Una vez descubiertas las necesidades de información de la organización, se procede a su análisis para obtener los indicadores y las perspectivas que se tomarán en cuenta durante la implementación del Data Mart. Los indicadores son, generalmente, valores numéricos que se desean analizar y las perspectivas son los puntos de vista sobre los cuales se examinan los indicadores para responder las preguntas analizadas durante la primera actividad, finalmente, el modelo conceptual permitirá reunir la información hasta el momento obtenida y presentará un primer bosquejo de alto nivel del alcance del proyecto. Este modelo se representa a través de un gráfico que consta de tres partes, la primera son las perspectivas sobre las cuales se realizará el análisis, en la posición central se grafica la relación que tienen todas estas perspectivas y finalmente como resultado de esta relación se miden los indicadores resultantes como se aprecia en la siguiente ilustración.

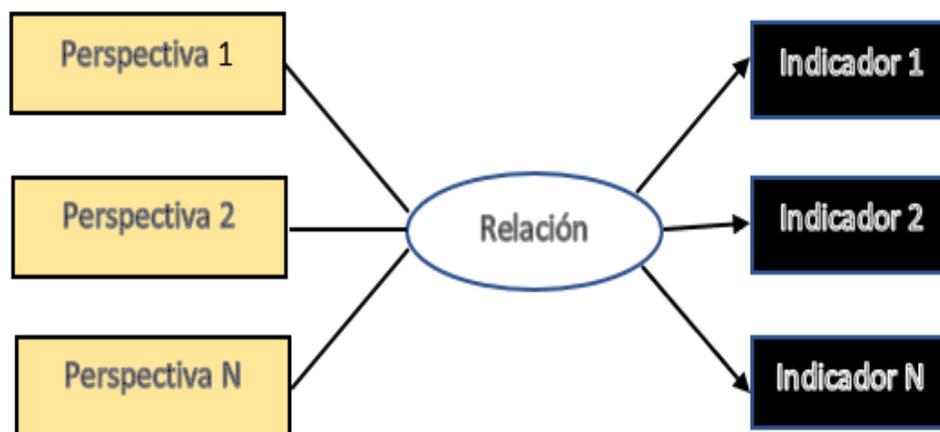


Figura 2.14. Modelo Conceptual (Bernabéu, 2010)

“Unidades vendidas de cada producto a cada cliente en un tiempo determinado”  
**INDICADOR** **PERSPECTIVAS**

“**Monto total de ventas** de cada **producto** a cada **cliente** en un **tiempo** determinado”

**INDICADOR**
**PERSPECTIVAS**

Figura 2.15. Indicadores y perspectivas (Bernabéu, 2010)

### 2.2.5.2 ANÁLISIS DE LOS PROCEDIMIENTOS DE TRANSACCIONES EN LÍNEA (OLTP)

Según Bernabéu (2010), menciona que en el análisis de los OLTP se realizará el análisis de las fuentes de datos de las cuales se proveerá de datos al modelo planteado en el punto anterior. Para esto se llevan a cabo las siguientes actividades: a) Conformar indicadores, b) establecer correspondencias, c) Definir el nivel de granularidad y finalmente la d) Creación del modelo conceptual ampliado. Para comenzar, en la actividad conformar indicadores, se define las funciones de cálculo para los indicadores. Los indicadores están compuestos de hechos y una función de sumariación para agregar los resultados. Seguido, se realiza la correspondencia de los datos requeridos por el Data Mart y con los datos que existen en el OLTP, esto facilitará su próxima extracción y asegura que se cuente con los datos requeridos para la construcción del Data Mart. A continuación, se definirán los campos que tendrán cada perspectiva y el nivel de detalle (granularidad) con el que los datos serán importados al Data Mart. Finalmente obtenemos un modelo conceptual ampliado que incluye los campos de las perspectivas y las fórmulas de cálculo de cada indicador definido.

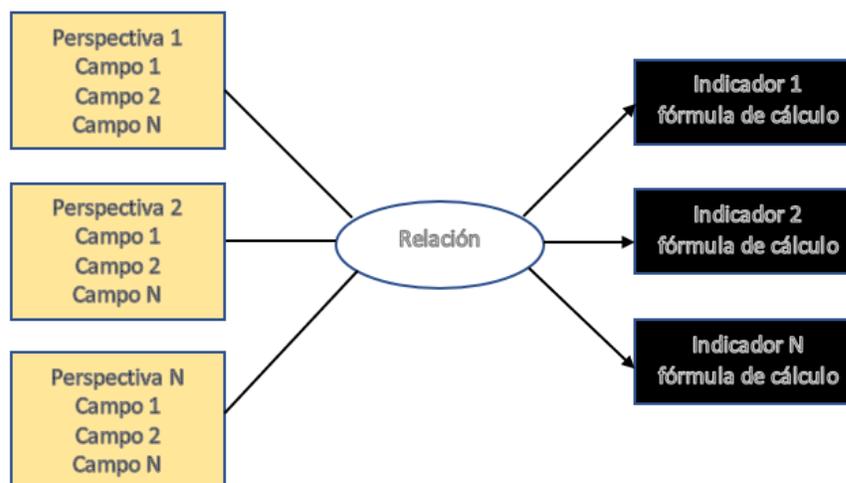


Figura 2.16. Modelo Conceptual ampliado (Bernabéu, 2010)

“El proceso transaccional en línea OLTP (on-line Transactional Processing), es un tipo de proceso especialmente rápido en el que las solicitudes de los usuarios son resueltas; naturalmente, ello implica la concurrencia de un “mecanismo” que permite el procesamiento de varias transacciones a la vez” (Torres, 2007).

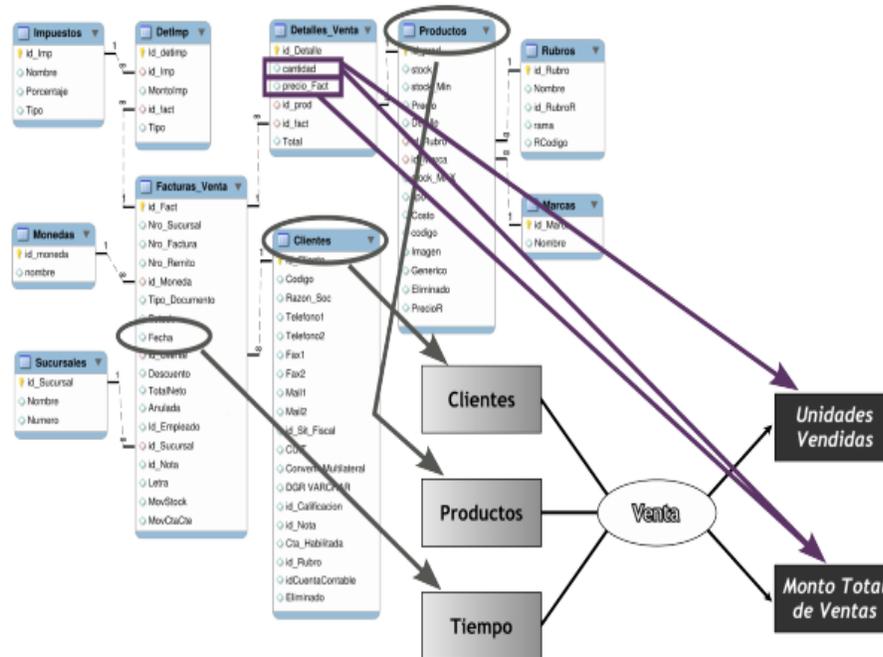


Figura 2.17. Caso práctico, Correspondencia (Bernabéu, 2010).

“Los sistemas transaccionales, cuyas siglas en inglés son OLTP (Online Transaction Process), son los sistemas que soportan las transacciones realizadas en una organización, los cuales se caracterizan porque son los encargados de soportar cada operación realizada en el negocio, por ello deben dar respuesta rápida a cada transacción realizada, además se encargan de recolectar los datos asociados a cada transacción y almacenar estos datos en bases de datos a través de sistemas de información. Si un sistema OLTP deja de funcionar puede ocasionar grandes pérdidas en una compañía por el hecho de no poder realizar transacciones” (Kimball, 2013).

Bernabéu (2009) afirma que los sistemas OLTP representa toda aquella información transaccional que genera la empresa en su accionar diario, además de las fuentes externas con la que puede llegar a disponer, estas fuentes de información son de características muy disímiles entre sí, en formato, procedencia, función, etc, entre los OLTP más habituales que

puedan existir en cualquier organización se encuentran los archivos de texto, hipertextos, hojas de cálculo, informes semanales.

### 2.2.5.3 MODELO LÓGICO DEL DATA WAREHOUSE

Bernabéu (2010) afirma que el modelo lógico del Data Warehouse se confeccionará de la estructura de DW, teniendo como base el modelo conceptual que ya ha sido creado. Para ello, primero se definirá el tipo de modelo que se utilizará y luego se llevarán a cabo las acciones propias al caso, para diseñar las tablas de dimensiones y de hechos. Finalmente, se realizarán las uniones pertinentes entre estas tablas. El modelo lógico del Data Warehouse consta de a) Tipo de Modelo lógico del DW. Se debe seleccionar cuál será el tipo de esquema que se utilizará para contener la estructura del depósito de datos, que se adapte mejor a los requerimientos y necesidades de los usuarios. Es muy importante definir objetivamente si se empleará un esquema en estrella, constelación o copo de nieve, ya que esta decisión afectará considerablemente la elaboración del modelo lógico; b) Tablas de dimensiones. Se elegirá un nombre que identifique la tabla de dimensión, se añadirá un campo que represente su clave principal, se redefinirán los nombres de los campos si es que no son lo suficientemente intuitivos; c) Tablas de hecho. En este paso, se definirán las tablas de hechos, que son las que contendrán los hechos a través de los cuales se construirán los indicadores de estudio; d) Uniones. Para los tres tipos de esquemas, se realizarán las uniones correspondientes entre sus tablas de dimensiones y sus tablas de hechos. Los nombres de los campos si es que no son lo suficientemente intuitivos.

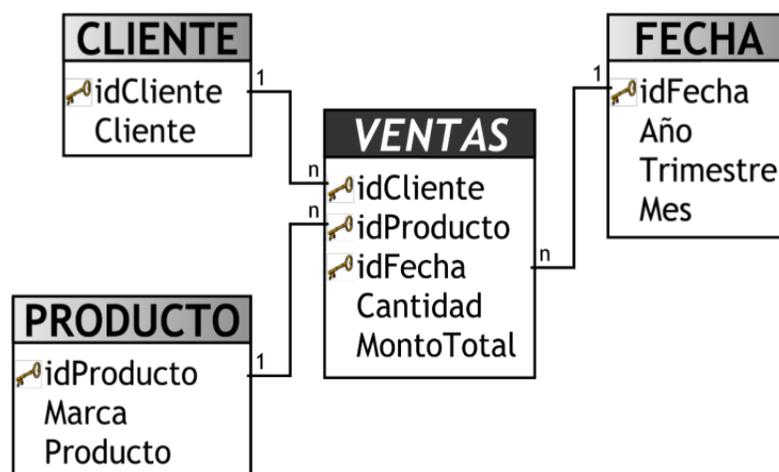


Figura 2.18. Uniones(Bernabéu, 2010).

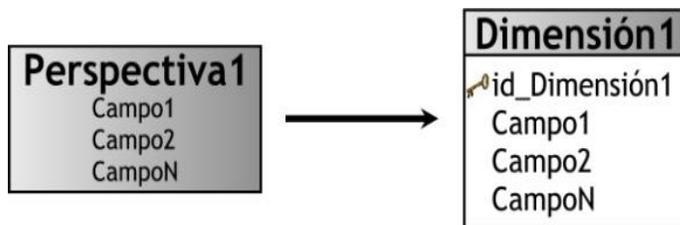


Figura 2.19. Tablas de dimensiones (Bernabéu, 2010).

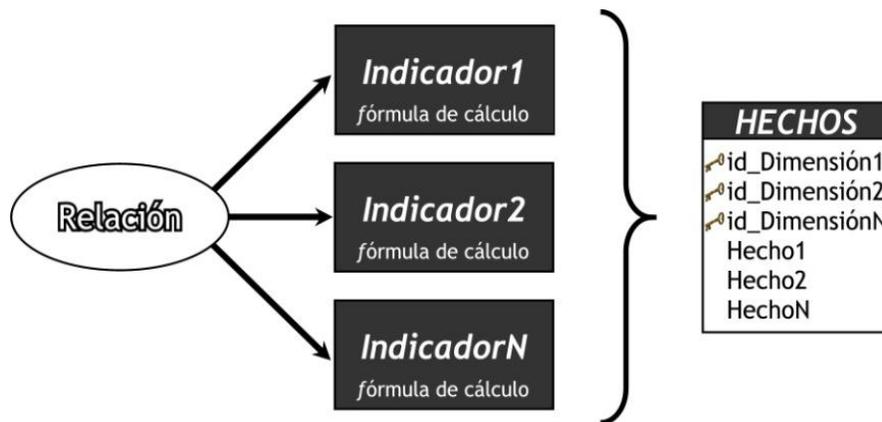


Figura 2.20. Tabla de hechos (Bernabéu, 2010).

Según Inmon (2002), un Data Warehouse es una colección de datos orientados a temas integrados, no volátiles y variante en el tiempo, organizados para soportar necesidades empresariales.

Según la definición del INEI (2005), un Data Warehouse se crea al extraer datos desde una o más bases de datos de aplicaciones transaccionales. La data extraída es transformada para eliminar inconsistencias y luego es cargada en el Data Warehouse. El proceso de transformar, crear el detalle de tiempo variante, resumir y combinar los extractos de datos, ayudan a crear el ambiente para el acceso a la información de la organización. Este nuevo enfoque orienta a las personas en todos los niveles de la organización a efectuar su toma de decisiones con más responsabilidad.

“El Data Warehouse debe tener un esquema que permita entregar al usuario la información que éste requiera, es necesario que se realice un proceso de análisis que permita al desarrollador definir las bases del proyecto que se pretende crear en base a los requerimientos iniciales del usuario, el modelo adecuado. Si la estructura inicial del Data Warehouse no satisface todas las necesidades iniciales, los desarrolladores se verán en la situación de tener que regresar al comienzo del análisis y redefinir la estructura, corrigiendo errores”

(Valdiviezo, Herrera y Jáuregui, 2007).

## TIPOS DE MODELAMIENTO DE UN DATAMART

### A. ESQUEMA ESTRELLA

“El esquema estrella forma un diagrama en forma de estrella teniendo en el centro de la estrella una o más tablas de hechos y las puntas de las estrellas a las tablas de dimensiones” (Ullman y Widom, 1999).

El esquema estrella es una estructura compuesta por una tabla central (tabla de hechos) y un conjunto de tablas organizadas alrededor de esta (tablas de dimensiones). En las puntas de la estrella se encuentran las tablas de dimensión que contienen los atributos de las aperturas que interesan al negocio que se pueden utilizar como criterios de filtro y son relativamente pequeñas. Cada tabla de dimensión se vincula con la tabla de hechos por un identificador (Kimball y Ross, 2013).

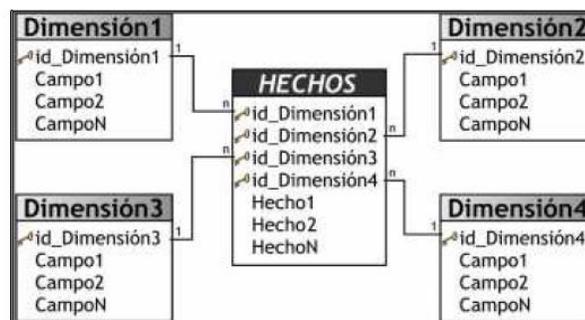


Figura 2.21. Esquema en estrella (Bernabéu, 2010)

### VENTAJAS DEL ESQUEMA ESTRELLA:

- a) Crea una base de datos con tiempos de respuesta rápido
- b) Diseño fácil de modificar.
- c) Simula como ven los datos, los usuarios finales.
- d) Simplifica la navegación.
- e) Facilita la interacción con herramientas.

### B. ESQUEMA COPO DE NIEVE

“El esquema copo de nieve es una extensión del esquema estrella en donde cada uno de los puntos de la estrella se divide en más puntos. En esta forma de esquema, las tablas de

dimensión del esquema estrella contienen más normas. Las ventajas que proporciona el esquema de copo son mejorar el desempeño de consultas debido a un mínimo almacenamiento en disco para los datos y mejorar el desempeño mediante la unión de tablas más pequeñas con normas. Así mismo el esquema copo de nieve incrementa la flexibilidad de las aplicaciones debido a la aplicación de normas y por lo tanto disminuye la granularidad de las dimensiones (Kimball y Ross, 2013).

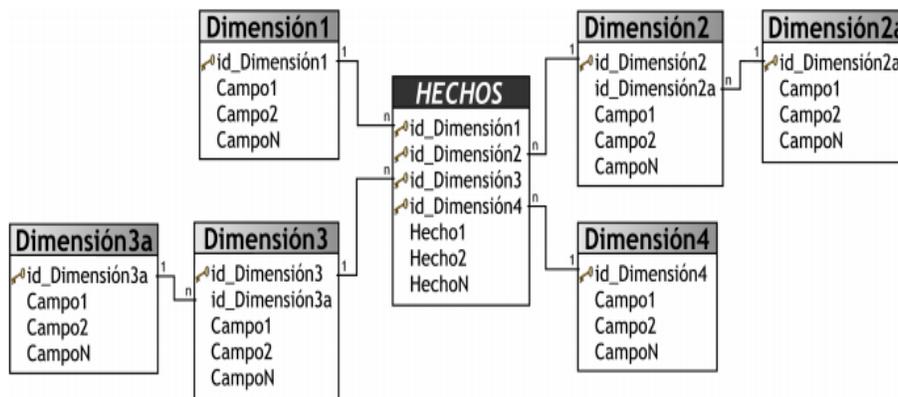


Figura 2.22. Esquema copo de nieve

### C. ESQUEMA CONSTELACIÓN

“Este modelo está compuesto por una serie de esquemas en entrella y formado por una tabla de hechos principal (“HECHOS\_A”) y por una o más tablas de hechos auxiliares (“HECHOS\_B”), las cuales pueden ser sumalizaciones de la principal. Dichas tablas yacen en el centro del modelo y están relacionadas con sus respectivas tablas de dimensiones (Bernabeu, 2010).

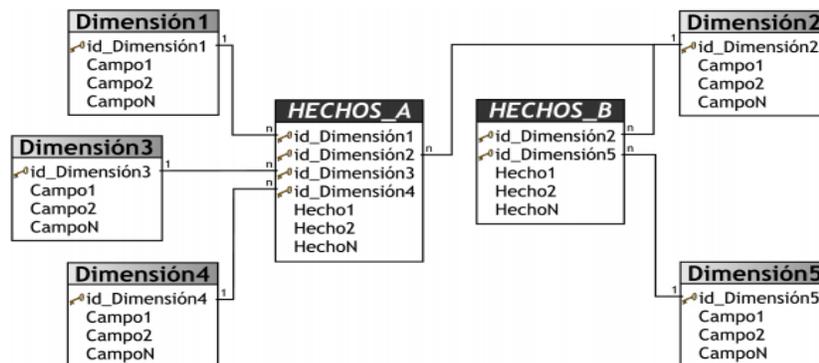


Figura 2.23. Esquema constelación o copo de estrellas (Bernabéu, 2010).

#### 2.2.5.4 DATA WAREHOUSE VS DATA MART

“ Los Data Mart son almacenes de datos especializados, diseñados para soportar necesidades de análisis específicas para un único departamento o área funcional de la empresa. Estos almacenes soportan menos usuarios y menos cantidades de datos que un Data Warehouse centralizado y por lo tanto pueden ser optimizados para cargar y recuperar la información de forma más rápida y eficaz que un Data Warehouse” (Moss y Atre, 2003).

“Un Data Warehouse es un sistema de control de información masivo, es decir engloba la parte global de una organización en el manejo de la información, esto permite administrar todas las áreas funcionales que el Datamart, esta sólo ocupa una parte del global, lo cual permite administrar una sola área en específico esto implica menos recursos y mayor efectividad en la minería de datos” (Bernabéu, 2010).

Tabla 2.2. Diferencias entre Data Warehouse vs DataMart (Yalán y Palomino, 2012)

	DATA WAREHOUSE	DATAMART
<b>Alcance</b>	Construido para satisfacer las necesidades de información de toda la organización	Construido para satisfacer las necesidades de un área de negocios específica.
<b>Objetivo</b>	Diseñado para optimizar la integración y la administración de los datos fuente.	Diseñado para optimizar la entrega de la información de soporte de decisiones.
<b>Características de los datos</b>	Administra grandes cantidades de datos históricos a nivel atómico.	Se concentra en administrar resúmenes y/o datos totalizados.
<b>Pertenencia</b>	Pertenece a toda la organización.	Pertenece al área de negocio al cual está orientado.
<b>Administración</b>	Es administrado por la unidad de sistema de la organización.	Es administrado por el personal de sistema de la unidad propietaria del Datamart.

Fuente: Yalán y Palomino, 2012

#### 2.2.5.5 PROCESO DE EXTRACCION, TRANSFORMACIÓN Y CARGA (ETL)

Según Lluís (2007), el proceso ETL consume entre el 60% y 80% de un proyecto de inteligencia de negocio, esta parte de la construcción de un Data Warehouse es muy costoso por lo cual consume gran parte del proyecto, por lo cual requiere recursos, estrategias y habilidades especializadas.

Según Sivaganesh, Chandra (2012), afirman que el proceso ETL es el responsable de extraer la información de los datos y finalmente la carga en el almacén de datos. El implementar el

escenario de un ETL es un proceso por etapas, que contiene sus relaciones, atributos y las transformaciones entre ellos. Los procesos ETL manejan gran cantidad de datos.

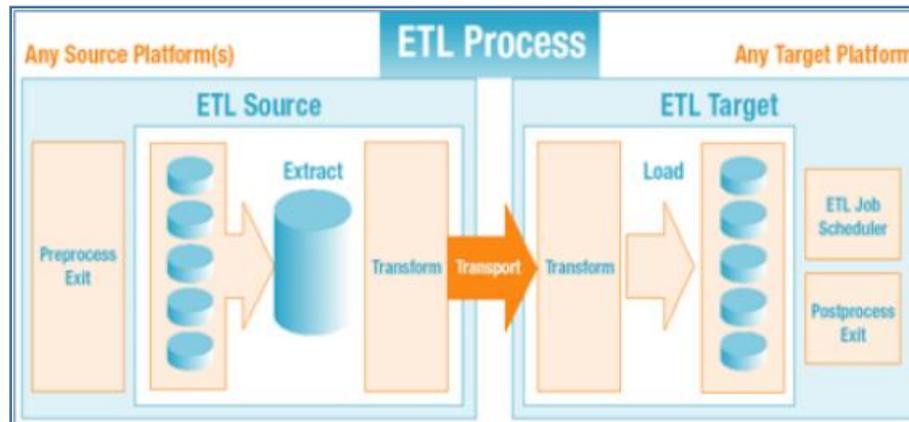


Figura 2.24. Flujo de trabajo ETL (Sirvaganesh y Chandra, 2012)

Bernabéu (2010) menciona que la integración de datos se deberá proceder a poblarlo con datos, utilizando técnicas de limpieza y calidad de datos, procesos ETL, etc.; luego se definirán las reglas y políticas para su respectiva actualización, así como también los procesos que la llevarán a cabo. La integración de datos está constituida por:

**a. Carga inicial**

Debemos en este paso realizar la carga inicial al Data Mart, poblando el modelo de datos que hemos construido anteriormente. Para lo cual debemos llevar adelante una serie de tareas básicas, tales como limpieza de datos, calidad de datos, procesos ETL, etc. La realización de éstas tareas puede contener una lógica realmente compleja en algunos casos.

**b. Actualización**

Cuando se haya cargado en su totalidad el Data Mart, se deben establecer sus políticas y estrategias de actualización o refresco de datos.

Una vez realizado esto, se tendrán que llevar a cabo las siguientes acciones: Especificar las tareas de limpieza de datos, calidad de datos, procesos ETL, etc., que deberán realizarse para actualizar los datos del DW, además se debe especificar de forma general y detallada las acciones que deberá realizar cada software.

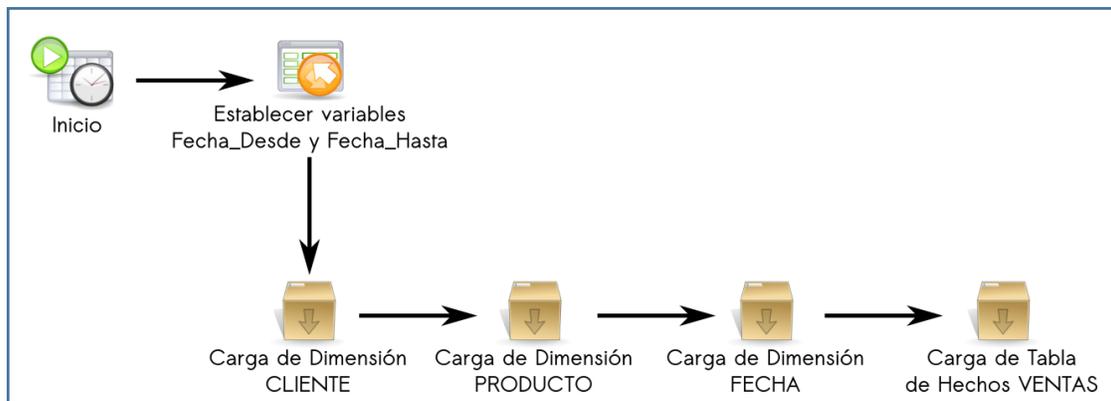


Figura 2.25. Caso práctico, carga Inicial (Bernabéu, 2010).

En Zambrano (2011) se indica que los procesos ETL (de las siglas en inglés Extraction, Transformation, Load) se encargan de las funciones de extracción de distintas fuentes de datos, sean estas transaccionales o externas, transformación, realizando tareas de limpieza y consolidación de datos y la carga del Data Warehouse o Data Mart.

Según Cabanillas (2011), la construcción del ETL consta de tres sub etapas principales: extracción, transformación y carga de datos; a) Extracción, durante esta sub etapa se siguen los procesos necesarios para obtener los datos que permiten efectuar la carga del modelo físico. b) Transformación, durante esta sub etapa se siguen los procesos para convertir los datos fuente a fin de calcular las métricas y mantener un formato estándar de los datos. c) Carga, durante esta sub etapa se siguen los procesos necesarios para poblar los Data Marts.

Según Kimball y Caserta (2004) afirman que un sistema ETL extrae los datos desde los sistemas fuente, verifica los estándares de calidad y consistencia que los datos requieran para ser procesados y los entrega de tal manera que a partir de ellos se pueda construir las herramientas para la toma de decisiones. Dicho proceso no solo se refiere al transporte de los datos de una estructura a otra, por el contrario, durante este proceso se agrega valor a los datos en las siguientes actividades: 1) Elimina errores y corrige datos faltantes, 2) Integra diferentes fuentes de información y 3) Estructura los datos para que puedan ser usadas por las herramientas de análisis, de esta manera garantiza resultados positivos en el procesamiento de la información mediante datos cuantificables que permiten reflejar las buenas practicas en una implementación.

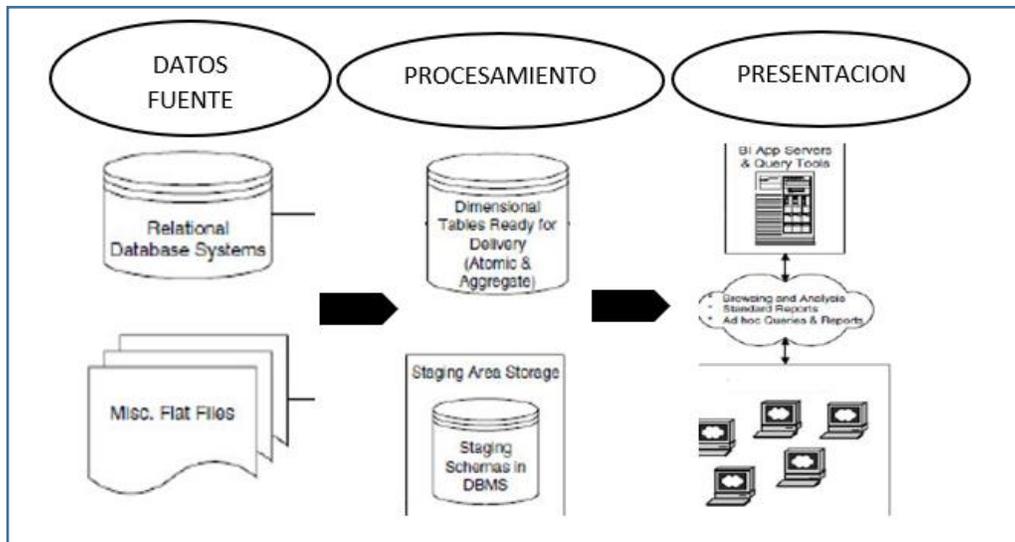


Figura 2.26. capa de presentación de un Data Warehouse (Kimball y Caserta, 2004)

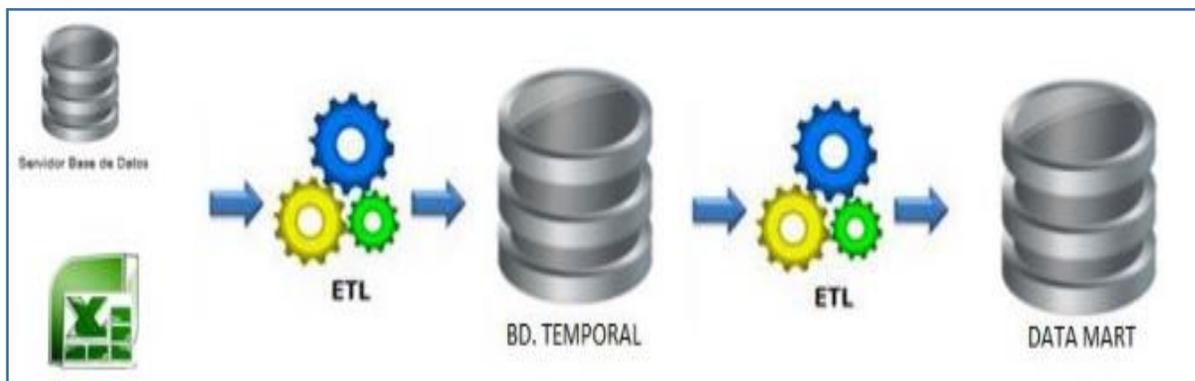


Figura 2.27. Proceso ETL (Esparza y Álvarez, 2012).

### 2.2.6. PROCESAMIENTO ANÁLITICO EN LINEA OLAP (ON-LINE ANALYTICAL PROCESSING)

Según Bernabéu (2010), OLAP es el acrónimo para procesamiento analítico en línea (On-Line Analytical Processing) que representa una forma de ver los datos resumidos históricos en estructuras multidimensionales (conocidos como cubos OLAP) que permitan el acceso rápido a los mismos.

Según Laudon (2008), OLAP soporta el análisis de datos multidimensionales, el cual permite a los usuarios ver los mismos datos en diferentes formas utilizando múltiples dimensiones, obteniendo respuestas en línea a preguntas específicas en un lapso de tiempo sumamente

rápido aún cuando los datos están almacenados en base de datos sumamente grande. OLAP representa las relaciones entre los datos y cubos de datos para permitir un análisis de datos más complejos.

Según Leme (2004), OLAP es un conjunto de herramientas direccionadas a los usuarios finales para acceder a las informaciones del Data Warehouse y que así sea posible crear informes que auxilien la toma de decisión.

“OLAP son aplicaciones que se encargan de analizar datos del negocio para generar información táctica y estratégica que sirve de soporte para la toma de decisiones logrando su máxima eficiencia y flexibilidad operando sobre bases de datos multidimensionales. Se basan en los cubos OLAP, que se construyen agregando, según los requisitos de cada área o departamento, las dimensiones y los indicadores necesarios de cada cubo relacional” (Laudon, 2008).

Según Colaco y Turban (2009), afirman que las tecnologías OLAP es la multidimensionalidad de los datos. Esto significa que los datos son modelados en una estructura conocida como cubo, que en cada dimensión representa temas como por ejemplo, producto, cliente y tiempo.

Según Vitt y Luckevich (2002), OLAP proporciona un modelo de datos intuitivo y conceptual, para que los usuarios que no tengan experiencia como analistas puedan comprender y relacionar los datos mostrados. Este modelo es llamado análisis multidimensional, siendo habilitado para ver los datos a través de múltiples filtros o dimensiones.

Existen variaciones de OLAP según la cantidad de datos y la eficiencia requerida. OLAP no se recomienda para consultas complejas y que recorran muchas tablas.

**A. MOLAP (Multidimensional online analytical processing).**

Según Vitt y Luckevich (2002), Molap brinda el mayor rendimiento de recuperación

de información; porque los datos son estructurados y colocados en niveles especiales que se encuentran en un servidor central.

“Su premisa base es el pre cálculo de sus estructuras, de modo que siempre se hacen las mismas consultas sobre una misma estructura, se carece de una base de datos relacional. Cada vez que un sistema cambia algún sector de sus estructuras vuelve a cargar gran parte de sus dimensiones lo cual lo hace poco dinámico, pero a su vez con una gran capacidad de respuesta” (Moss y Atre, 2003).

“El sistema MOLAP permite almacenar físicamente los datos en estructuras multidimensionales de manera que la representación externa y la interna coincidan, disponiendo de estructuras de almacenamiento específicas (Arrays) y técnicas de comparación de datos que favorecen el rendimiento del Data Warehouse (Bernabéu, 2010).

Según Torres (2007) el MOLAP tiene las siguientes ventajas y desventajas:

#### **Ventajas**

- a) Excelente performance: los cubos MOLAP son contruidos para tener una rápida recuperación de datos y esta optimizado para operaciones.
- b) Puede realizar cálculos complejos, ya que todos los cálculos han sido pre generados cuando el cubo se crea. Por lo tanto, los cálculos complejos se almacenan y regresan su resultado rápidamente.
- c) Extracto eficaz de los datos alcanzados con la pre- estructuración de datos agregados.

#### **Desventajas**

- a) El paso de proceso (carga de los datos) puede ser absolutamente muy largo, especialmente en volúmenes grandes de los datos. Esto es remediado generalmente haciendo solamente el proceso incremental, es decir, procesando solamente los datos que han cambiado (generalmente los nuevos datos).
- b) Limitado en el monto de datos a ser manejados. Porque todos los cálculos son contruidos cuando se genera el cubo, no es posible incluir grandes cantidades de datos en el cubo en sí mismo. Esto no quiere decir que los datos del cubo no deriven de una gran cantidad de datos. Si es posible pero en este caso solo la información

de alto nivel puede estar incluida en ella.

- c) El acercamiento de MOLAP introduce redundancia de datos.

## **B. ROLAP (Relational online analytical processing)**

“En los sistemas ROLAP, los cubos multidimensionales se generan dinámicamente al instante de realizar las diferentes consultas, haciendo de esta manera el manejo de cubos transparente (Bernabéu, 2010).

Según Vitt Luckevich (2002), permite tomar ventaja de uno de sus más grandes beneficios, el almacenamiento de inmensas cantidades de datos. El rendimiento de recuperación de la información para ROLAP frecuentemente no es tan rápido como otras opciones de almacenamiento. Rolap es recomendado para consultas pesadas que no se usan muy a menudo.

“Acrónimo de Relational Online Analytical Processing, almacena los datos en un motor relacional logrando una mejor flexibilidad mediante los tipos de análisis disponibles, tener menor tiempo de respuesta para la elaboración de reportes, análisis de una enorme cantidad de datos. Se implementa sobre tablas físicas diseñadas siguiendo un modelo en estrella o copo de nieve” (Tamayo, M. y Moreno F., 2006).

“ROLAP se utiliza para ahorrar espacio de almacenamiento cuando se trabaja con grandes conjuntos de datos que se consultan con poca frecuencia; por ejemplo, datos exclusivamente históricos (Kimball y Ross, 2013).

Según Torres (2007) el Rolap tiene las siguientes ventajas y desventajas:

### **Características principales:**

- a) Almacena la información en una base de datos relacional
- b) Utiliza índices de mapas de bits.
- c) Posee optimizadores de consultas.
- d) Cuenta con extensiones de SQL (drill-up, drill-down,etc).

### **Desventajas**

- a) Performance bajo. Ya que Rolap es esencialmente múltiples Querys de SQL en la

base de datos relacional, el tiempo de respuesta se alarga entre el tamaño de la base de datos, mientras sea más grande será mas lenta.

- b) Limitada funcionalidad SQL. Ya que la tecnología Rolap utiliza básicamente sentencias SQL ó queries de la base de datos relacional, SQL no aporta todas las necesidades de consultas multidimensionales, Rolap son limitadas a lo que el sistema soporte. Se han desarrollado últimamente herramientas externas que permiten utilizar formulación más compleja que pueda cubrir parte de estas deficiencias.

### **C. HOLAP (Hybrid online analytical processing).**

Según Kimball y Ross (2013), afirman que Holap hace que las agregaciones se almacenen en una estructura multidimensional y los datos a nivel de detalle, en una base de datos relacional como lo hace el almacenamiento Rolap, obteniendo así el rendimiento de las consultas.

Según Vitt y Luckevich (2002), es un híbrido entre Molap y Rolap, Holap no es realmente un modo diferente de almacenamiento de datos. Más bien es la habilidad para diseminar los datos a través de bases de datos relacionales y multidimensionales con la finalidad de obtener lo mejor de ambos sistemas.

“Acrónimo Hybrid Online Analytical Process, almacena datos con las dos técnicas anteriores, utilizando Molap que ofrece análisis sobre los datos agregados, métricas o indicadores precalculados y Rolap que ofrece escalabilidad, cálculo en tiempo real de reportes requeridos por usuarios, concurrencia y administración madura de los datos” (Tamayo, M. y Moreno F., 2006).

Según Bernabéu (2010), menciona que Holap constituye un sistema híbrido entre Molap y Rolap, que combina estas dos implementaciones para almacenar algunos datos en un motor relacional y otros en una base de datos multidimensional, esto debido a la cantidad de datos que estas manejen durante el proceso de análisis de la información, mientras tenga una división se manejará de manera optima los resultados transaccionales del proceso de transformación.

Tabla 2.3. Diferencias entre Molap, Rolap, Holap (Tamayo M. y Moreno F., 2006).

	<b>MOLAP</b>	<b>ROLAP</b>	<b>HOLAP</b>
<b>Datos</b>	Detalle y precalculados	Detalle y agregados	Detalle y agregados
<b>Estructura</b>	Matrices comprimidas	Tablas relacionales	Multidimensional
<b>Administración</b>	Especialista en BDMD	Administrador BD	Administrador BD
<b>Acceso</b>	Lenguaje Especializado	SQL	SQL

Fuente: Tamayo M. y MorenoF, 2006

### 2.2.7. SISTEMA GESTOR DE BASE DE DATOS (SGBD)

“Conjunto de datos con un significado propio, pueden ser numéricos, alfabéticos o alfanuméricos, este es la unidad mínima de información. Un dato dentro de una base de datos responde a la función objeto, atributo, valor” (Juárez, 2006, p.45).

“Es un conjunto de datos interrelacionados, recolectados, que satisfacen las necesidades de información de una comunidad determinada de usuarios, estos representan las bases de datos” (Arreguin, 2006, p.9).

Nevado (s.f.), un SGBD tiene como funciones principales: la descripción, manipulación y control, que en conjunto permiten: a) Definir los datos a los distintos niveles de abstracción (físico, lógico y externo), b) Manipular los datos en la base de datos, permitiendo insertar, modificar, borrar y consultar los datos, c) Mantener la integridad de la base de datos, y d) Control de la privacidad y seguridad de los datos en la base de datos.

“Representa algún aspecto del mundo real, llamado universo de discurso del cual provienen los datos” (Aguilar, 2006, p.17).

“La existencia de los lenguajes denominados algebra relacional y cálculo relacional, son la base del éxito comercial de los SGBDR (Sistemas gestores de Base de Datos Relacionales), cuyo análisis radica en el proceso analítico entre la información existente de medio” (Cobo, s.f., p. 49).

Luque et al. (2002) afirman que un SGBD es una colección de programas de aplicaciones que proporciona al usuario de la base de datos los medios necesarios para realizar las siguientes tareas:

- a) Definición de los datos a los distintos niveles de abstracción (físico, lógico y externo).
- b) Manipulación de los datos en la base de datos, es decir, la inserción, modificación, borrado y acceso o consulta a los mismos.
- c) Mantenimiento de la integridad de la base de datos, integridad en cuanto a los datos en sí, sus valores y las relaciones entre ellos.
- d) Control de la privacidad y seguridad de los datos en la base de datos.

Los medios necesarios para el establecimiento de todas aquellas características exigibles a una base de datos.

## **BASE DE DATOS**

Según Gómez (2007), una base de datos es un conjunto de datos que pertenecen al mismo contexto, almacenados sistemáticamente para su posterior uso, es una colección de datos estructurados según un modelo que refleje las relaciones y restricciones existentes en el mundo real. Los datos que han de ser compartidos por diferentes usuarios y aplicaciones, deben mantenerse independientes de éstas, y su definición y descripción han de ser únicas estando almacenadas junto a los mismos.

Según Opperl y Sheldon (2009), a pesar de la falta de una definición absoluta, la mayoría de las fuentes coinciden en que una base de datos, por lo menos es una colección de datos organizada en un formato estructurado que es definido como metadatos que describe esa estructura. Puede pensar en los metadatos como información sobre los datos almacenados, que define cómo se almacenan éstos en una base de datos.

Según McLeod (2000), afirman que la base de datos está formada por tablas en las cuales se almacenan los datos. Así mismo, las tablas están formadas por columnas y filas a las cuales se les conoce como campos y registros, donde cada uno de ellos representa un valor importante dentro de la tabla.

“Una base de datos de un Sistema de Información es la representación integrada de los conjuntos de entidades correspondientes a las diferentes entidades tipo del Sistema de Información y de sus interrelaciones. Ésta representación informática (o conjunto estructurado de datos) debe poder ser utilizada de forma compartida por muchos usuarios de distintos tipos” (Camps et al., 2005).

Según Batai (1994.), las bases de datos presentan una serie de características propias, las cuales son: 1) Independencia de datos, 2) Eficiencia, 3) Normalización. Estas definen a una base de datos relacional consistente dentro de una arquitectura.

### **2.2.7.1 BASE DE DATOS RELACIONAL**

“Una colección de tablas a cada una de las cuales se asigna un nombre único. Cada tabla tiene una estructura similar” (Korth y Silberschatz, 1993).

De acuerdo a Osorio (2008), es un paradigma que se ha adoptado en las tecnologías de la información, ninguno como el modelo relacional de las bases de datos se ha consolidado de una manera tan categórica y unánime, pudiéndose decir que la actual orientación a objetos debe su éxito a la consolidación de este modelo en la implementación de las bases de datos.

De acuerdo a Heurtel (2009), una base de datos relacional presenta una organización de los datos en el modelo relacional desarrollada en 1970 por Edgar Codd. Es la estructura más extendida actualmente; En una base de datos relacional, los datos se organizan en tablas enlazadas de manera lógica. Una tabla incluye columnas (o campos) que describen una fila (o registro). La relación entre las tablas se establecen mediante una columna.

Según Cood (1970), afirma que una base de datos relacional se utiliza para modelar problemas reales y administrar datos dinámicamente. Es una forma sencilla de representar los datos en este modelo y es la representación por medio de tablas, una para cada relación.

(IBM, 1920), es un conjunto de tablas y se manipula de acuerdo con el modelo de datos relacional. Contiene un conjunto de objetos que se utilizan para almacenar y gestionar los

datos, así como para acceder a los mismos.

### **2.2.8. POBLACIÓN**

Según Tamayo (1997), la población se define como la totalidad del fenómeno a estudiar donde las unidades de población poseen una característica común la cual se estudia y da origen a los datos de la investigación.

Lepkowski (Como se citó en Hernandez Sampieri, Fernández y Baptista, 2014) Una población es el conjunto de todos los casos que concuerdan con una serie de especificaciones.

### **2.2.9. CENSO**

“No siempre, pero en la mayoría de las situaciones si realizamos el estudio en una muestra. Sólo cuando queremos efectuar un censo debemos incluir todos los casos (personas, animales, plantas, objetos) del universo o la población” (Hernández Sampieri et al., 2014).

## **CAPÍTULO III**

### **METODOLOGÍA DE LA INVESTIGACIÓN**

#### **3.1. TIPO Y NIVEL DE LA INVESTIGACIÓN**

##### **A. TIPO DE INVESTIGACIÓN**

Para (Supo, 2015), los estudios observacionales son estudios sin intervención debido a que el investigador no tiene ninguna participación en la medición de los datos, o que cualquier acto de medición realizada por el investigador no modifica los resultados de la investigación.

Para (Supo, 2015), los estudios retrospectivos son estudios donde se cuenta con la información en el que el investigador realizara el estudio de esta, dicha información ayudara al investigador obtener datos relevantes.

Para (Sampieri, 2003), un estudio transversal es cuando se recolectan datos en un solo momento, en un tiempo único, su propósito es describir variables y su incidencia de interrelacion en un momento dado.

##### **B. NIVEL DE INVESTIGACIÓN**

Según (Bernal, 2010), la investigación descriptiva es la capacidad para seleccionar las características fundamentales del objeto de estudio y su descripción detallada. Esta investigación se soporta fundamentalmente en técnicas como la encuesta, la entrevista, la observación y la revisión documental. Algunos temas de investigación descriptiva son: a) estudios de carácter diagnóstico, b) diseños de guías, modelos, productos, prototipos, etc, c) estudio de mercado, d) estudio de descripción de conducta, e) estudio de tiempo y movimiento, etc.

Para (Supo, 2015), la investigación descriptiva es el primero de los niveles cuantitativos. Aquí, el análisis estadístico es indispensable a la hora de hacer cálculos para completar los

objetivos de estudio. Los estudios descriptivos son “univariados” por que el análisis de datos se centra en una sola variable.

Se implementará un DataMart el cual ayudará identificar indicadores que permitan describir la morosidad dentro del área en estudio de la institución “Seda Ayacucho” , usando la metodología Hefesto. Por estas consideraciones el nivel de investigación es descriptivo.

### **3.2. DISEÑO DE LA INVESTIGACIÓN**

La información para el estudio, se ha recolectado registros digitales existentes para el año 2018 en SEDA Ayacucho, esta información recolectada es procesada mediante la técnica Hefesto y los resultados que se obtienen serán presentados mediante indicadores para la morosidad de clientes.

### **3.3. POBLACIÓN Y MUESTRA**

#### **POBLACIÓN**

La población estuvo conformada por 580,000 facturaciones del servicio de agua y alcantarillado, 250,000 reclamos, 50,000 conexiones, información que está en el registro de las bases de datos de Seda Ayacucho del año 2018.

#### **MUESTRA**

No existe muestra, porque se utiliza toda la población , siendo un censo.

### **3.4. VARIABLES E INDICADORES**

#### **3.4.1. DEFINICIÓN CONCEPTUAL DE LAS VARIABLES**

##### **VARIABLE**

**Información táctica.-** Es la información operacional que ayuda en la actividad de control, toma de decisiones y desempeño actual en una organización para instrumentar programas de planeación estratégicas y planes específicos para las áreas funcionales.

##### **VARIABLES DESCRIPTIVAS**

**Cobranza.-** Proceso mediante el cual se hace efectiva la percepción de un pago en concepto de la prestación de un servicio en base a un tiempo determinado.

**Deuda.-** Es una obligación de pago que presenta un cliente frente a la entidad que le presta

algún servicio y está direccionado a un tiempo preestablecido.

**Cliente.-** Es el activo de una empresa cuya importancia radica en satisfacer las necesidades de esta, conduciendo así al crecimiento organizacional.

### **3.4.2. DEFINICIÓN OPERACIONAL DE LAS VARIABLES**

#### **VARIABLE DE INTERES**

X: Información táctica

#### **VARIABLES DESCRIPTIVAS**

Y1: Cobranza

Y2: Deuda

Y3: Cliente

La operacionalización de las variables, se muestra en el anexo A.

### **3.5. TÉCNICAS E INSTRUMENTOS PARA RECOLECTAR INFORMACIÓN**

#### **3.5.1. TÉCNICAS**

Se utilizó la técnica de análisis documental para recolectar datos en relación a la variable información táctica y, a las variables descriptivas cobranza, deuda y cliente.

#### **3.5.2. INSTRUMENTOS**

Se utilizó la ficha de análisis documental , ficha que permite obtener datos a partir de los cubos olap, se muestra en el anexo B.

### **3.6. HERRAMIENTAS PARA EL TRATAMIENTO DE DATOS**

Las herramientas tecnológicas que utilizamos, pertenecen a la línea de Microsoft, debido a que los usuarios por lo general utilizan parte de las herramientas de dicha empresa, Microsoft ofrece a los usuarios una alta gama de tecnologías que facilita en la elaboración de proyectos, enfocado a una línea en específico, esto con la intención de optimizar el tiempo de desarrollo. En función a este aspecto seleccionamos una serie de tecnologías enfocado en los procesos de desarrollo de un Datamart que nos va a permitir consolidar toda la información en estas, lo cual se muestra en la siguiente tabla 3.1:

Tabla 3.1. Herramientas tecnológicas para el tratamiento de datos

<b>SOFTWARE</b>	<b>FABRICANTE</b>	<b>SERVICIO</b>
<b>Windows Server 2012 r12</b>	Microsoft Corporation	Es la versión más reciente de Microsoft del sistema operativo Windows Server, Es la versión para servidores de Windows 8, producida por Microsoft Corporation.
<b>Sql server business intelligence development studio 2012</b>	Microsoft Corporation	Es el entorno que utilizará para desarrollar cubos de Procesamiento analítico en línea (OLAP) y modelos de minería de datos en SQL Server Analysis Services.
<b>Microsoft Sql Server 2012</b>	Microsoft Corporation	Es la última versión del sistema operativo Windows para servidores, ofrece el marco para instalar el servidor web y el servidor de datos.
<b>Visual Studio 2010</b>	Microsoft Corporation.	Es un Entorno Integrado de Desarrollo para Sistemas operativos Windows. Soporta varios lenguajes de programación.
<b>Microsoft Excel</b>	Microsoft Corporation	Es una hoja de cálculo desarrollada por Microsoft cuyo trabajo esta en base a datos numéricos.
<b>VMware Workstation</b>	EMC Corporation	Es un sistema de virtualización, es decir permite a una maquina física ejecutar sistemas operativos simultáneamente.

Fuente: Elaboracion propia.

### 3.7. TÉCNICAS PARA APLICAR LA METODOLOGÍA HEFESTO

Observando la revisión literaria desarrollada en el capítulo II, sección 2.2.5,

formulamos el proceso, que considera las fases para desarrollar el datamart que aplica el método Hefesto, como se muestra en las tablas 3.2 a 3.5

Tabla 3.2. Análisis de requisitos. (Lapa, 2016)

<b>TAREA</b>	<b>ENTREGABLE</b>	<b>TÉCNICA</b>	<b>RESPONSABLES</b>
<b>Identificar preguntas</b>	Lista de preguntas	Entrevistas Formulación de preguntas Observaciones	Analista de negocio, desarrollador
<b>Identificar indicadores y perspectivas de análisis</b>	Lista de indicadores y perspectivas	Entrevistas Encuestas	Desarrollador
<b>Realizar el modelo conceptual</b>	Diagrama del modelo conceptual inicial.	Construir el esquema en estrella	Analista de negocio, desarrollador

Fuente: Lapa, 2016

Tabla 3.3. Análisis de los OLTP

<b>TAREA</b>	<b>ENTREGABLE</b>	<b>TÉCNICA</b>	<b>RESPONSABLES</b>
<b>Conformar indicadores</b>	Lista de indicadores cuantificados.	Determinación de función matemática o estadística	Analista de negocio, desarrollador
<b>Establecer correspondencia</b>	Prueba de existencia de datos	Comparación.	Analista de negocio, desarrollador
<b>Determinar el nivel de granularidad</b>	Nivel de granularidad de las perspectivas	Análisis de las perspectivas	Analista de negocio, desarrollador
<b>Realizar el modelo conceptual ampliado</b>	Diagrama de Modelo Conceptual Ampliado	Ampliación del modelo conceptual con cada perspectiva	Desarrollador

Fuente: Lapa, 2016

Tabla 3.4. Modelo lógico del datamart

<b>TAREA</b>	<b>ENTREGABLE</b>	<b>TÉCNICA</b>	<b>RESPONSABLES</b>
<b>Realizar el modelo lógico</b>	Diagrama del modelo lógico	Esquema estrella	Desarrollador
<b>Diseñar las tablas de dimensiones</b>	Diagrama de las tablas de dimensiones	Definir nombre a la tabla de dimensión Añadir un campo de la clave principal Redefinir nombres de campos	Analista de negocio, desarrollador
<b>Realizar las tablas de hechos</b>	Diagrama de las tablas de hechos	Definir nombre a la tabla de hechos Definir la clave primaria Añadir campos de hechos como indicadores	Analista de negocio, desarrollador
<b>Hacer uniones</b>	Diagrama del esquema de uniones	Esquema estrella	desarrollador

Fuente: Lapa, 2016

Tabla 3.5. Proceso de ETL

<b>TAREA</b>	<b>ENTREGABLE</b>	<b>TÉCNICA</b>	<b>RESPONSABLES</b>
<b>Extracción, transformación de la carga inicial</b>	Carga del almacén intermedio	Limpieza de datos Procesos ETL	Desarrollador
<b>Actualización de las tablas</b>	Datos cargados en el Datamart	Carga de todas las tablas y dimensiones	Analista de negocio, desarrollador
<b>Creación de cubos multidimensionales</b>	Cubo multidimensional	Creación de indicadores, atributos y jerarquías.	Desarrollador

Fuente: Lapa, 2016

## CAPÍTULO IV

### RESULTADOS DE LA INVESTIGACIÓN

#### 4.1. RESULTADOS APLICANDO EL MÉTODO HEFESTO

##### 4.1.1. ANÁLISIS DE REQUISITOS

Según el procedimiento desarrollado en las tablas 3.2 al 3.5, se indica la metodología Hefesto, descrita en el capítulo II sección 2.2.5; fase de análisis de requisitos, se obtiene la lista de preguntas indicadores, perspectivas y el diagrama del modelo conceptual.

Tabla 4.1. Lista de preguntas.

ÍTEM	PREGUNTA	FINALIDAD
<b>1</b>	¿Qué cantidad de clientes pagan su deuda dentro del plazo promedio de cobranza durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>2</b>	¿Cuántos clientes no pagan su deuda dentro del plazo promedio de cobranza durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>3</b>	¿Cuál es el porcentaje de facturas gestionadas durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>4</b>	¿Qué cantidad de la deuda presenta una eficiencia de cobranza durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>5</b>	¿Cual es el porcentaje de clientes que presentan cobranza activa?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos

<b>6</b>	¿Cuál es el porcentaje de cobranza activa de la deuda??	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>7</b>	¿Qué cantidad de clientes presentan deuda deudada durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>8</b>	¿Qué cantidad de clientes son gestionados?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>9</b>	¿Qué cantidad de pagos efectuados se realizó durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>10</b>	¿Qué monto total se recaudo durante el año en base a la facturación neta?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>11</b>	¿Cuál es el porcentaje de la deuda vencida durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>12</b>	¿Cuál es porcentaje recuperado de la deuda vencida durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>13</b>	¿Qué cantidad de clientes presentan reclamos?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>14</b>	¿Qué cantidad reclamos son procesados?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>15</b>	¿Cuál es el porcentaje de clientes facturado por medicion?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos

<b>16</b>	¿Qué porcentaje de la facturación por medición es cubierto durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>17</b>	¿Cuál es la cantidad total de conexiones activas?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos
<b>18</b>	¿Qué cantidad de clientes presentan conexión activa durante el año?	Conociendo la cantidades podrá realizar la monitorización de indicadores y emitir reportes oportunos

Fuente: Elaboracion propia

Tabla 4.2. Lista de indicadores

<b>N°</b>	<b>INDICADORES</b>
<b>IND 1</b>	Pagos fuera del plazo promedio de cobranza
<b>IND 2</b>	% de Eficiencia de cobranza.
<b>IND 3</b>	% de Cobranza gestionada.
<b>IND 4</b>	% de clientes gestionados por deuda.
<b>IND 5</b>	% de pagos efectuados.
<b>IND 6</b>	% de la deuda vencida.
<b>IND 7</b>	% de reclamos totales.
<b>IND 8</b>	% de conexiones activas.

Fuente: Elaboracion propia

Tabla 4.3. Lista de perspectivas

<b>N°</b>	<b>Perspectivas</b>
<b>1</b>	Factura
<b>2</b>	Cliente
<b>3</b>	Pagos
<b>4</b>	Tiempo
<b>5</b>	Reclamos
<b>6</b>	Conexión

Fuente: Elaboracion propia

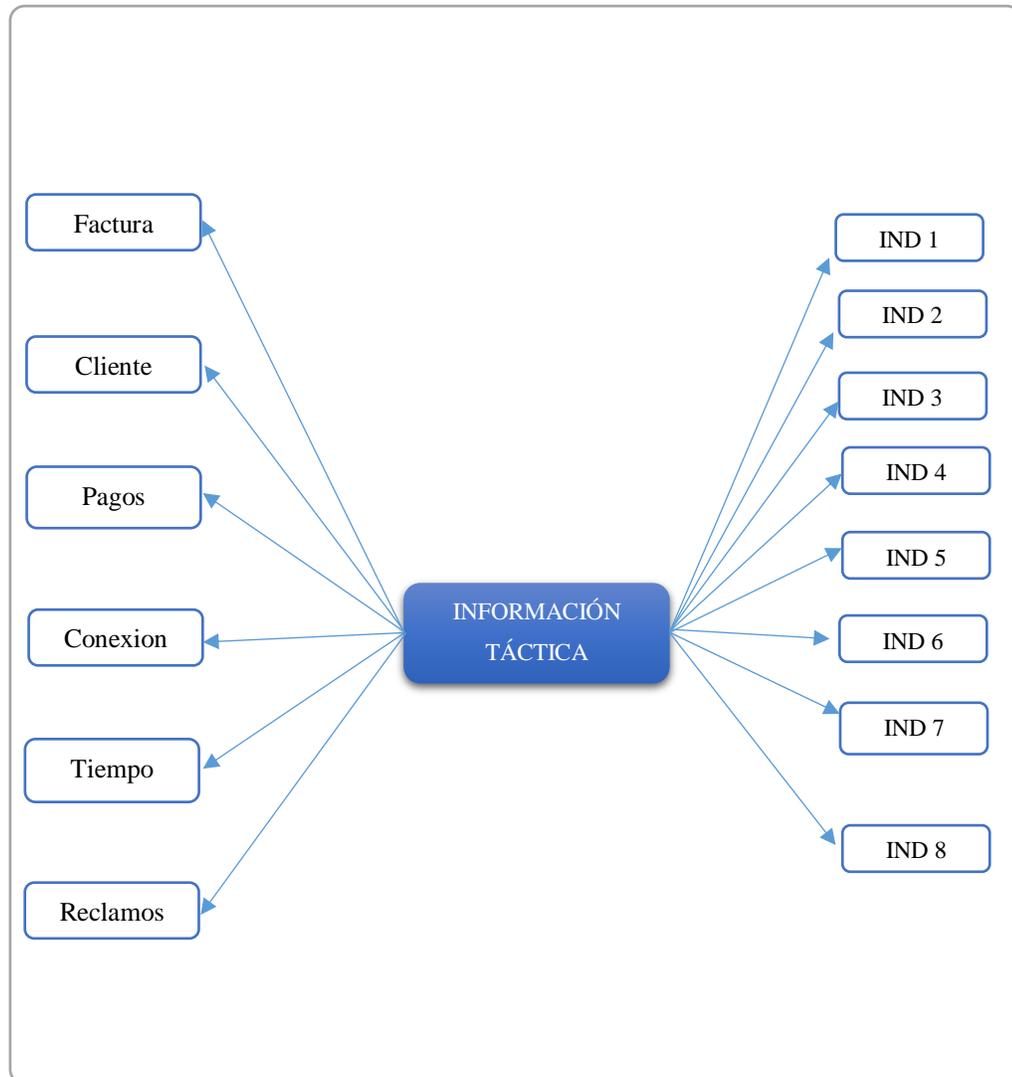


Figura 4.1. Diagrama del modelo conceptual inicial

#### 4.1.2. FASE DEL ANÁLISIS PROCESAMIENTO DE TRANSACCIONES EN LÍNEA (OLTP).

Aplicamos la técnica para la fase de análisis de los procesamientos de transacciones en línea (OLTP), presentada en la tabla Nro. 3.3 según la teoría del capítulo III y la sección 2.2.5 del capítulo II, en donde se obtendrá los indicadores sumariados, agregar llaves y atributos, prueba de existencia de datos y diagrama de modelo conceptual ampliado, donde este último definirá la arquitectura en la implementación del datamart.

Tabla 4.4. Lista de indicadores cuantificados

N°	INDICADORES	HECHOS	FUNCIÓN DE AGREGACIÓN
IND 1	Pagos fuera del plazo promedio de cobranza.	(Pagos fuera de fecha) / (Total de pagos)	SUM, COUNT
IND 2	% de eficiencia de cobranza.	(Cobranza efectuada) / (Facturación neta)	SUM, COUNT
IND 3	% de cobranza gestionada.	(Nro. de facturas no canceladas) / (Nro. de facturas emitidas)	SUM, COUNT
IND 4	% de clientes gestionados por deuda.	(Nro. de facturas gestionadas) / (Nro. de facturas afecto a deuda)	SUM, COUNT
IND 5	% de pagos efectuados.	(Nro. Facturas canceladas) / (Nro. De clientes)	SUM, COUNT
IND 6	% de la deuda vencida.	(Total de la deuda vencida) / (Total de la deuda)	SUM, COUNT
IND 7	% de reclamos totales.	(Nro. De reclamos totales) / (Nro. de conexiones activas)	SUM, COUNT
IND 8	% de conexiones activas.	(Nro. de conexiones activas) / (Nro. de conexiones totales)	SUM, COUNT

Fuente: Elaboración propia.

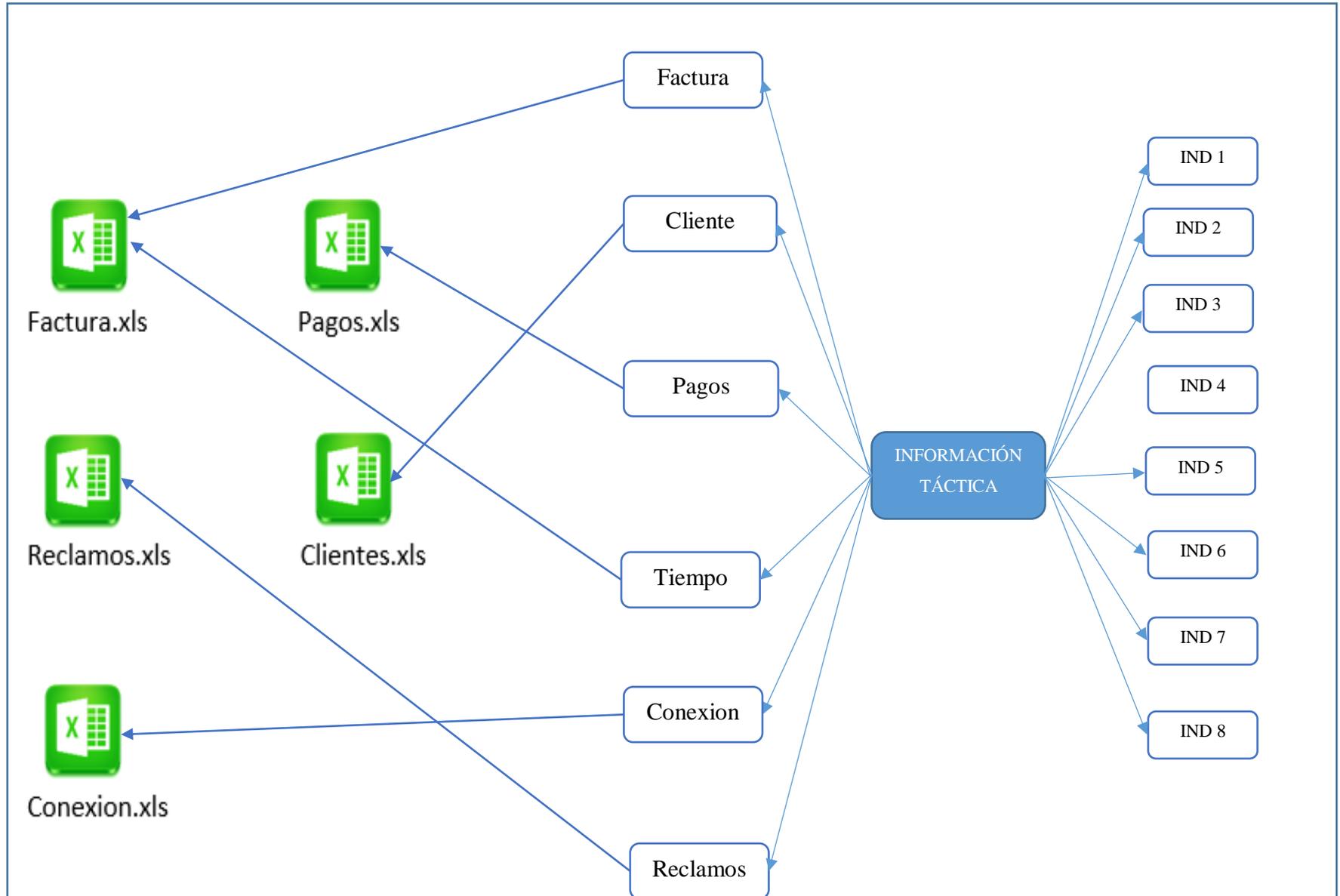


Figura 4.2. Prueba de existencia de datos, correspondencia

Las relaciones encontradas con la prueba de existencia de datos que se muestra en la figura N° 4.2, se detallan a continuación en la tabla 4.5.

Tabla 4.5. Relaciones de la prueba de existencia

N°	RELACIONES
1	El archivo_Factura_ se relaciona con la perspectiva Factura
2	El archivo_Factura se relaciona con la perspectiva Tiempo
3	El archivo_Pagos se relaciona con la perspectiva Pagos
4	El archivo_Reclamos se relaciona con la perspectiva Reclamos
5	El archivo_Clientes se relaciona con la perspectiva Clientes
6	El archivo_Conexion se relaciona con la perspectiva Conexion

Fuente: Elaboracion propia

## NIVEL DE GRANULARIDAD

En las tablas 4.6 a 4.10 se describe el nivel de granularidad de cada perspectiva para la fase de Análisis OLTP.

Tabla 4.6. Nivel de granularidad de la perspectiva Factura

Factura		
Nro.	NOMBRE	DESCRIPCIÓN
1	cod_emp	Codido de empleado
2	cod_suc	Código de sucursal
3	cod-cliente	Código del cliente
4	nro_facturacion	Número de facturación del cliente
5	cod_ciclo	Código del ciclo de facturación
6	año_conexion	Año del servicio en curso
7	mes_conexion	Mes del servicio en curso
8	cod_sect	Código del sector
9	tipo_doc	Tipo de documentación
10	serie_doc	Serie de documentación
11	nro_doc	Número de documentación
12	etd_registro	Estado del registro de la conexión
13	tipo_serv	Tipo del servicio
14	sin_mora	No presenta mora en el servicio
15	Fecha_lecturaant	Fecha de lectura anterior
16	Fecha_lecturault	Fecha de lectura última
17	imp_deudatotal	Importe de la deuda total
18	imp_mesagu	Importe del servicio por mes
19	int_mesagu	Interés del servicio por mes

20	imp_mesotros	Importe por otros conceptos por mes
21	imp_mesmora	Importe de la mora por mes
22	imp_mesigv	Importe del igv por mes
23	imp_mestotal	Importe total del mes
24	mes_refinanc	Refinanciación del servicio por mes
25	dia_pago	Dia del pago del servicio
26	imp_mescta	Importe de cuentas por mes
27	dia_pagocta	Dia del pago por cuotas
28	int_mesigv	Interés del igv generado por mes
29	fecha_corte	Fecha de corte del servicio
30	mes_atraz	Mes atrasado por falta de pago del servicio
31	fecha_registro	Fecha de emisión de la factura
32	fecha_reapertura	Fecha de reapertura de la conexión del servicio
33	fecha_prorroga	Fecha de prórroga del pago
34	imp_mescargofijo	Importe del cargo fijo mensual
35	int_mescargofijo	Interés del cargo fijo mensual
36	deuda_refinancixvencer	Deuda refinanciada por vencer
37	mes_saldoafavor	Saldo mensual del cliente
38	saldo_afavorxvencer	Saldo del cliente a vencer
39	imp_deudareclamo	Importe de la deuda reclamada activa
40	no_facturado	Facturas no emitidas por conexiones clandestinas
41	nro_incobrible	Numero de factura incobrible
42	fecha_vencimiento	Fecha de vencimiento de la factura emitida

Fuente: Elaboracion propia

Tabla 4.7. Nivel de granularidad de la perspectiva Pago

Pago		
Nro.	NOMBRE	DESCRIPCIÓN
1	cod_emp	Código de empleado
2	cod_sede	Código de sede
3	nro_pago	Numero de pago de la conexión del servicio
4	cod_suc	Código de la sucursal
5	cod_cliente	Código del cliente
6	cod_ciclo	Código del ciclo de de facturación del cliente
7	propietario_serv	Id del cliente activo con el servicio
8	dirección_serv	Id de la dirección del cliente activo
9	caja_sede	Id de caja en atención al cliente
10	cod_usuario	Código de usuario
11	tipo_doc	Tipo de documentación
12	seri_edoc	Serie de documentación
13	dia_pago	Dia en el que el cliente efectua el pago
14	pag_anulado	Pago anulado o cancelado
15	imp_subtotal	Importe subtotal de la deuda

16	imp_igv	Importe del igv de la facturación
17	imp_total	Importe total de la deuda
18	imp_efectivo	Importe de pago efectivo
19	eventual_serv	Observaciones
20	fecha_reg	Fecha de emisión de las facturas
21	tipo_servicio	Tipos de servicio al cliente
22	nro_prepago	Numero prepago del cliente
23	cod_usua	Código de usuario
24	pide_vuelto	Registro de movimientos en los pagos

Fuente: Elaboracion propia

Tabla 4.8. Nivel de granularidad de la perspectiva Reclamos

Reclamos		
Nro.	NOMBRE	DESCRIPCIÓN
1	nro_reclamo	Número de registro de reclamo
2	cod_cliente	Código del cliente
3	reclamante_usu	Id del cliente en atencion
4	fecha_reclamo	Fecha de reclamo
5	cod_reclamo	Código de reclamo
6	fecha_resp	Fecha de respuesta
7	reslt_reclamo	Resultado del reclamo
8	fecha_atencion	Fecha de atención
9	tipo_reclamo	Tipo de reclamo
10	cod_emp	Código de empleado
11	cod_sede	Código de sede
12	cod_suc	Código de sucursal
13	cod_oficina	Código de oficina
14	fecha_notifi	Fecha de notificacion
15	fecha_reuni	Fecha de reunión- atención del reclamo
16	estado_rec	Estado del reclamo en curso
17	cod_usua	Código de usuario en atencion
18	c_urbaniza	Registro de la urbanización en atencion
19	fecha_maxrecl	Fecha máximo del reclamo
21	Est_procesado	Proceso del reclamo en curso
22	Fecha_entregada	Fecha de finalización del reclamo

Fuente: Elaboracion propia

Tabla 4.9. Nivel de granularidad de la perspectiva conexion

Conexion		
Nro.	NOMBRE	DESCRIPCIÓN
1	cod_emp	Código de empleado
2	cod_suc	Código de sucursal
3	cod_cliente	Código del cliente en atencion

4	cod_ciclo	Código del ciclo de conexion
5	año_ser	Año del servicio por conexión activa
6	mes_serv	Mes del servicio por conexión activa
7	cod_sector	Código del sector
8	tipo_doc	Tipo de documentacion
9	estado_conexion	Estado de la conexión del servicio
10	tipo_conexion	Tipo de conexión del servicio
11	lectura_anterior	Lectura anterior de la facturacion
12	lectura_ult	Lectura ultima de la facturacion
13	consumo_ser	Registro del consumo de la conexion
14	consumo_fact	Registro del consumo facturado
15	lectura_prom	Registro de la lectura promedio
16	fecha_lectant	Fecha de la lectura anterior
17	fecha_lectult	Fecha de la lectura ultima

Fuente: Elaboracion propia

Tabla 4.10. Nivel de granularidad de la perspectiva Clientes

Clientes		
Nro.	NOMBRE	DESCRIPCIÓN
1	cod_emp	Código de empleado
2	cod_cliente	Código del cliente en atención
3	cod_suc	Código de sucursal
4	cod_sector	Código del sector
5	Cod_mza	Código de ubicación de manzana
6	nro_lote	Registro del numero de lote de la ubicacion
7	nro_ordenru	Número de orden rural
8	cod_ciclo	Código del ciclo
9	propietario_ser	Id del propietario de la conexión en servicio
10	descrip_vv	Descripción de vivienda del cliente
11	id_estadser	Estado de la conexión de servicio
12	tipo_ser	Tipo de servicio de conexión del cliente
13	nro_med	Numero del medidor del cliente
14	tipo_activ	Tipo de activación
15	Obse_conexion	Observación de la conexión en servicio
16	cod_reserv	Código del reservorio
17	descrip_sectorial	Descripción sectorial del cliente
18	descrip_estadomed	Descripción de estado del medidor del cliente
19	cod_depar	Código de departamento
20	cod_prov	Código de provincia
21	cod_distrito	Código de distrito

Fuente: Elaboracion propia

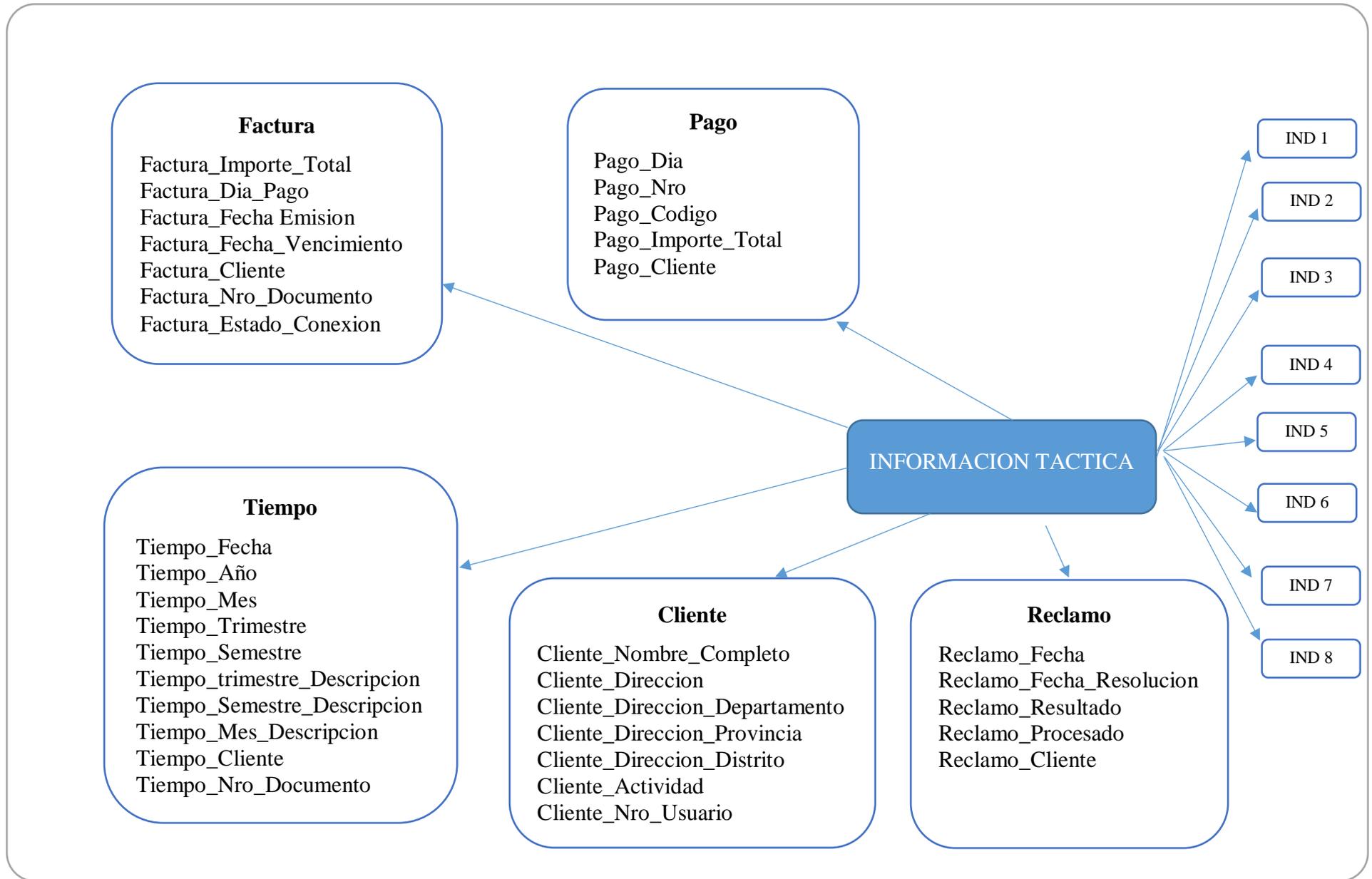


Figura 4.3. Diagrama de modelo conceptual ampliado

#### 4.1.3. FASE DE MODELO LÓGICO DEL DATAMART

La fase de Modelo lógico de datamart presentado en el capítulo II y la técnica referenciada en la tabla presentada en la tabla 3.4, permite obtener el entregable, Diagrama del modelo lógico.

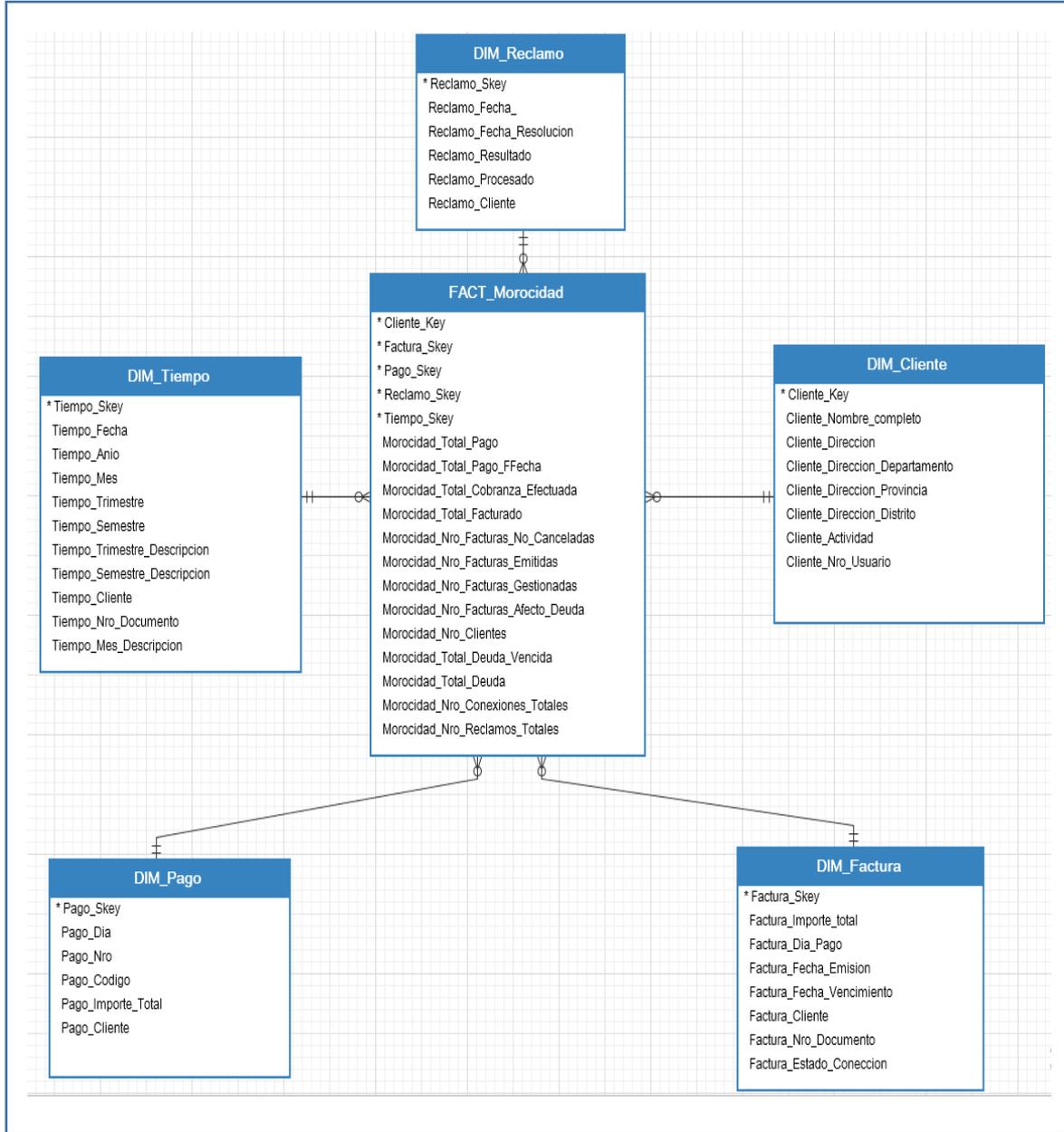


Figura 4.4. Diagrama de modelo lógico

## TABLAS DE DIMENSIONES

Se crea la base de datos “DB\_DATAMART\_MOROCIDAD”, donde se involucra la información digital del año 2018 proporcionada por la entidad , que contine información sobre Facturaciones, pagos , con ellos se procederá a realizar el datamart.

### PERSPECTIVA FACTURA

La nueva tabla de dimensión tendrá el nombre DIM\_Factura

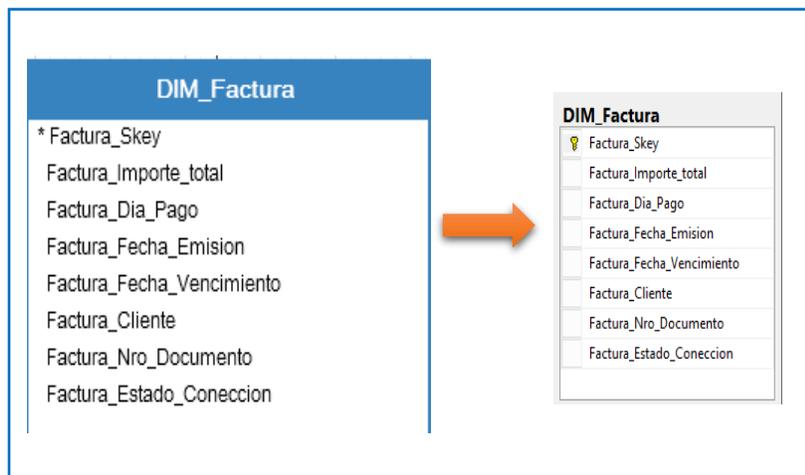


Figura 4.5. Diagrama de la tabla DIM\_Factura

### PERSPECTIVA CLIENTE

La nueva tabla de dimensión tendrá el nombre DIM\_Cliente

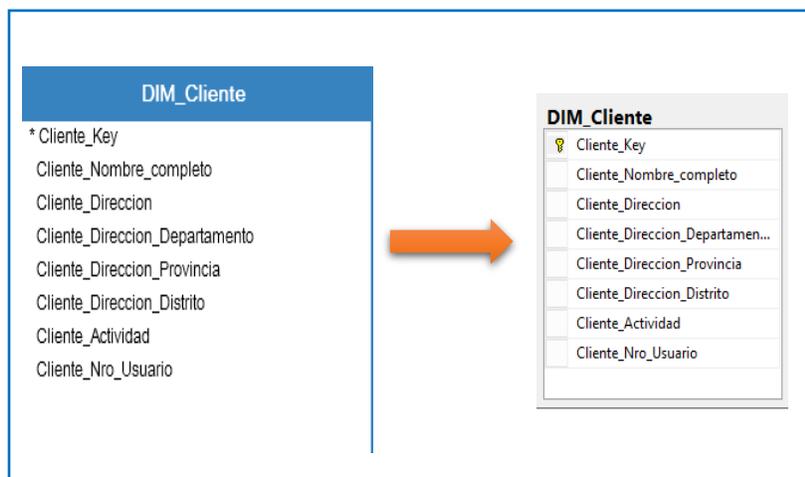


Figura 4.6. Diagrama de la tabla DIM\_Cliente

## PERSPECTIVA PAGOS

La nueva tabla de dimensión tendrá el nombre DIM\_Pago

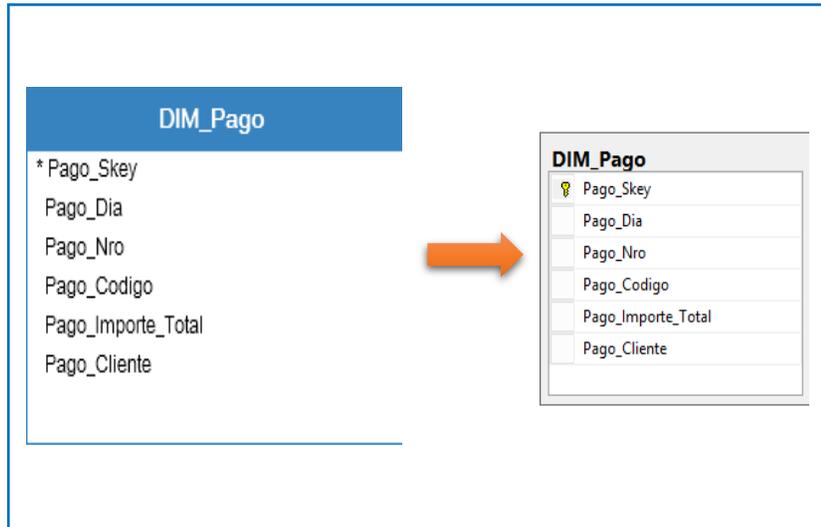


Figura 4.7. Diagrama de la tabla DIM\_Pago

## PERSPECTIVA TIEMPO

La nueva tabla de dimensión tendrá el nombre DIM\_Tiempo

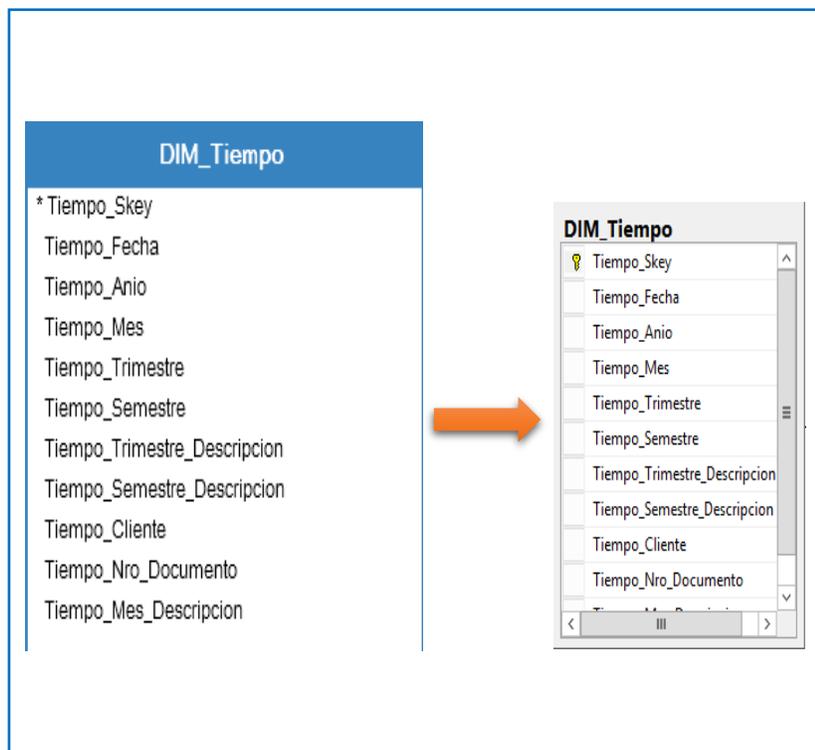


Figura 4.8. Diagrama de la tabla DIM\_Tiempo

## PERSPECTIVA RECLAMOS

La nueva tabla de dimensión tendrá el nombre DIM\_Reclam

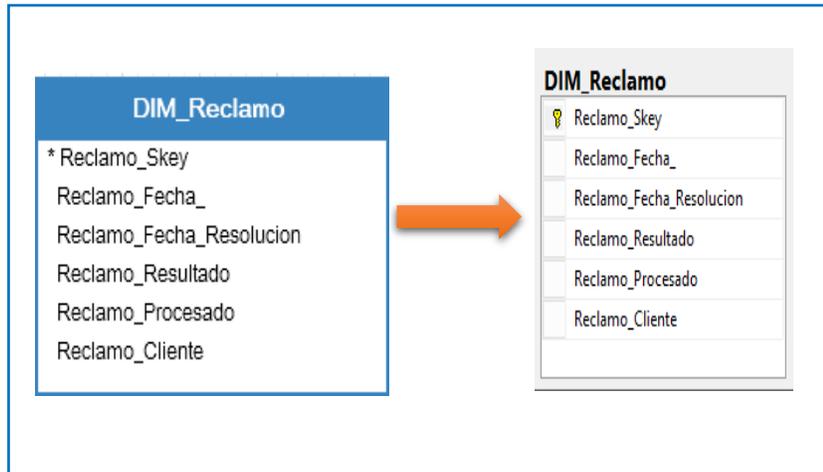


Figura 4.9. Diagrama de la tabla DIM\_Reclam

## TABLAS DE HECHOS

En las siguientes figuras identificaremos las tablas de hechos, los cuales representaran la información analizada, definiremos las tablas de hecho con los indicadores ya definidos en el modelo conceptual.

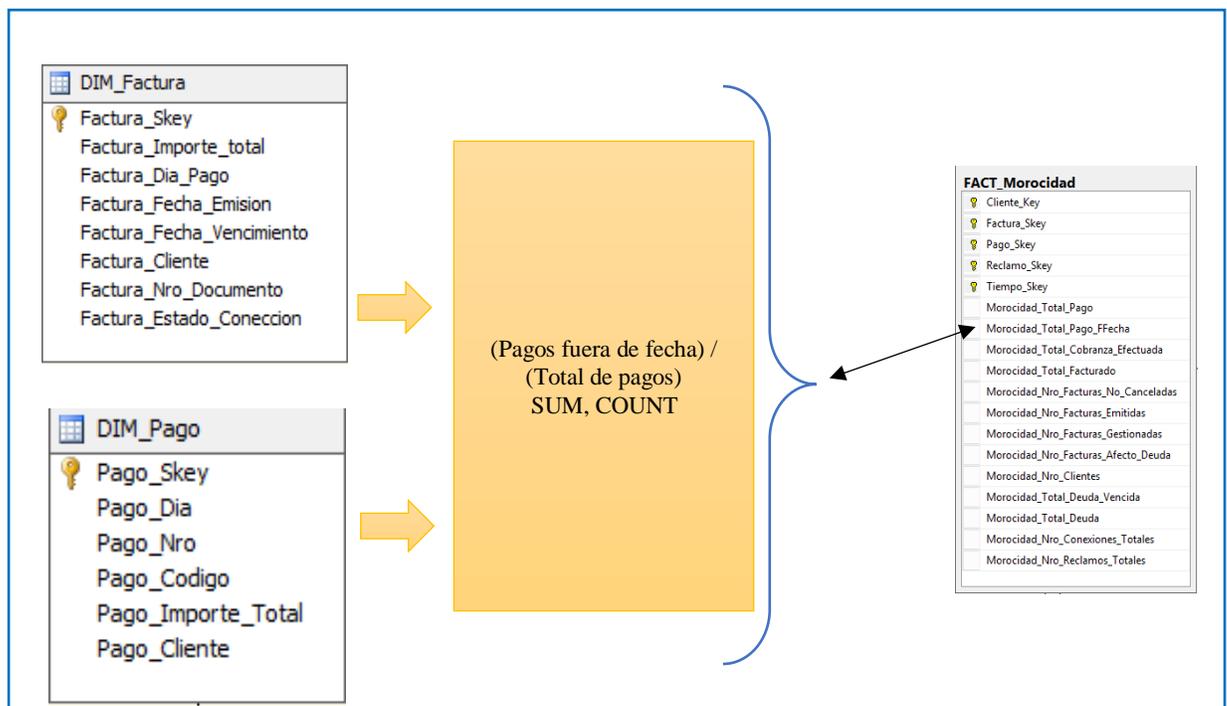


Figura 4.10. Diagrama de la tabla de hechos Morocidad\_Total\_Pago\_FFecha

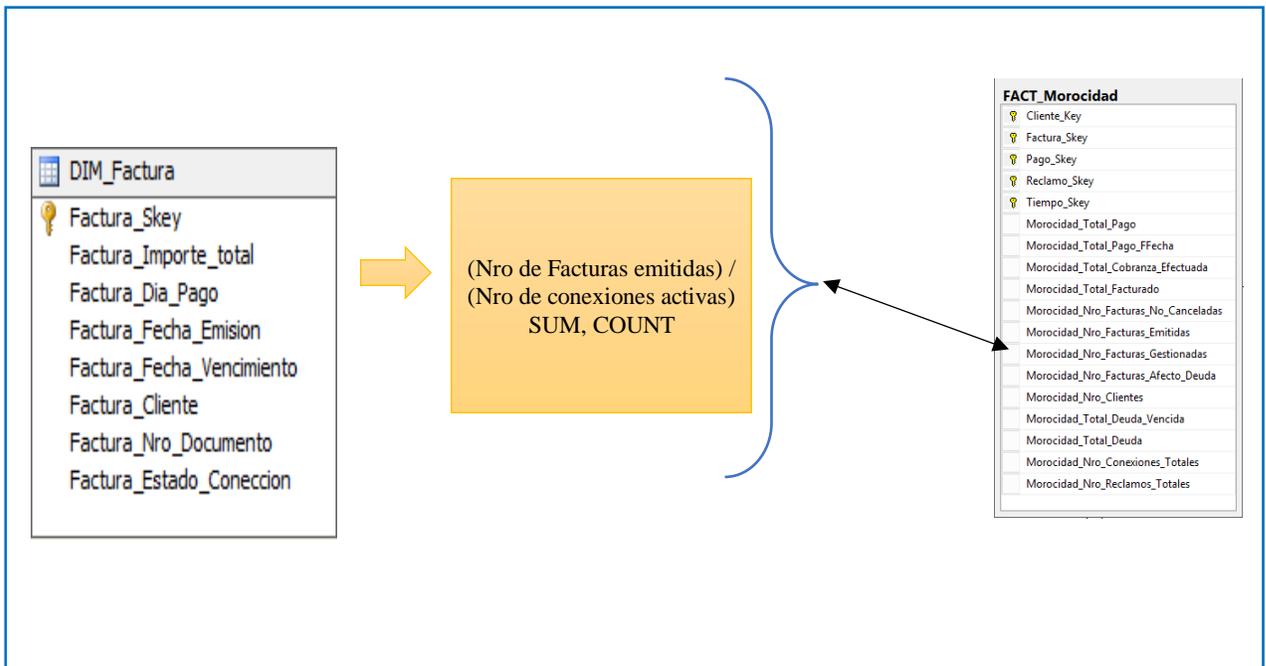


Figura 4.11. Diagrama de la tabla de hechos Morocidad\_Nro\_Facturas\_Gestionadas

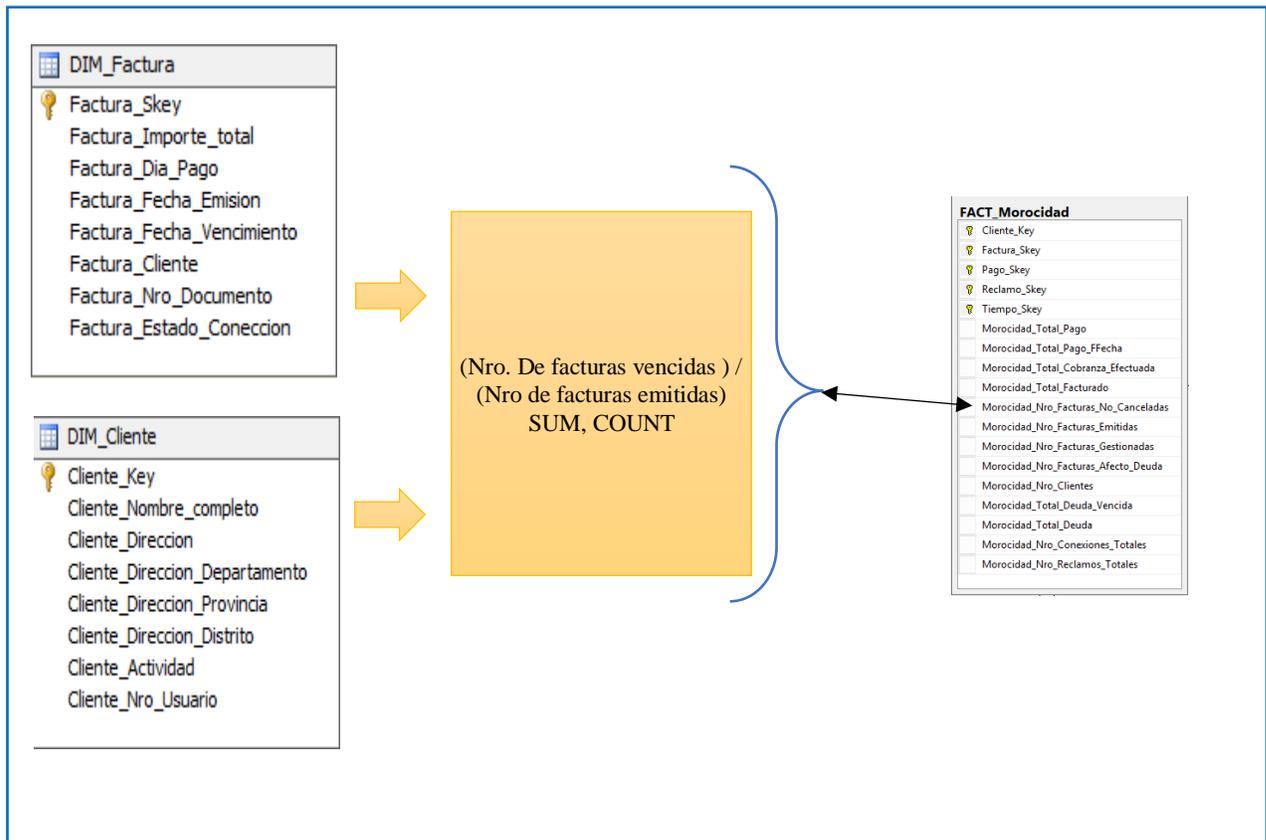


Figura 4.12. Diagrama de la tabla hechos Morocidad\_Nro\_Facturas\_No\_Canceladas

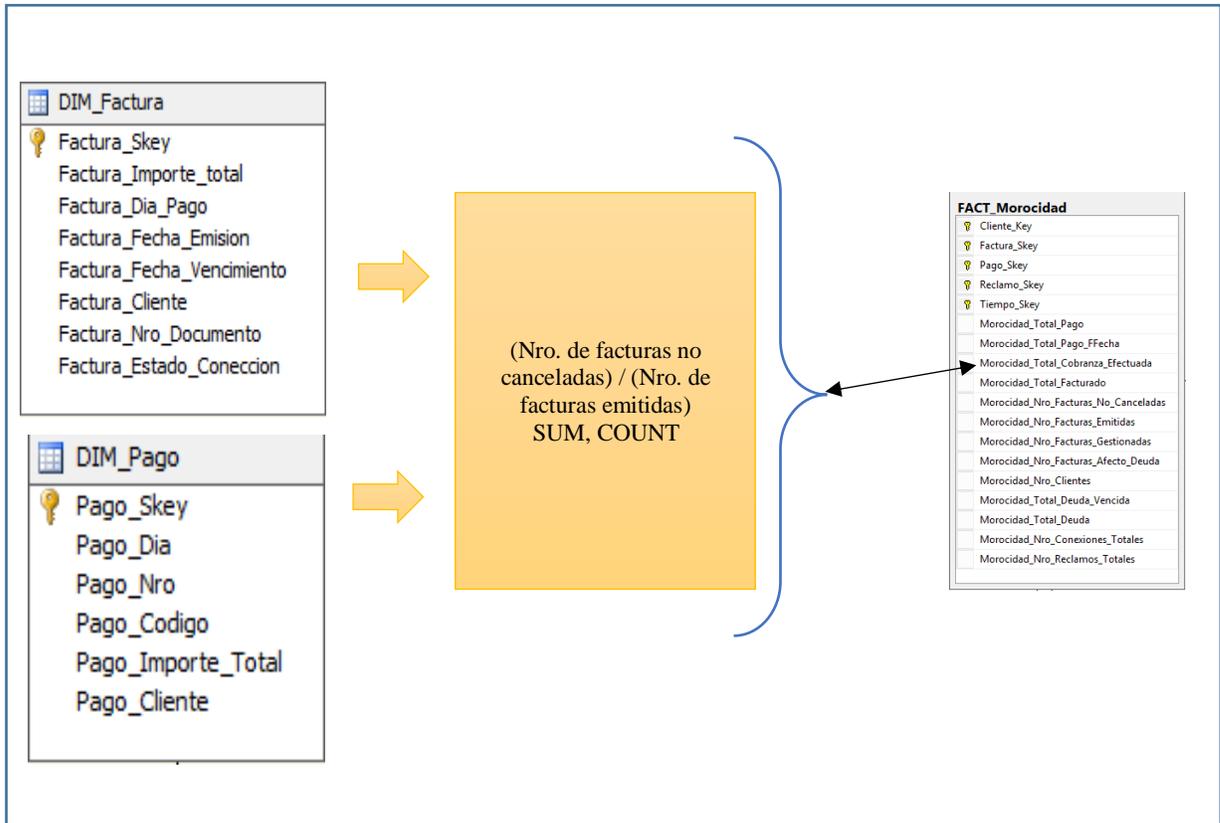


Figura 4.13. Diagrama de la tabla hechos Morocidad\_Total\_Cobranza\_Efectuada

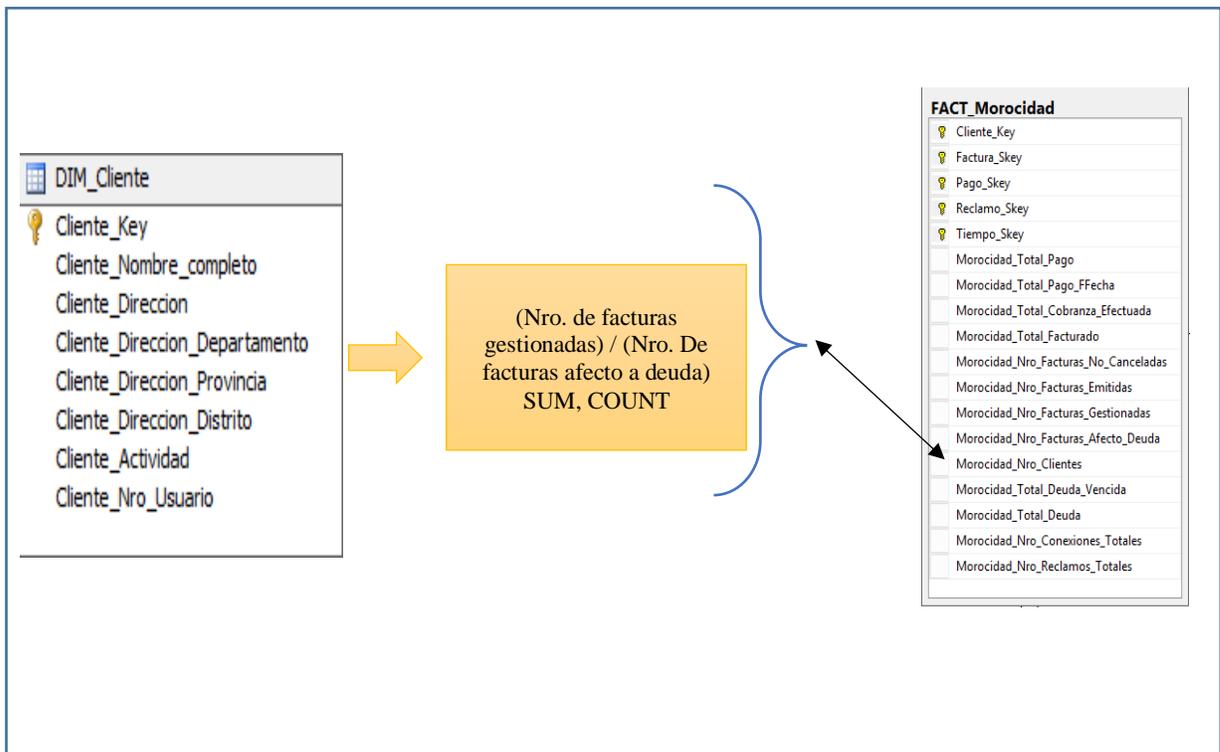


Figura 4.14. Diagrama de la tabla de hechos Morocidad\_Nro\_Clientes

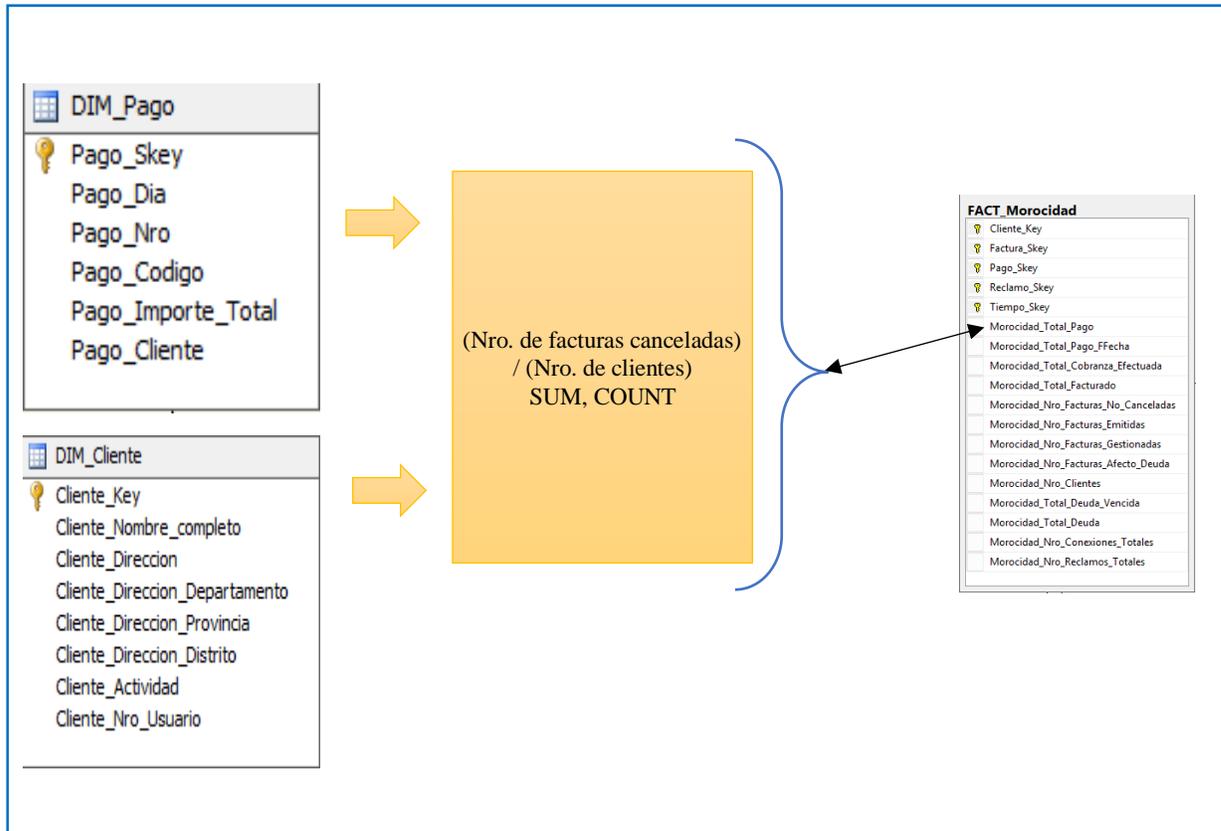


Figura 4.15. Diagrama de la tabla de hechos Morocidad\_Total\_Pago

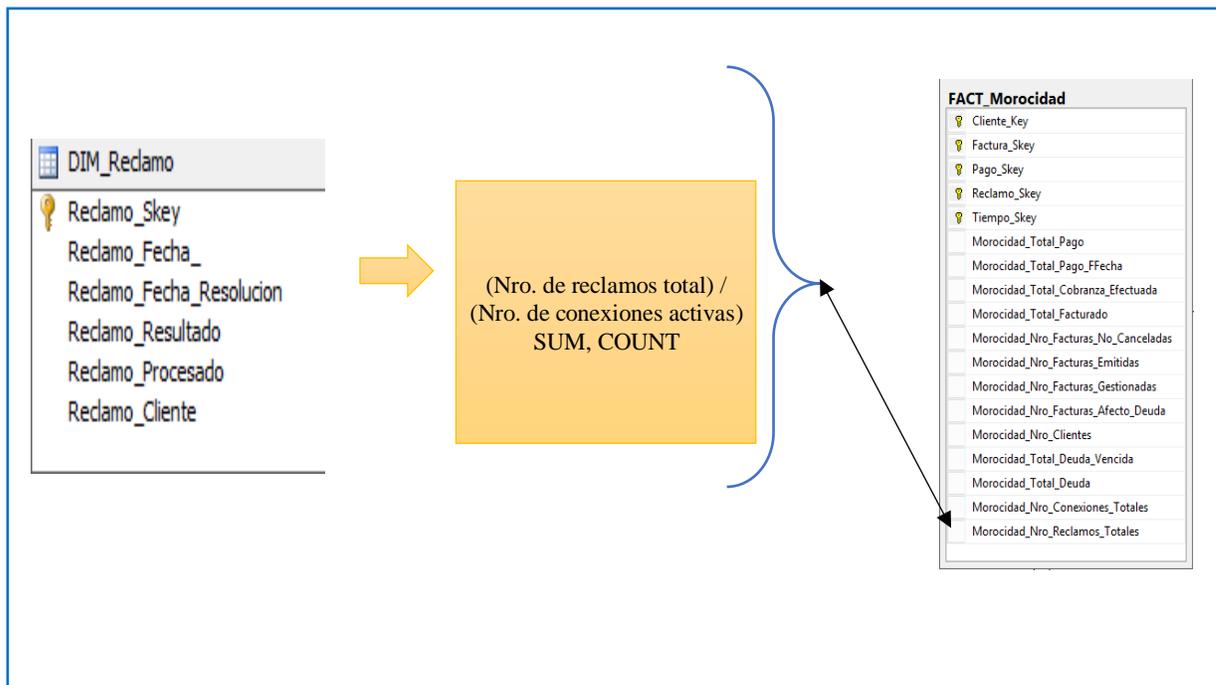


Figura 4.16. Diagrama de la tabla de hecho Morocidad\_Nro\_Reclamos\_Totales

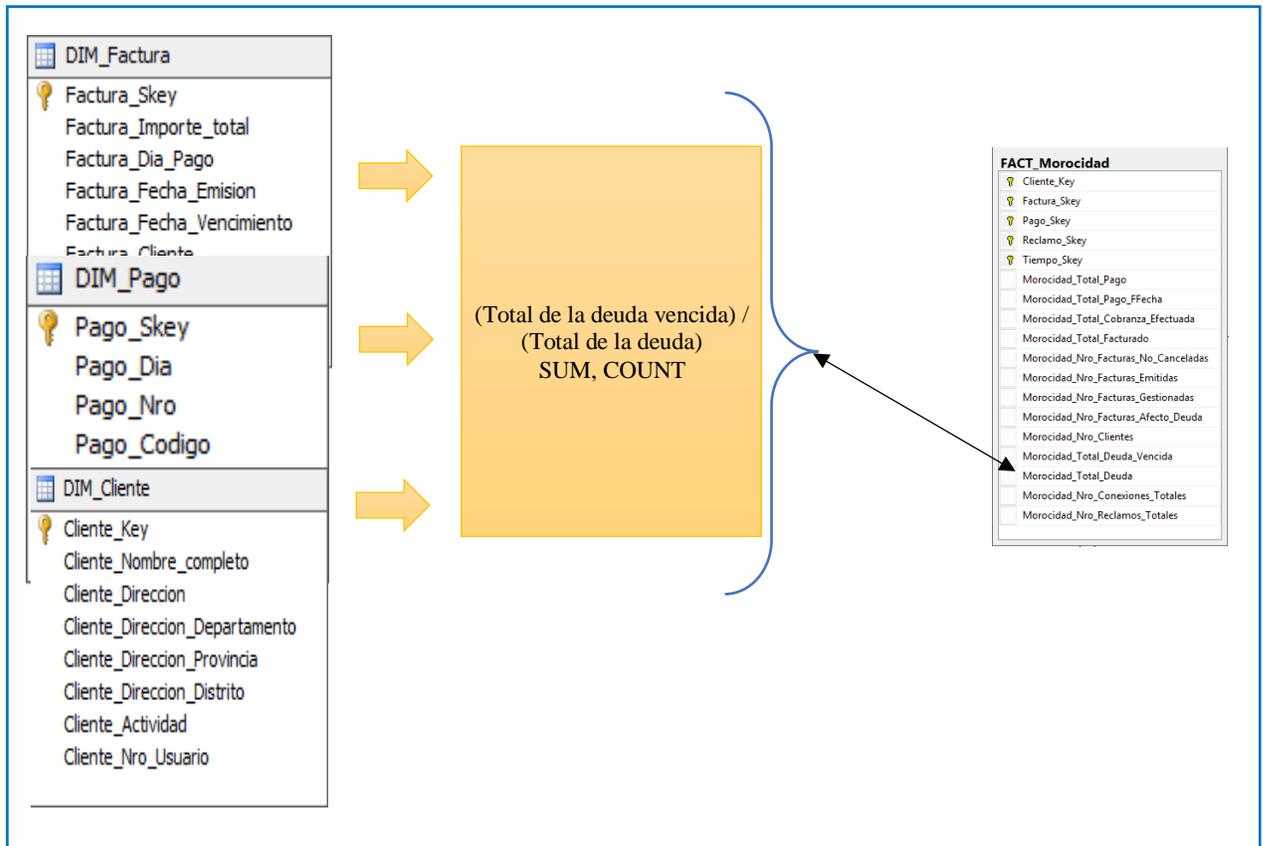


Figura 4.17. Diagrama de la tabla de hecho Morocidad\_Total\_Deuda

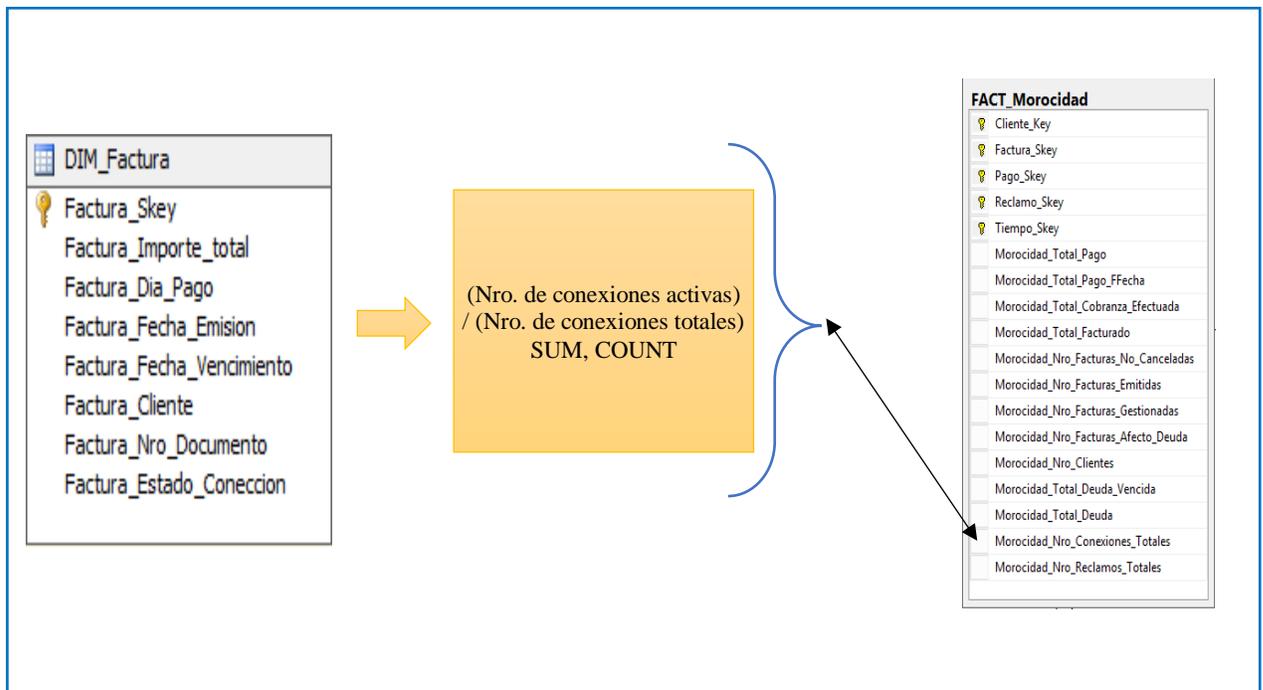


Figura 4.18. Diagrama de la tabla de hecho Morocidad\_Nro\_Conexiones\_Totales

# UNIONES

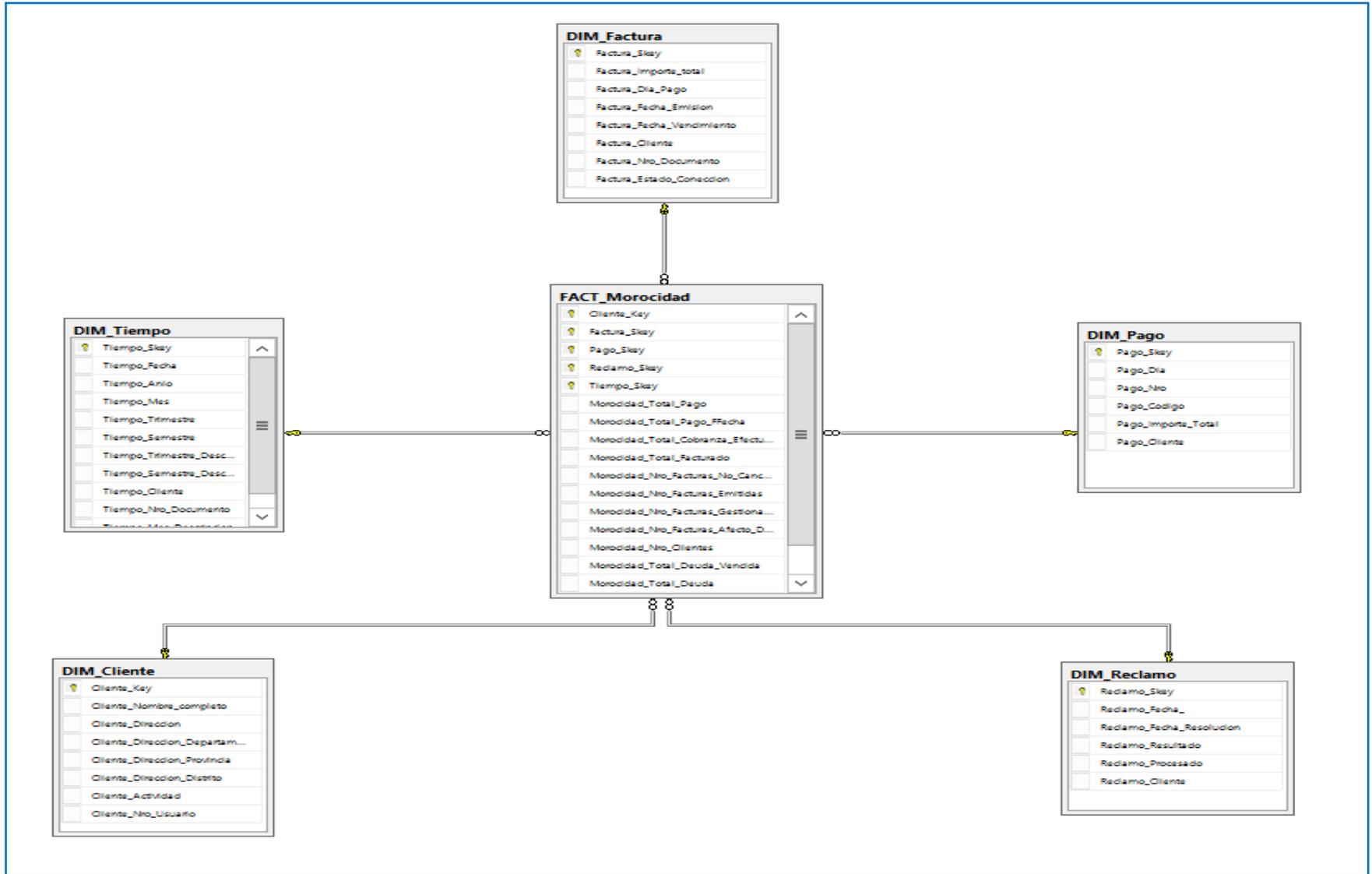


Figura 4.19. Diagrama esquema de uniones, esquema estrella de la base de datos

#### 4.1.4. FASES DEL PROCESO ETL

##### CARGA INICIAL

Para la carga de nuestro modelo se procede a llenar cada una de las dimensiones de nuestro modelo.

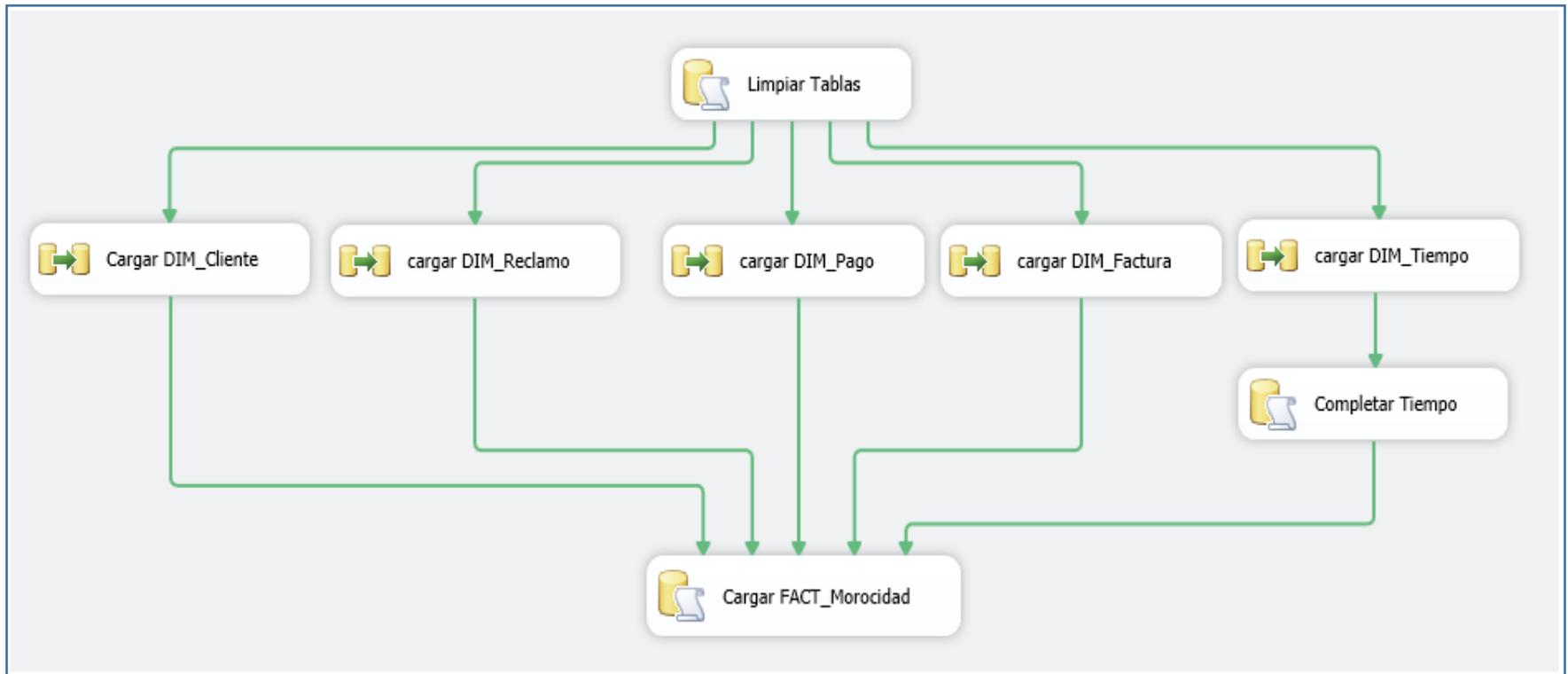


Figura 4.1. Proceso ETL para la carga inicial

Las tareas que lleva a cabo este proceso son:

- a. **Limpiar Tablas:** Borra todos los registros, tablas y relaciones de la base de datos DB\_DataMart\_Morocidad.
- b. **Carga de dimensión Cliente:** Se ejecuta el proceso de la carga de la dimensión Cliente, se detallará más adelante.
- c. **Carga de dimensión Reclamo:** Se ejecuta el proceso de la carga de la dimensión Reclamo, se detallará más adelante.
- d. **Carga de dimensión Pago:** Se ejecuta el proceso de la carga de la dimensión Pago, se detallará más adelante.
- e. **Carga de dimensión Factura:** Se ejecuta el proceso de la carga de la dimensión Factura, se detallará más adelante.
- f. **Carga de dimensión Tiempo:** Se ejecuta el proceso de la carga de la dimensión Tiempo, se detallará más adelante.
- g. **Carga de tabla d hechos Morocidad\_Total\_Pago:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Total\_Pago.
- h. **Morocidad\_Total\_Pago\_FFecha:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Total\_Pago\_FFecha.
- i. **Morocidad\_Total\_Cobranza\_Efectuada:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Total\_Cobranza\_Efectuada.
- j. **Morocidad\_Total\_Facturado:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Total\_Facturado.
- k. **Morocidad\_Nro\_Facturas\_No\_Canceladas:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Nro\_Facturas\_No\_Canceladas.
- l. **Morocidad\_Nro\_Facturas\_Emitidas:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Nro\_Facturas\_Emitidas.
- m. **Morocidad\_Nro\_Facturas\_Gestionadas:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Nro\_Facturas\_Gestionadas.
- n. **Morocidad\_Nro\_Facturas\_Afecto\_Deuda:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Nro\_Facturas\_Afecto\_Deuda.
- o. **Morocidad\_Nro\_Clientes:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Nro\_Clientes.

- p. **Morocidad\_Total\_Deuda\_Vencida:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Total\_Deuda\_Vencida.
- q. **Morocidad\_Nro\_Conexiones\_Totales:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Nro\_Conexiones\_Totales.
- r. **Morocidad\_Nro\_Reclamos\_Totales:** Se ejecuta el proceso de la carga de la tabla de hechos Morocidad\_Nro\_Reclamos\_Totales.

A continuación, se especificarán las tareas llevadas a cabo para cargar las dimensiones y hechos para el proceso de carga se utiliza la herramienta de Integration Services Connections Project de SQL server Business Intelligence Development Studio

### **PROCESO ETL**

Mediante las figuras de tal a tal, mostramos el código para poder migrar la información existente a las dimensiones creadas para generar los cubos.

Para lo cual se creó la Base de Datos “DB\_DataMart\_Morocidad”, desde donde se obtendrá la información necesaria de la morocidad.

```
USE [DB_DataMart_Morocidad]
GO
/***** Object: StoredProcedure [dbo].[sp_limpiar_datos]  Script Date: 08/08/2019 5:26:49
*****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER procedure [dbo].[sp_limpiar_datos]
as
begin

truncate table FACT_Morocidad
delete from DIM_Cliente
delete from DIM_Pago
delete from DIM_Reclamo
delete from DIM_Tiempo
delete from DIM_Factura
end
```

Figura 4.2. Proceso ETL para limpieza de base de dato de DB\_DataMart\_Morocidad

```

USE [DB_DataMart_Morocidad]
GO
/***** Object: StoredProcedure [dbo].[sp_cargar_tiempo]  Script Date: 08/08/2019 6:08:20
*****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER procedure [dbo].[sp_cargar_tiempo]
as
begin
    update DIM_Tiempo set DIM_Tiempo.Tiempo_Trimestre=cast(case
when Tiempo_Mes in('01','02','03') then '01'
when Tiempo_Mes in('04','05','06') then '02'
when Tiempo_Mes in('07','08','09') then '03'
when Tiempo_Mes in('10','11','12') then '04'
end as nvarchar), DIM_Tiempo.Tiempo_Semestre=cast(case
when Tiempo_Mes in('01','02','03','04','05','06') then '01'
when Tiempo_Mes in('07','08','09','10','11','12') then '02'
end as nvarchar),DIM_Tiempo.Tiempo_Trimestre_Descripcion=cast(case
when Tiempo_Mes in('01','02','03') then 'PRIMER TRIMESTRE'
when Tiempo_Mes in('04','05','06') then 'SEGUNDO TRIMESTRE'
when Tiempo_Mes in('07','08','09') then 'TERCER TRIMESTRE'
when Tiempo_Mes in('10','11','12') then 'CUARTO TRIMESTRE'
end as nvarchar),DIM_Tiempo.Tiempo_Semestre_Descripcion=cast(case
when Tiempo_Mes in('01','02','03','04','05','06') then 'PRIMER SEMESTRE'
when Tiempo_Mes in('07','08','09','10','11','12') then 'SEGUNDO SEMESTRE'
end as nvarchar),DIM_Tiempo.Tiempo_Mes_Descripcion=CAST(case Tiempo_Mes
when '01' then 'ENERO'
when '02' then 'FEBRERO'
when '03' then 'MARZO'
when '04' then 'ABRIL'
when '05' then 'MAYO'
when '06' then 'JUNIO'

```

```

when '07' then 'JULIO'
when '08' then 'AGOSTO'
when '09' then 'SETIEMBRE'
when '10' then 'OCTUBRE'
when '11' then 'NOVIEMBRE'
when '12' then 'DICIEMBRE'
end as nvarchar)
end

```

Figura 4.3. Proceso ETL para cargar tiempo

```

USE [DB_DataMart_Morocidad]
GO
/***** Object: StoredProcedure [dbo].[PS_CARGAR_FACTMOROCIDAD]  Script Date:
08/08/2019 6:14:37 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER procedure [dbo].[PS_CARGAR_FACTMOROCIDAD]
AS
BEGIN

INSERT INTO FACT_Morocidad([Cliente_Key]
,[Factura_Skey],[Pago_Skey],[Reclamo_Skey],[Tiempo_Skey],[Morocidad_Total_Pago],[Morocid
ad_Total_Pago_FFecha],[Morocidad_Total_Cobranza_Efectuada],
[Morocidad_Total_Facturado],[Morocidad_Nro_Facturas_No_Canceladas],[Morocidad_Nro_Factu
ras_Emitidas],[Morocidad_Nro_Facturas_Gestionadas],[Morocidad_Nro_Facturas_Afecto_Deuda],

[Morocidad_Nro_Clientes],[Morocidad_Total_Deuda_Vencida],[Morocidad_Total_Deuda],[Moroc
idad_Nro_Conexiones_Totales],[Morocidad_Nro_Reclamos_Totales])
SELECT C.[Cliente_Key],F.[Factura_Skey],P.[Pago_Skey],R.[Reclamo_Skey],T.[Tiempo_Skey],
CAST(P.Pago_Importe_Total as double precision), CAST(F.Factura_Importe_total as double
precision)-CAST(P.Pago_Importe_Total as double precision),

```

```

CAST(P.Pago_Importe_Total as double precision), CAST(F.Factura_Importe_total as double
precision), CAST(case when CAST(F.Factura_Importe_total as double precision)-
CAST(P.Pago_Importe_Total as double precision) > 0 then 1 when CAST(F.Factura_Importe_total
as double precision)-CAST(P.Pago_Importe_Total as double precision) = 0 then 0 end as double
precision),
1, CAST(case when CAST(F.Factura_Importe_total as double precision)-
CAST(P.Pago_Importe_Total as double precision) > 0 then 0 when CAST(F.Factura_Importe_total
as double precision)-CAST(P.Pago_Importe_Total as double precision) = 0 then 1 end as double
precision),
CAST(case when CAST(P.Pago_Importe_Total as double precision)=0 then 1 when
CAST(P.Pago_Importe_Total as double precision)>0 then 0 end as double precision),1,CAST(case
when CAST(P.Pago_Importe_Total as double precision) =0 then CAST(F.Factura_Importe_total as
double precision)-CAST(P.Pago_Importe_Total as double precision)
when CAST(P.Pago_Importe_Total as double precision) >0 then 0 end as double precision),
CAST(F.Factura_Importe_total as double precision)-CAST(P.Pago_Importe_Total as double
precision),
CAST(case F.Factura_Estado_Conexion when 'ACTIVO' then 1 when 'INACTIVO' then 0 end as
int),CAST(case R.Reclamo_Resultado when 'SI' then 1 when 'NO' then 0 end as int) FROM
DIM_Pago P
INNER JOIN DIM_Factura F ON F.Factura_Nro_Documento=P.Pago_Nro
INNER JOIN DIM_Tiempo T ON T.Tiempo_Nro_Documento=P.Pago_Nro
INNER JOIN DIM_Cliente C ON C.Cliente_Nro_Usuario=P.Pago_Cliente
INNER JOIN DIM_Reclamo R ON R.Reclamo_Cliente=P.Pago_Cliente
end

```

Figura 4.4. Proceso ETL para cargar FACT\_MOROCIDAD

Mediante las figuras de tal a tal, mostramos la migración de la información digital proporcionada por la entidad a las dimensiones creadas para generar los cubos.

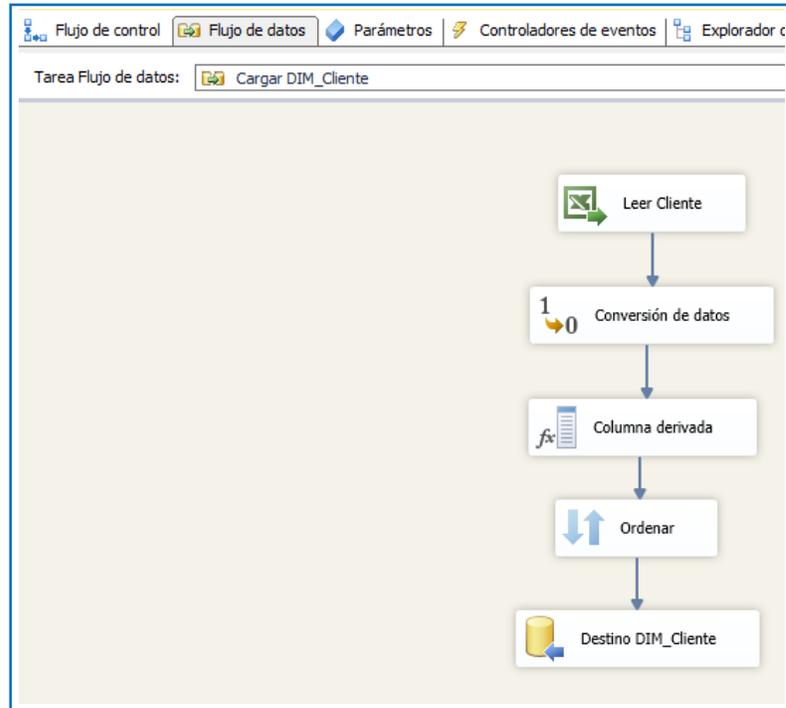


Figura 4.5. Proceso ETL para la tabla DIM\_Cliente

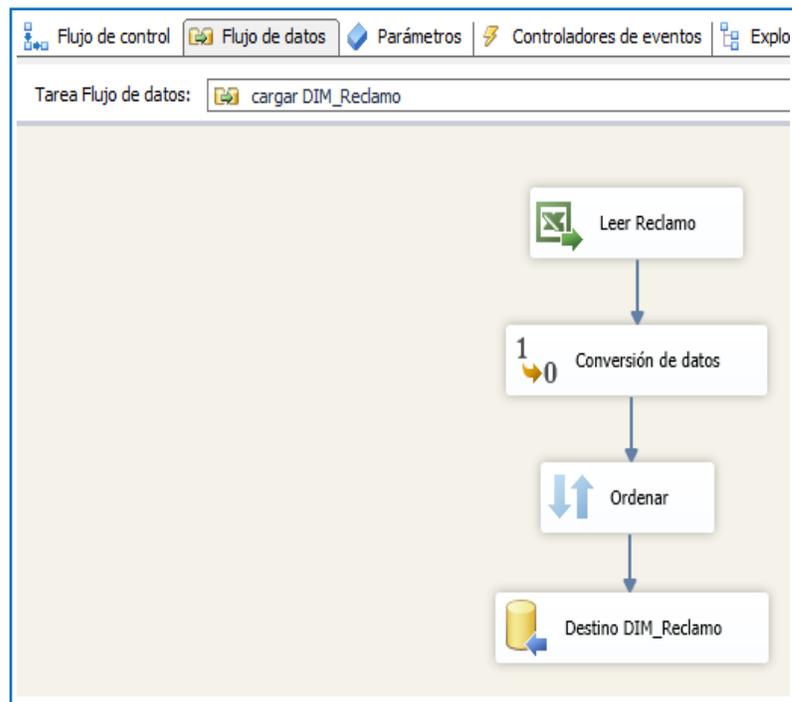


Figura 4.6. Proceso ETL para la tabla DIM\_Cliente

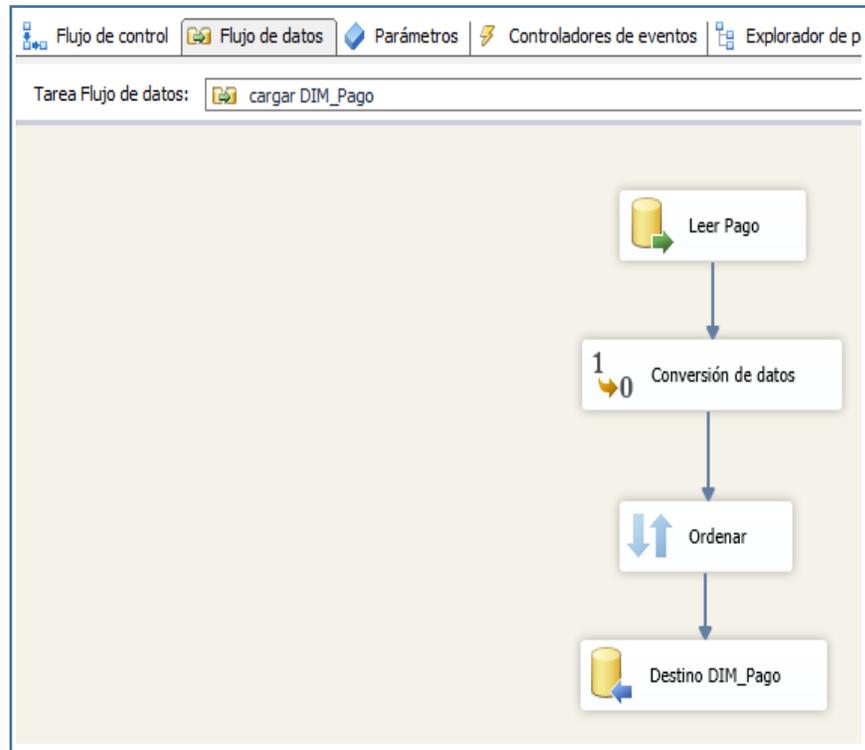


Figura 4.7. Proceso ETL para la tabla DIM\_Pago

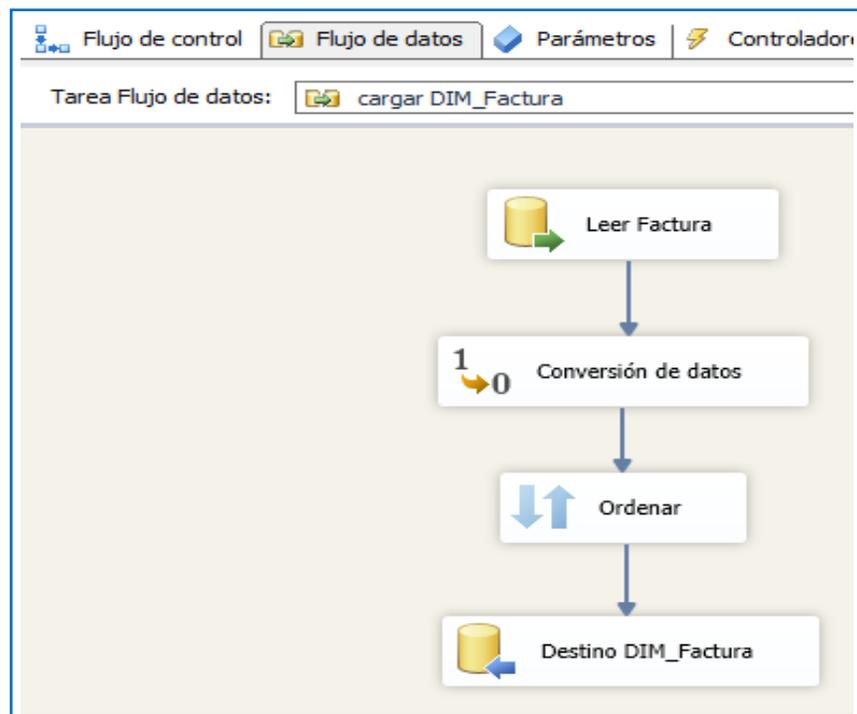


Figura 4.8. Proceso ETL para la tabla DIM\_Factura

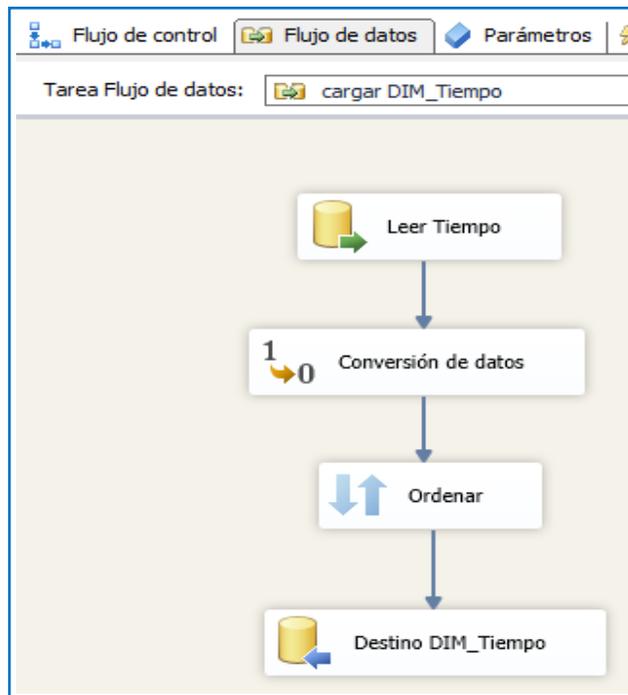


Figura 4.9. Proceso ETL para la tabla DIM\_Tiempo

## ACTUALIZACIÓN

Cuando se realice el cargado total del Datamart, se deben establecer sus políticas y estrategias de actualización o refresco de datos.

Una vez realizado esto, se tendrán que llevar a cabo las siguientes acciones:

- a) Especificar las tareas de limpieza de datos, calidad de datos, procesos ETL, etc., que deberán realizarse para actualizar los datos del DM.
- b) Especificar de forma general y detallada las acciones que deberá realizar cada software.

Las políticas que se establecen y las cuales se ha convenido con los responsables del area:

- a) La información se refrescará cada semestre.
- b) Los datos de la dimensión “tiempo” se cargarán de manera incremental teniendo en cuenta la fecha de la última actualización.
- c) Los datos de la tabla de hechos utilizados se cargarán de manera incremental teniendo en cuenta la fecha de la última actualización que corresponden al último mes a partir de la fecha.

#### 4.1.5. IMPLEMENTACIÓN DE CUBOS MULTIDIMENSIONALES

Se implementará el cubo multidimensional en Visual Studio 2012, es una herramienta parte del SQL SERVER, que permite trabajar proyectos de inteligencia de negocios.

#### CONFIGURAR EL ORIGEN DE DATOS

Realizará la conexión de los datos externos entre la plataforma de análisis de servicios con una plataforma externa, la interfaz nos pedirá una serie de parámetros que nos pedirá que mencionemos el servidor, así como el nombre de la de base datos “DB\_DataMart\_Morocidad”, para implementar el datamart.

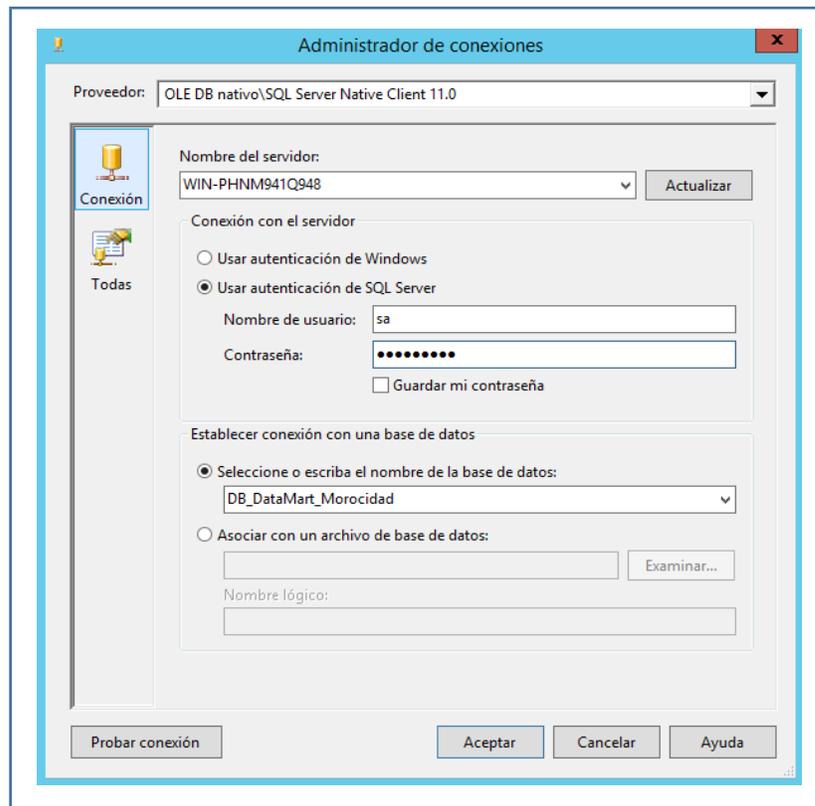


Figura 4.10. Administrador de conexiones

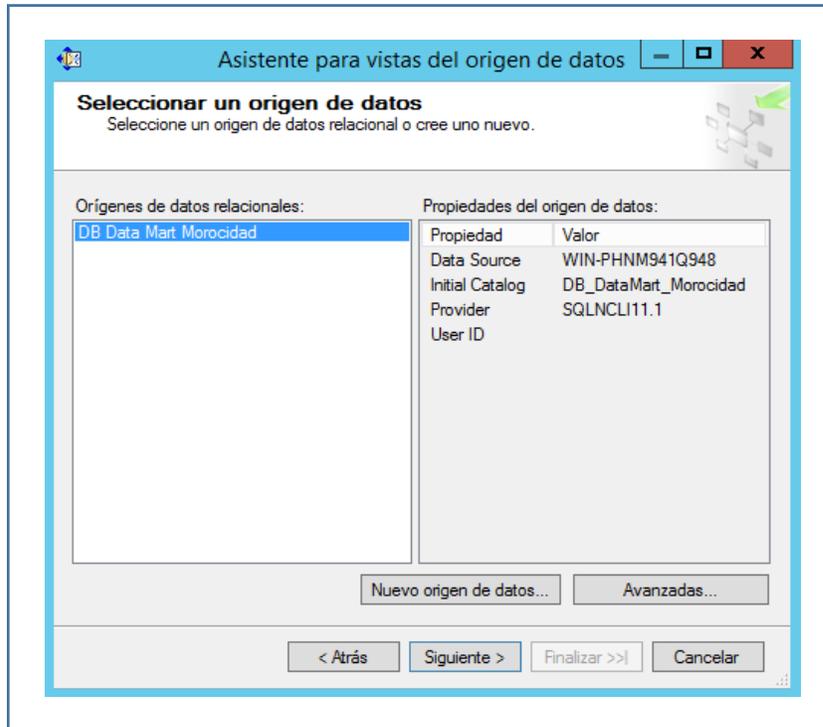


Figura 4.11. Iniciando la selección de la fuente de datos

## CONSTRUCCIÓN DE LA VISTA DE DATOS

Es una sección de la base de datos, que representa al conjunto de datos que forman parte de la vista del origen de datos, definiremos el origen de datos que queremos conectar, entonces optaremos el origen de datos que acabamos de realizar en el paso previo.

El asistente solicitará las tablas que van a formar parte de nuestra vista y cómo se observa, nos muestra todas las tablas de dimensiones, así como las tablas de hechos creadas en la fase de modelo lógico para la creación del cubo.

Dicho asistente permitirá unir las dimensiones y la tabla FACT\_Morocidad cuya importancia esta en los indicadores que muestra dentro de ella.

Esta combinación permitirá la fusión de toda la información cargada en en dicho asistente dentro de las vistas y tablas.

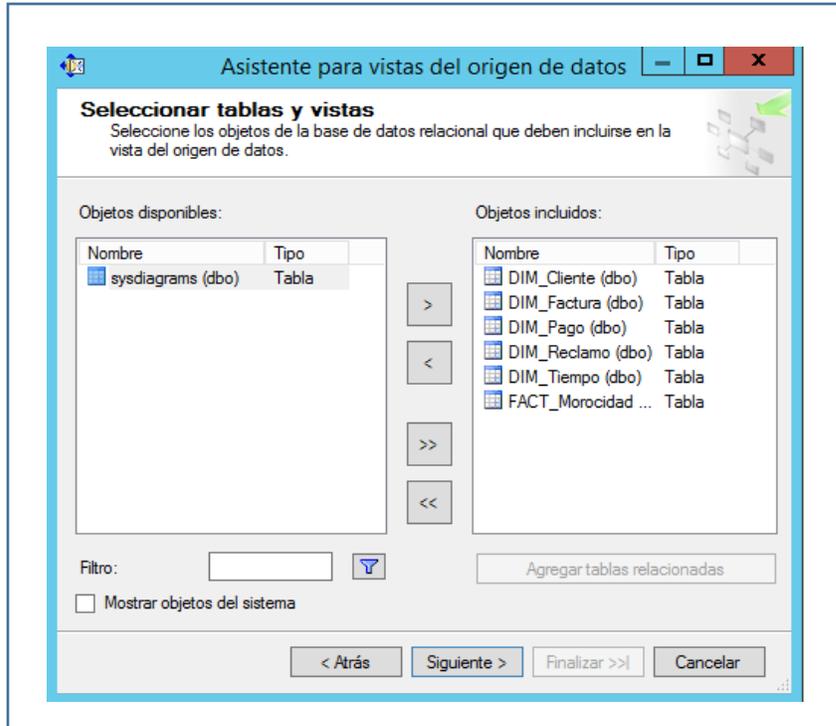


Figura 4.12. Seleccionando las dimensiones y la tabla principal

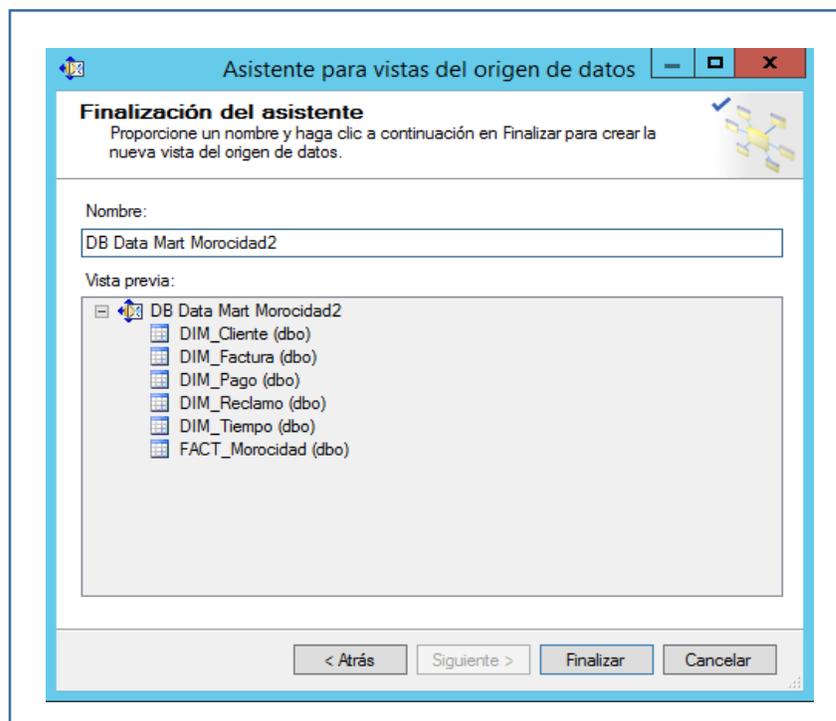


Figura 4.13. Tablas de dimensiones y la Principal

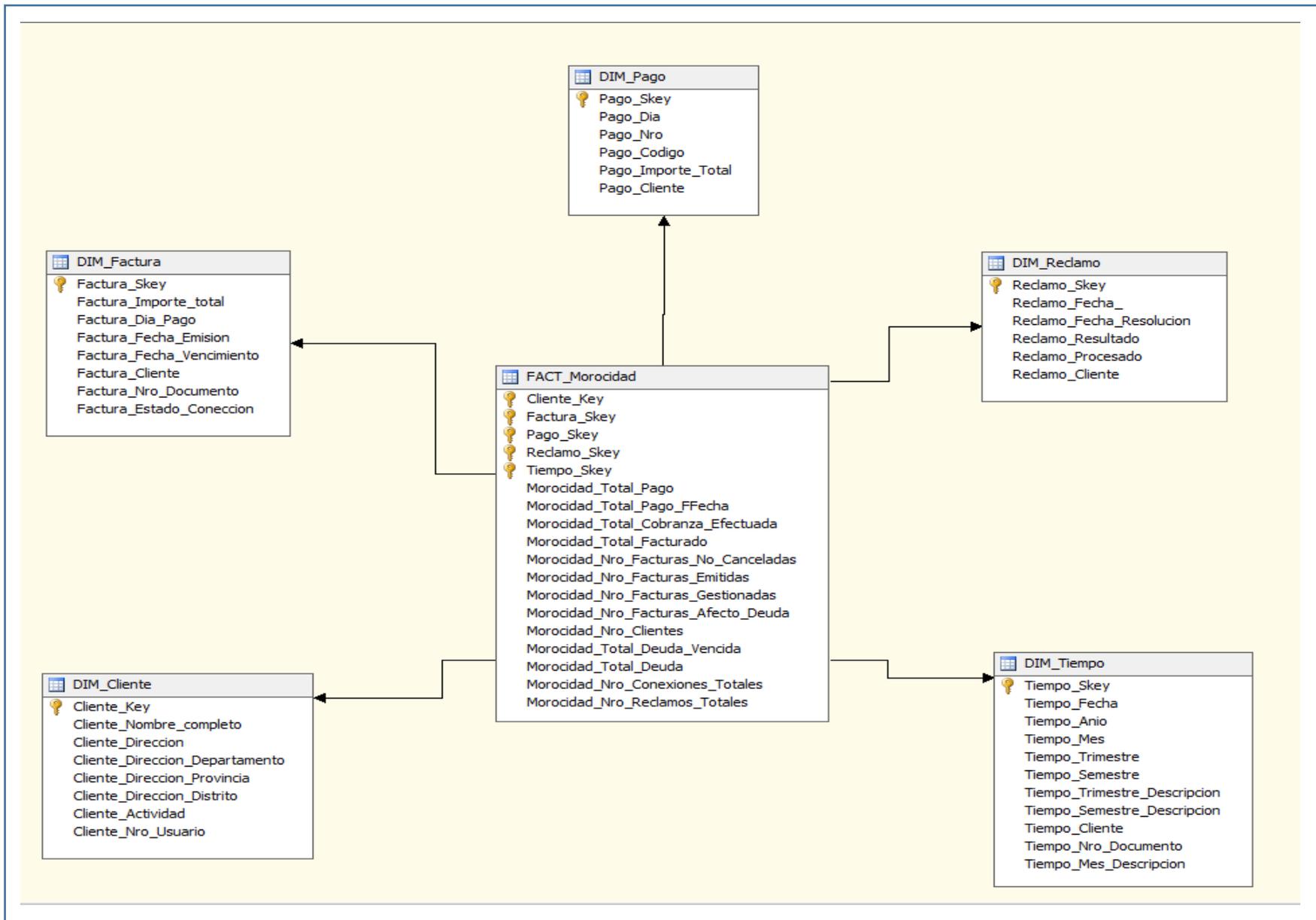


Figura 4.14. Diagrama de esquema de uniones

En la figura N° 4.14 se observa la primera vista creada, donde se encuentra las diferentes tablas de dimensiones y de hechos dentro de la tabla principal FACT\_Morocidad, que forman parte de nuestra vista, así como las diversas relaciones, entre las diferentes tablas de dimensiones y la tabla principal bajo el modelo estrella.

## DEFINICION DE LOS ATRIBUTOS DE LAS DIMENSIONES PARA EL CUBO

Se muestran tres secciones, la primera es la sección donde se encuentran los atributos que forman parte de las dimensiones, la segunda sección es la de jerarquías y la tercera es donde están las tablas fuentes la cual está contenida con la base de datos externa, de donde se extraen los atributos de la base de datos de las tablas de dimensión.

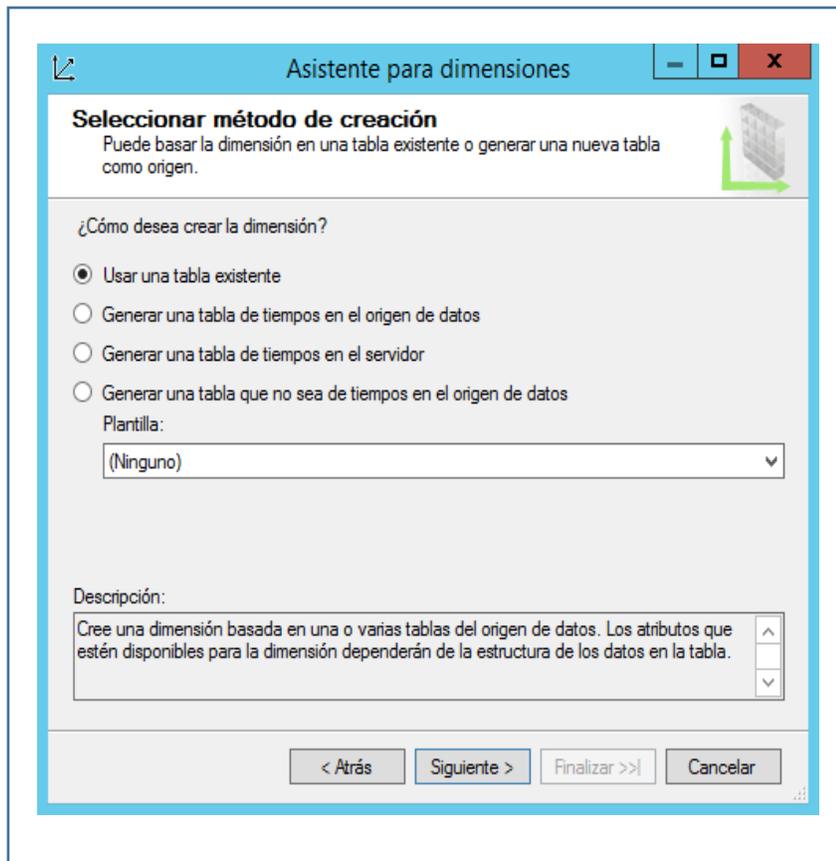


Figura 4.15. Asistente para dimensiones

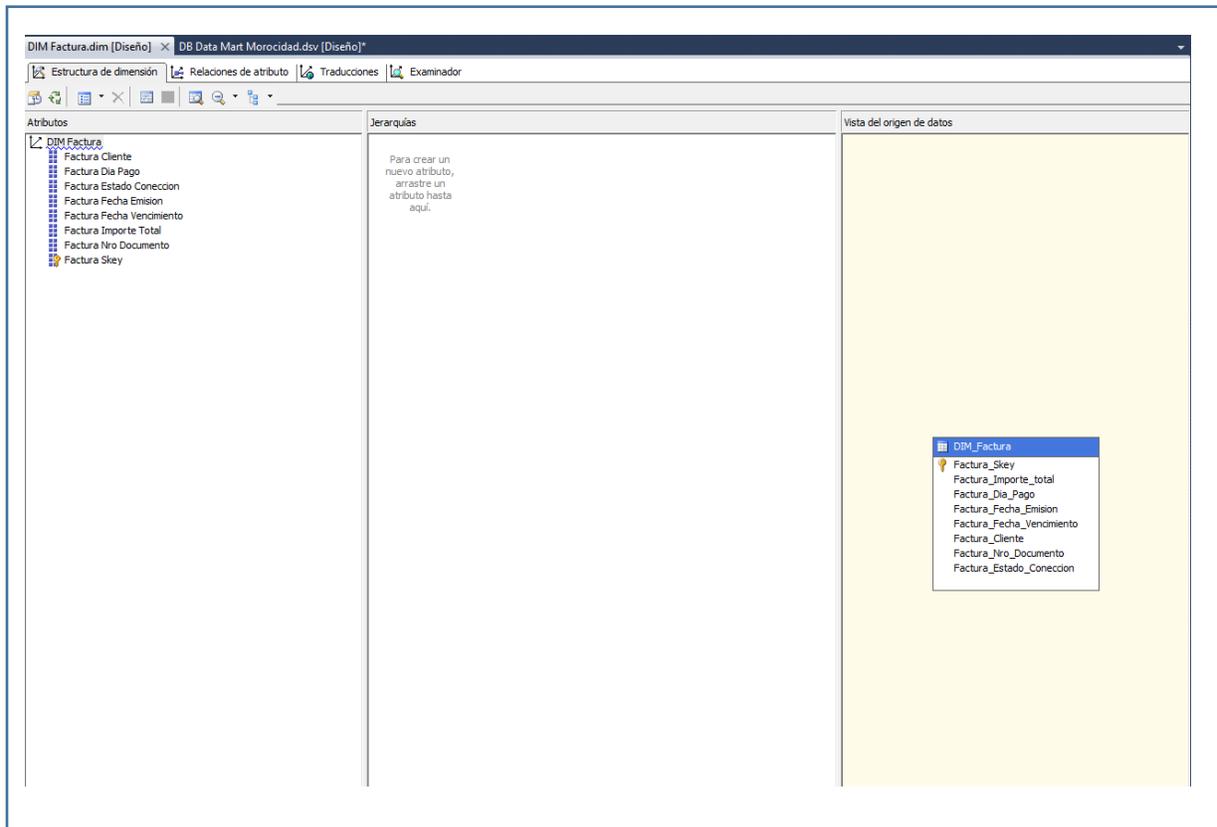


Figura 4.16. Definición de los atributos de la dimensión Factura

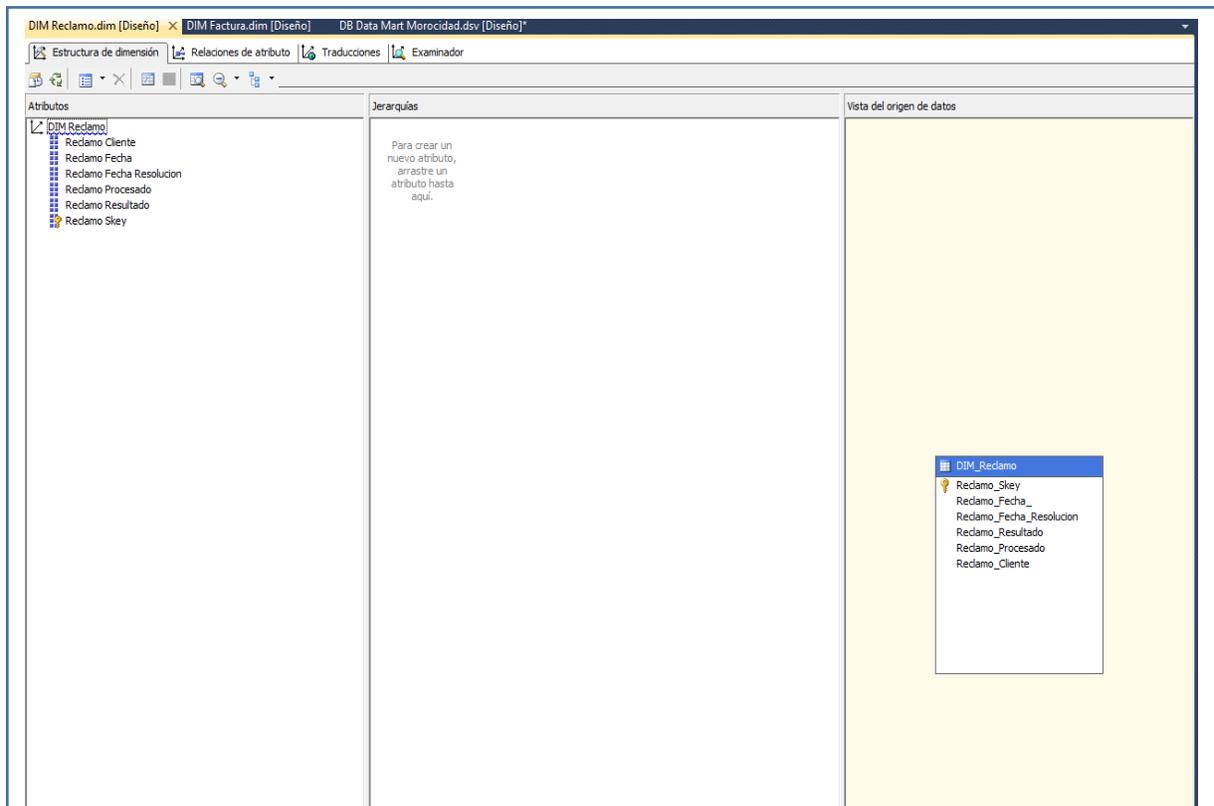


Figura 4.17. Definición de los atributos de la dimensión Reclamo

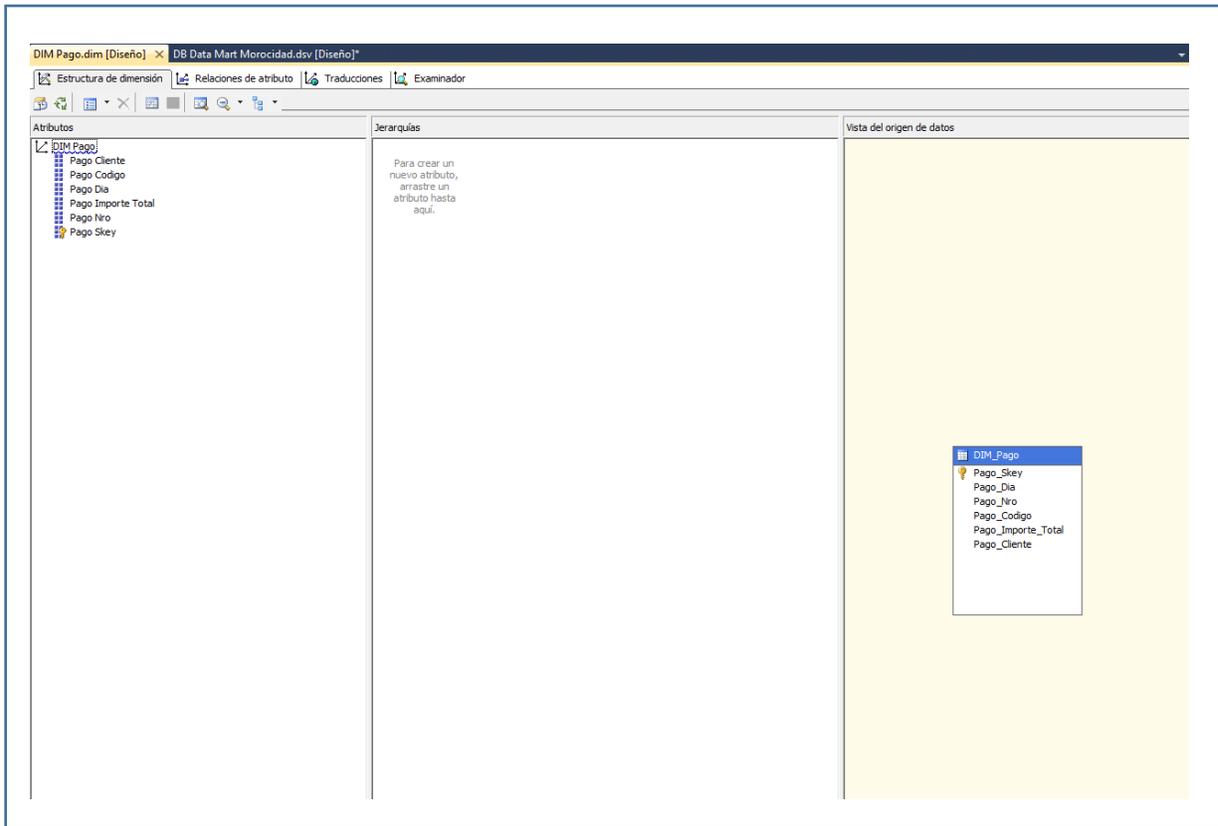


Figura 4.18. Definición de los atributos de la dimensión Pago

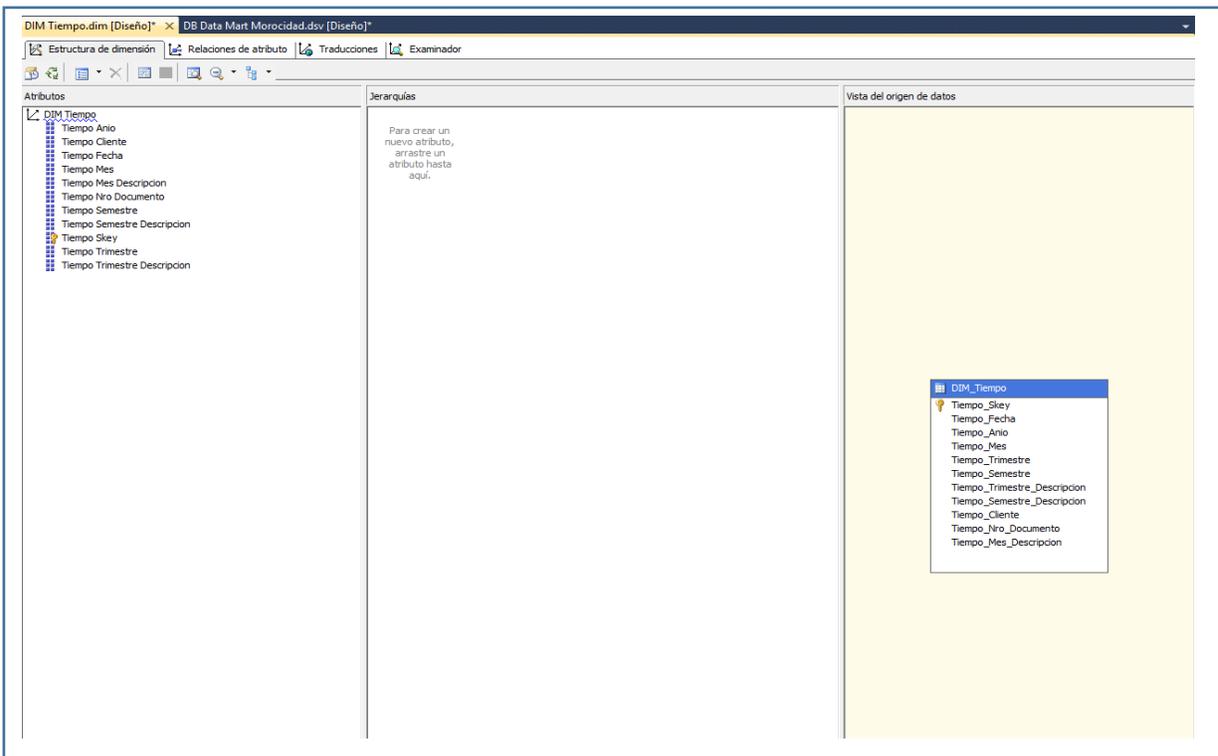


Figura 4.19. Definición de los atributos de la dimensión Tiempo

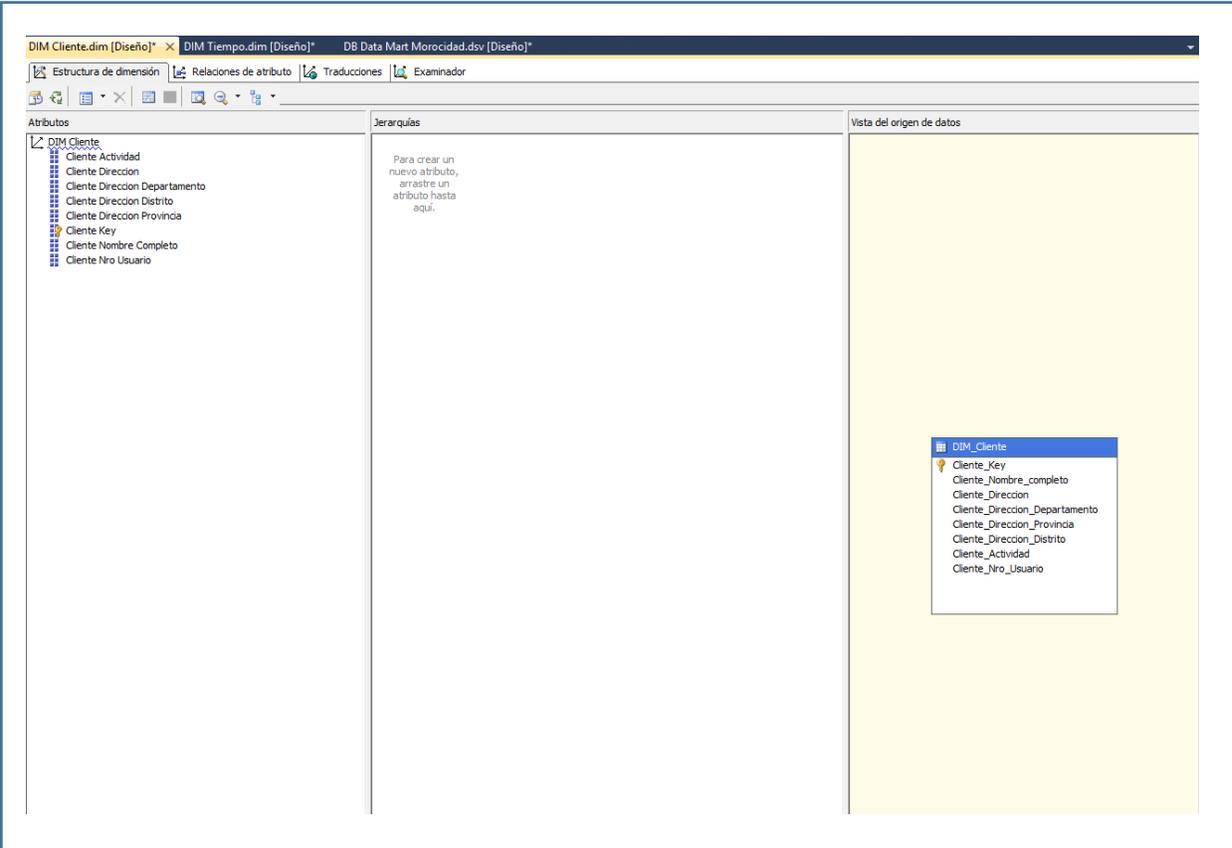


Figura 4.20. Definición de los atributos de la dimensión Cliente

**A. CONSTRUCCIÓN DEL CUBO MULTIDIMENSIONAL PARA CONSULTA , IND 1, IND 2, IND 3, IND 4, IND 5, IND 6, IND 7, IND 8.**

Nuestra vista construida en los pasos previos contiene toda la asociación de las diferentes tablas bajo el modelo estrella, para crear al cubo, necesitamos cargar la selección de tablas ya definidas, para la tabla principal.

Dicha carga seleccionada será actualizada mediante dicha herramienta para su procesamiento analítico y actualización de las mismas, la ejecución será positivo siempre y cuando la asociación sea la coorrecta respecto a las tablas seleccionadas.

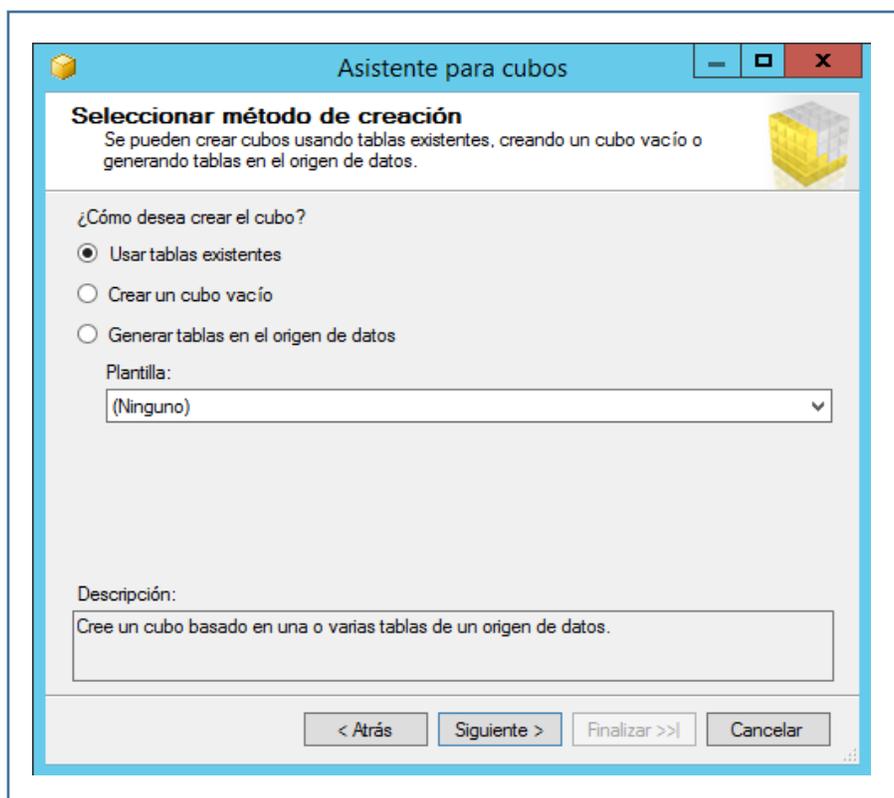


Figura 4.21. Asistente para creación de cubos

Se carga las tablas de hechos los cuales son: **Morocidad\_Total\_Pago**, **Morocidad\_Total\_Pago\_FFecha**, **Morocidad\_Cobranza\_Efectuada**, **Morocidad\_Total\_Facturado**, **Morocidad\_Nro\_Facturas\_No\_Canceladas**, **Morocidad\_Nro\_Facturas\_Emitidas**, **Morocidad\_Nro\_Facturas\_Gestionadas**, **Morocidad\_Nro\_Facturas\_Afecto\_Deuda**, **Morocidad\_Nro\_Clientes**, **Morocidad\_Total\_Deuda\_Vencida**, **Morocidad\_Nro\_Conexiones\_Totales**, **Morocidad\_Nro\_Reclamos\_Totales**, donde encontraremos los valores numéricos obtenidos de la Base de Datos.

Dicha carga será visualizada en la tabla principal fact-Morosidad cuya importancia radica en el análisis de dimensiones ya creadas y en la tabla de hecho principal.

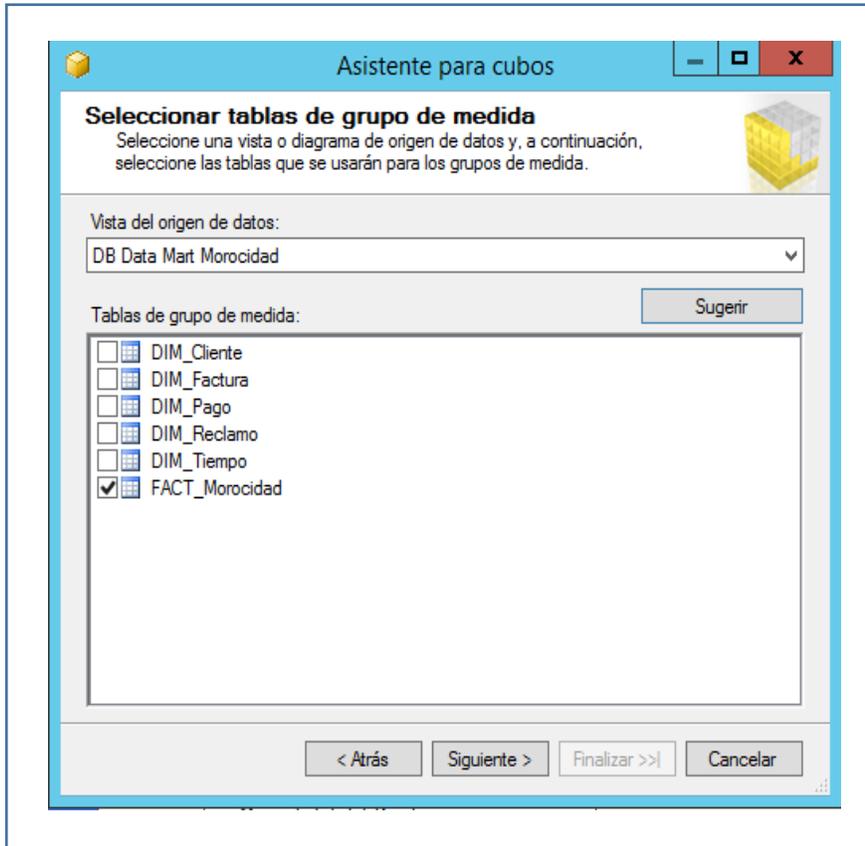


Figura 4.22. Selección del grupo de medida

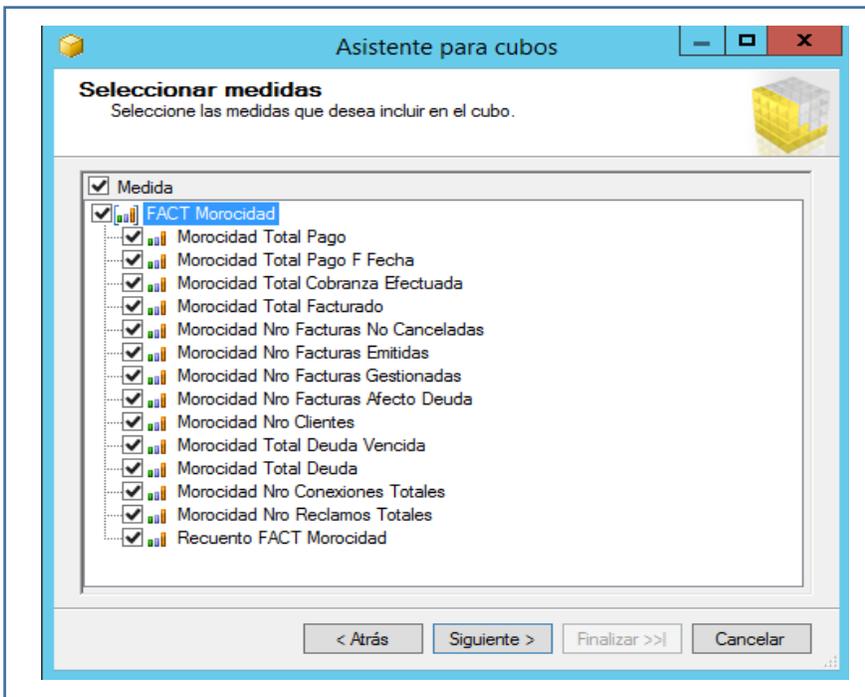


Figura 4.23. Creacion del cubo con el grupo de medida

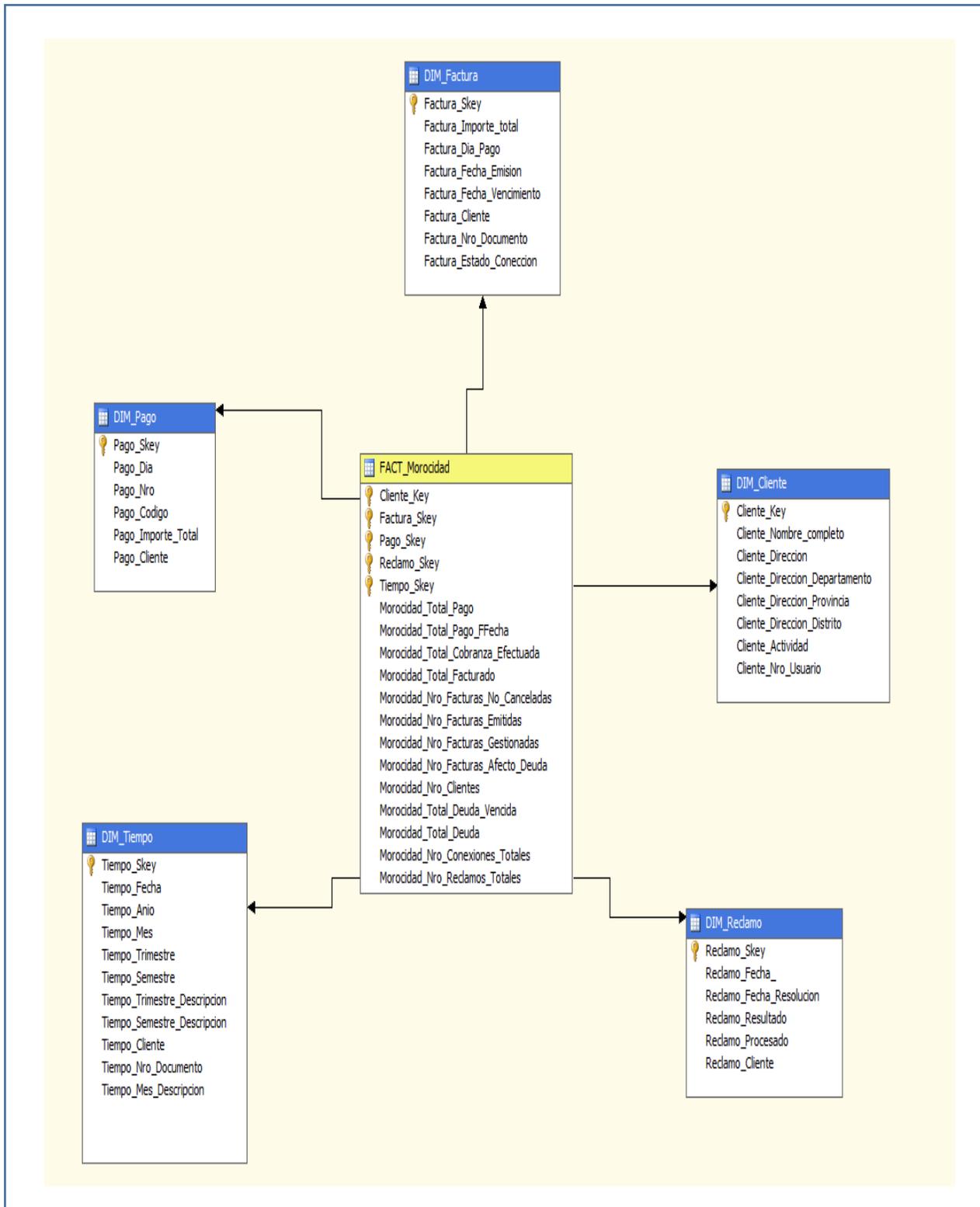


Figura 4.24. Configuración final de cubo con dimensiones y tabla de hechos.

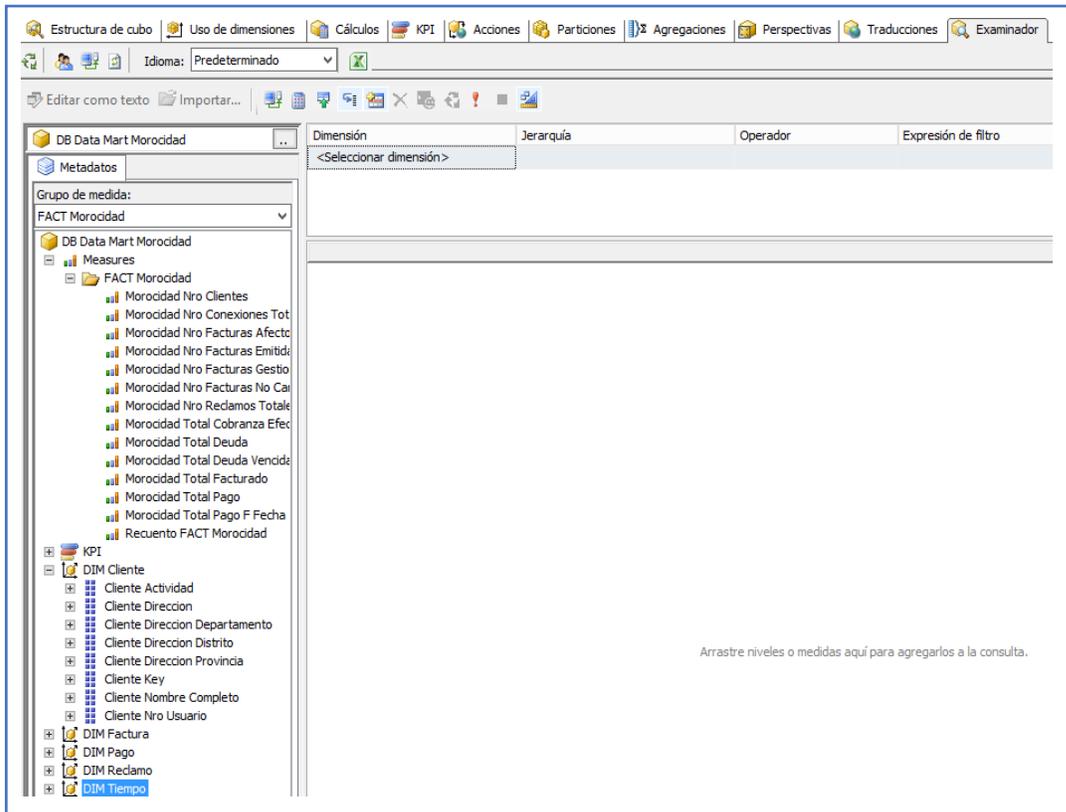


Figura 4.25. Panel de consultas del cubo creado

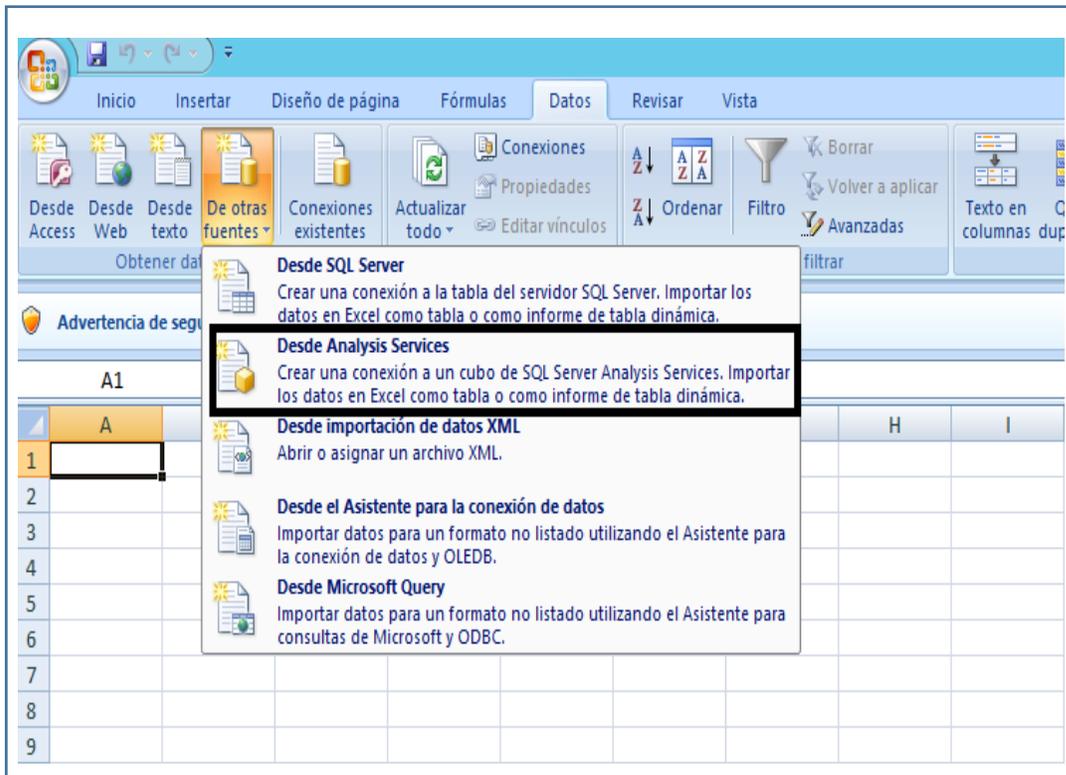


Figura 4.26. Importación del cubo mediante analysis services al excel

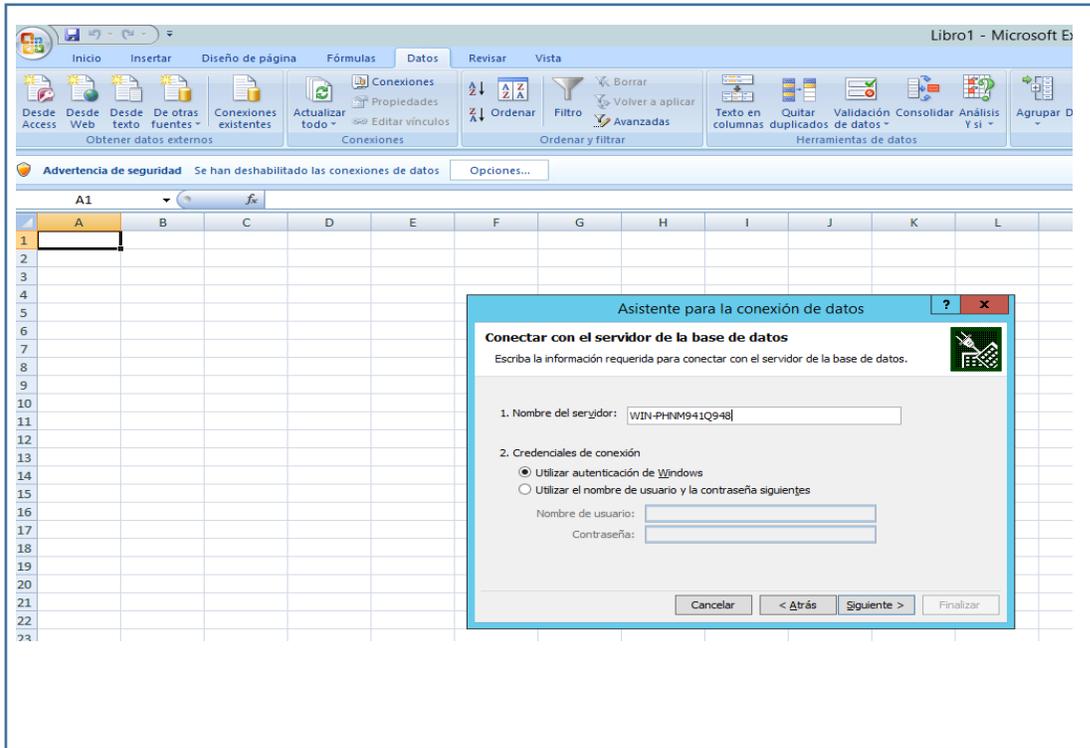


Figura 4.27. Asistente para la conexión de datos

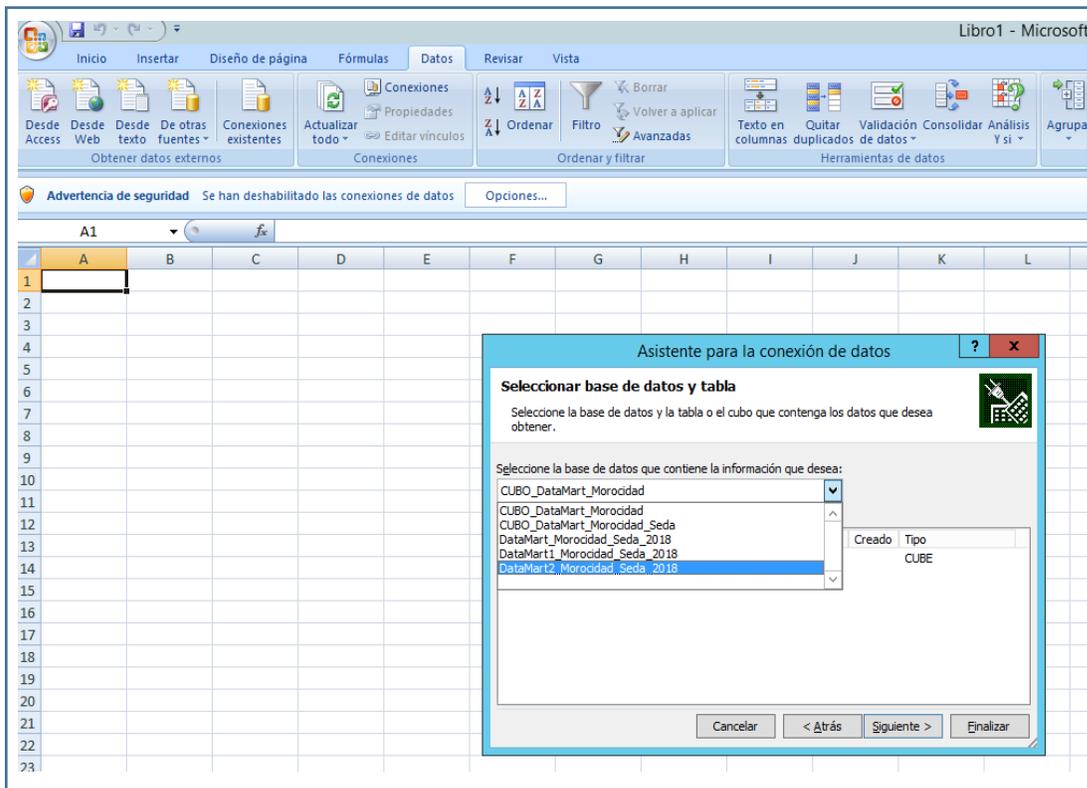


Figura 4.28. Selección de la base de datos creado

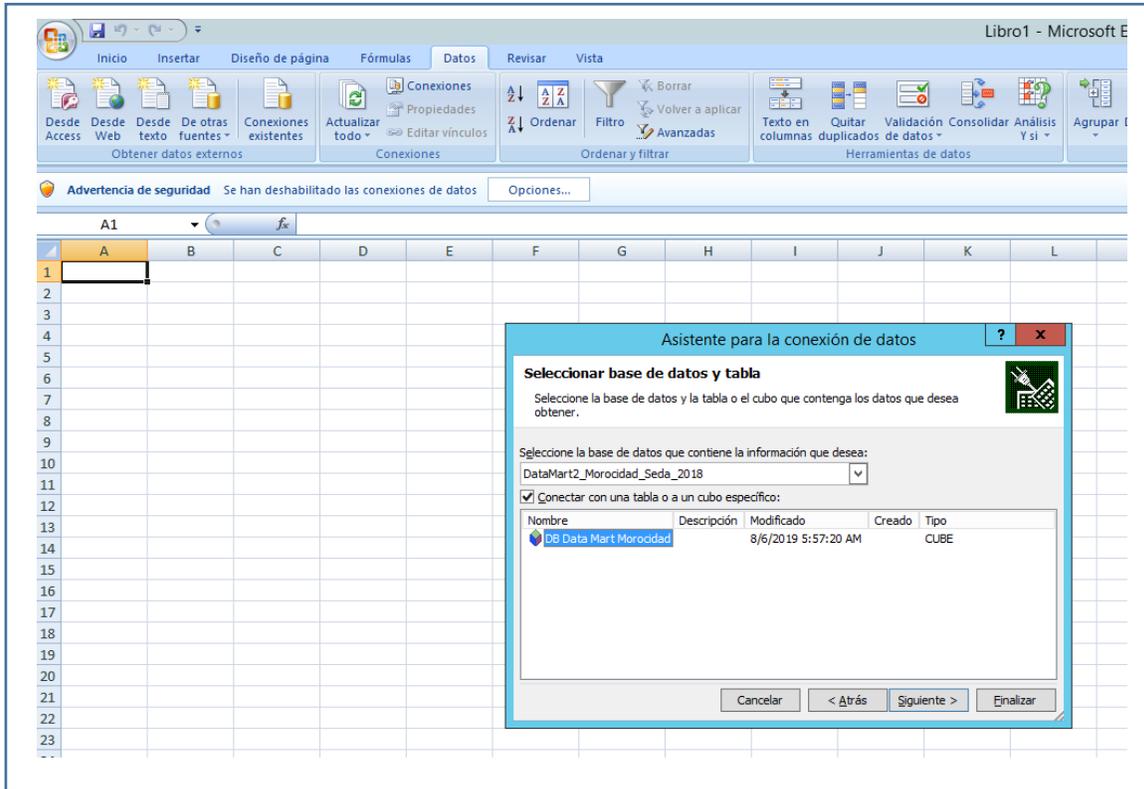


Figura 4.29. Selección del cubo ya creado

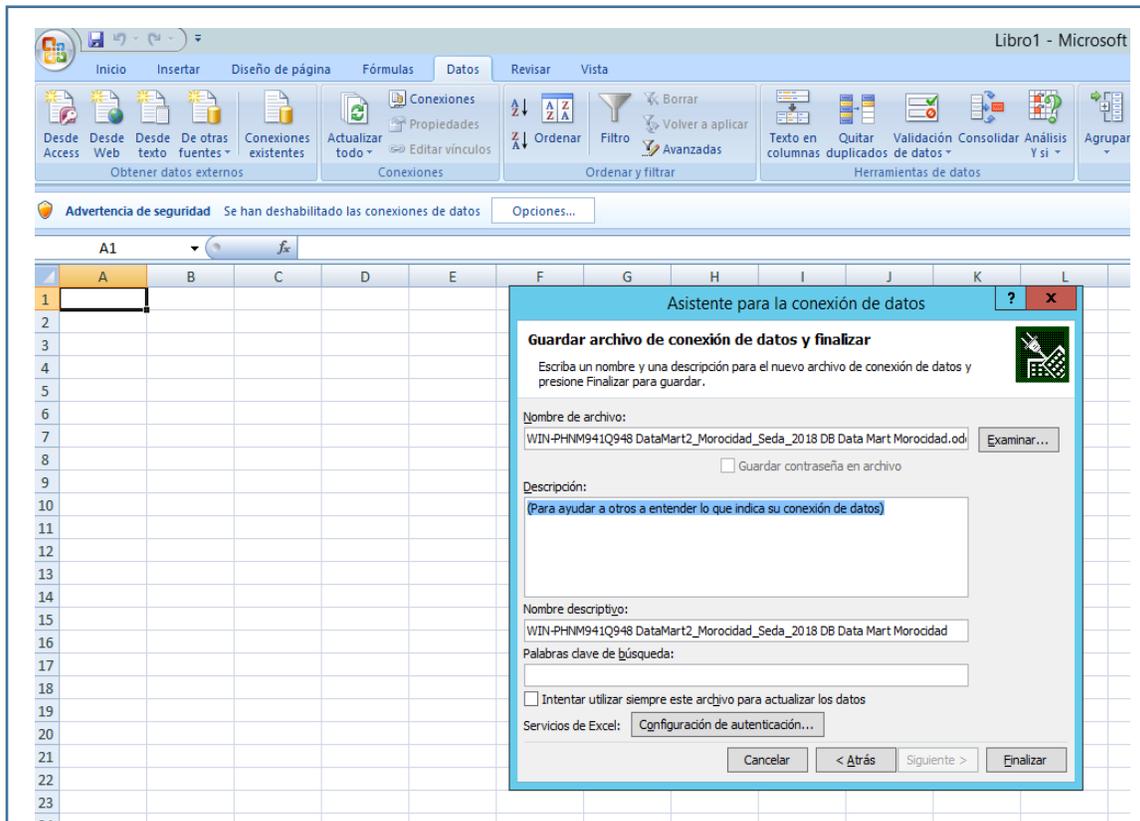


Figura 4.30. Asistente de confirmación de la conexión de datos

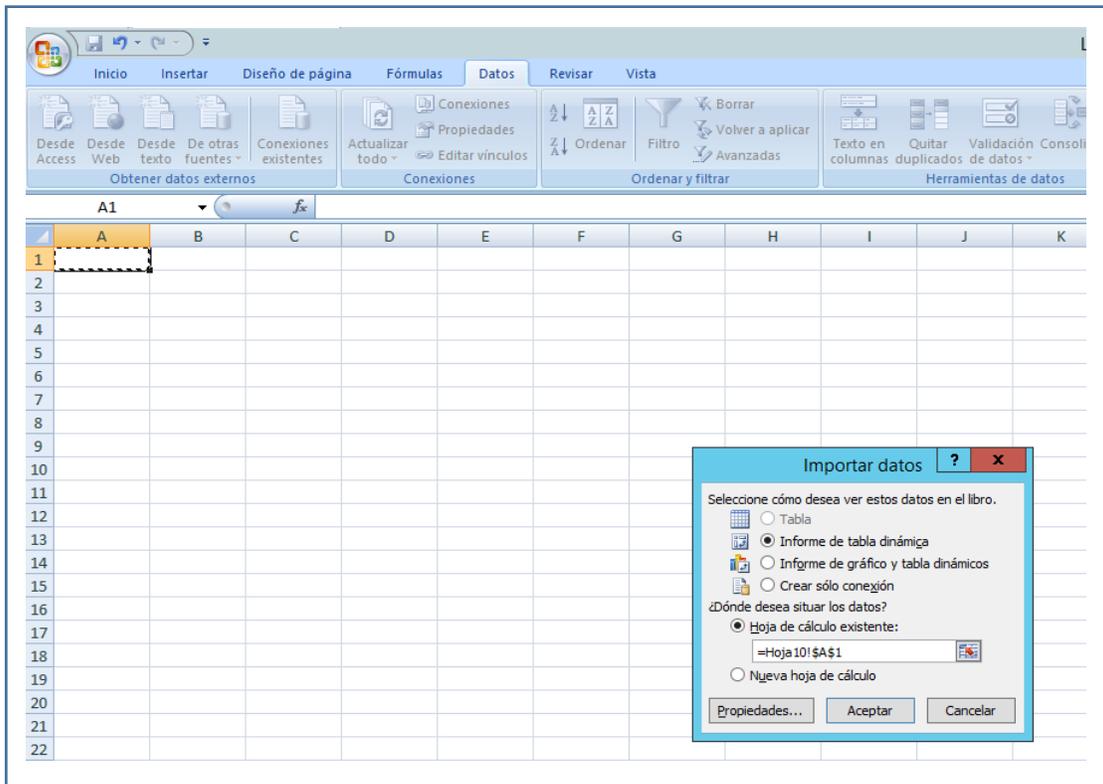


Figura 4.31. Selección del tipo de reporte de datos

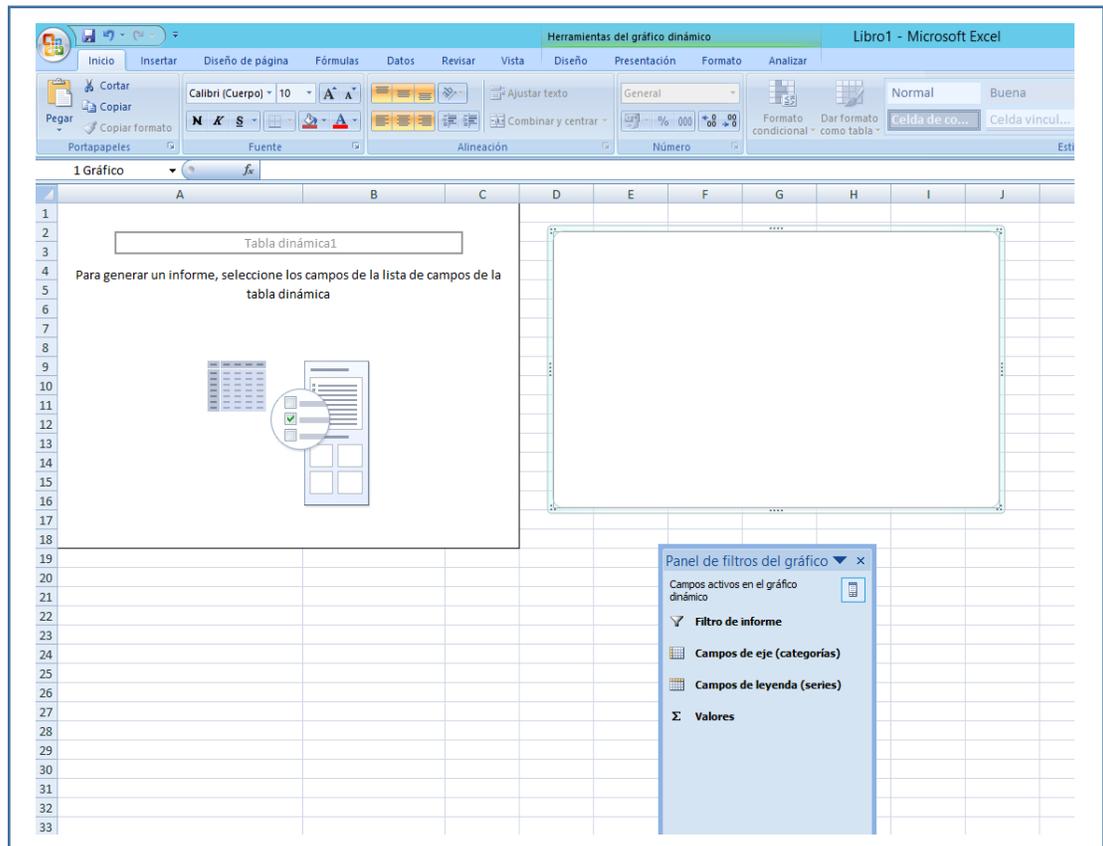


Figura 4.32. Panel de reportes del cubo en excel

**RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EXCEL POR EL INDICADOR PAGOS FUERA DEL PLAZO PROMEDIO DE COBRANZA:**

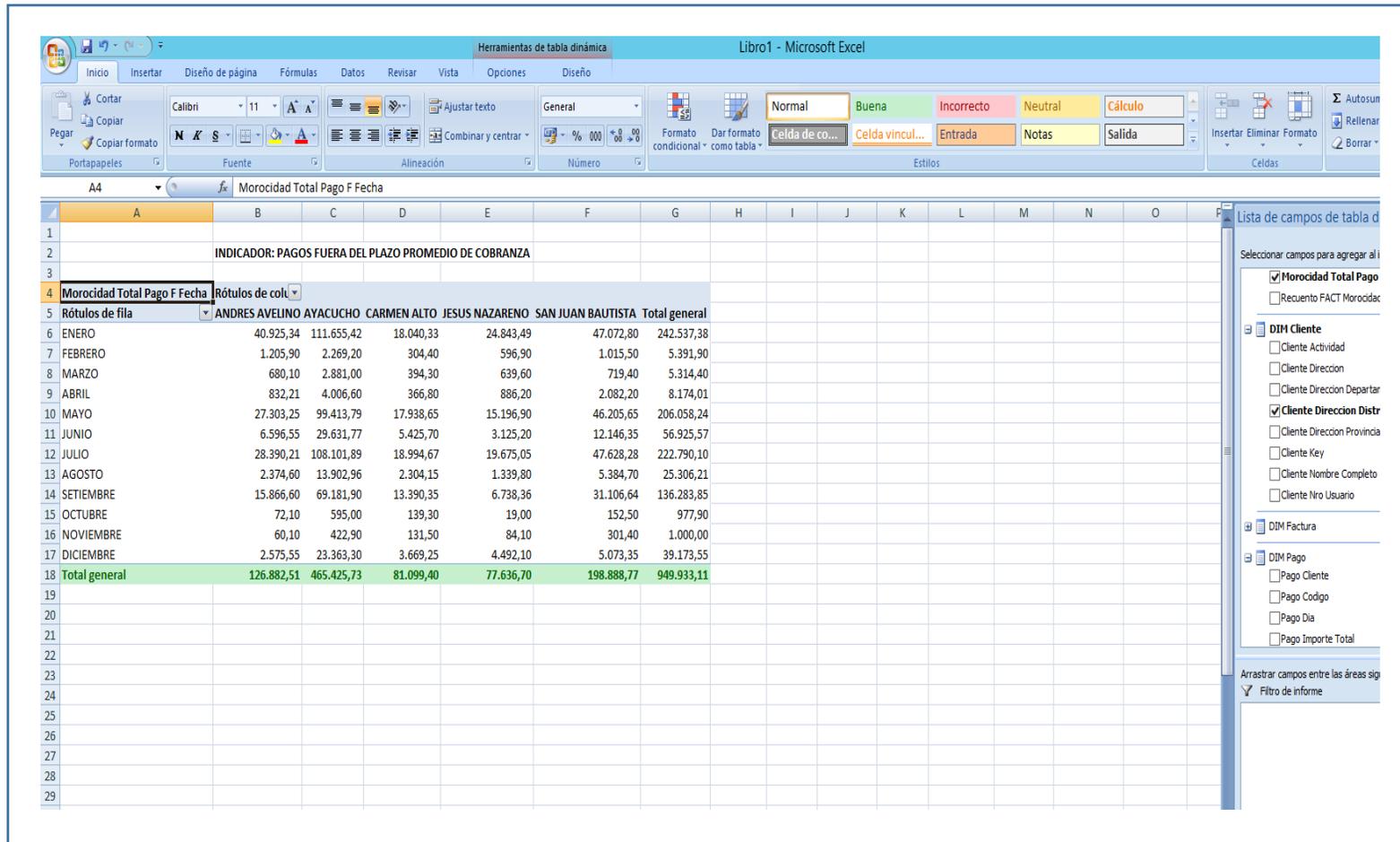


Figura 4.33. Pago total fuera de fecha por distritos y mes

## RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EXCEL POR EL INDICADOR EFICIENCIA DE COBRANZA:

Advertencia de seguridad Se han deshabilitado las conexiones de datos Opciones...

Rótulos de columna	Rótulos de fila	ACTIVO	Total general
	01	14211070,43	14211070,43
	PRIMER SEMESTRE	14211070,43	14211070,43
	ABRIL	2285528,74	2285528,74
	ENERO	2232606,81	2232606,81
	FEBRERO	2454497,56	2454497,56
	JUNIO	2438628,77	2438628,77
	MARZO	2416453,36	2416453,36
	MAYO	2383355,19	2383355,19
	02	14910104,06	14910104,06
	SEGUNDO SEMESTRE	14910104,06	14910104,06
	AGOSTO	2453868,28	2453868,28
	DICIEMBRE	2453951,71	2453951,71
	JULIO	2229420,55	2229420,55
	NOVIEMBRE	2595924,5	2595924,5
	OCTUBRE	2696789,12	2696789,12
	SEPTIEMBRE	2480149,9	2480149,9
	Total general	29121174,49	29121174,49

Lista de campos de tabla dinámica

Selecconar campos para agregar al informe

Σ FACT Morosidad

- Morosidad Nro Clientes
- Morosidad Nro Conexiones
- Morosidad Nro Facturas Af
- Morosidad Nro Facturas Entr
- Morosidad Nro Facturas Ge
- Morosidad Nro Facturas No
- Morosidad Nro Reclamos Tr
- Morosidad Total Cobranza
- Morosidad Total Deuda
- Morosidad Total Deuda Ver
- Morosidad Total Facturado
- Morosidad Total Pago
- Morosidad Total Pago F Fe
- Recuento FACT Morosidad

DIM Cliente

- Cliente Actividad

Figura 4.34. Cantidad total de la cobranza efectuada por trimestre respecto al estado de conexion.

**RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EXCEL POR EL INDICADOR COBRANZA GESTIONADA:**

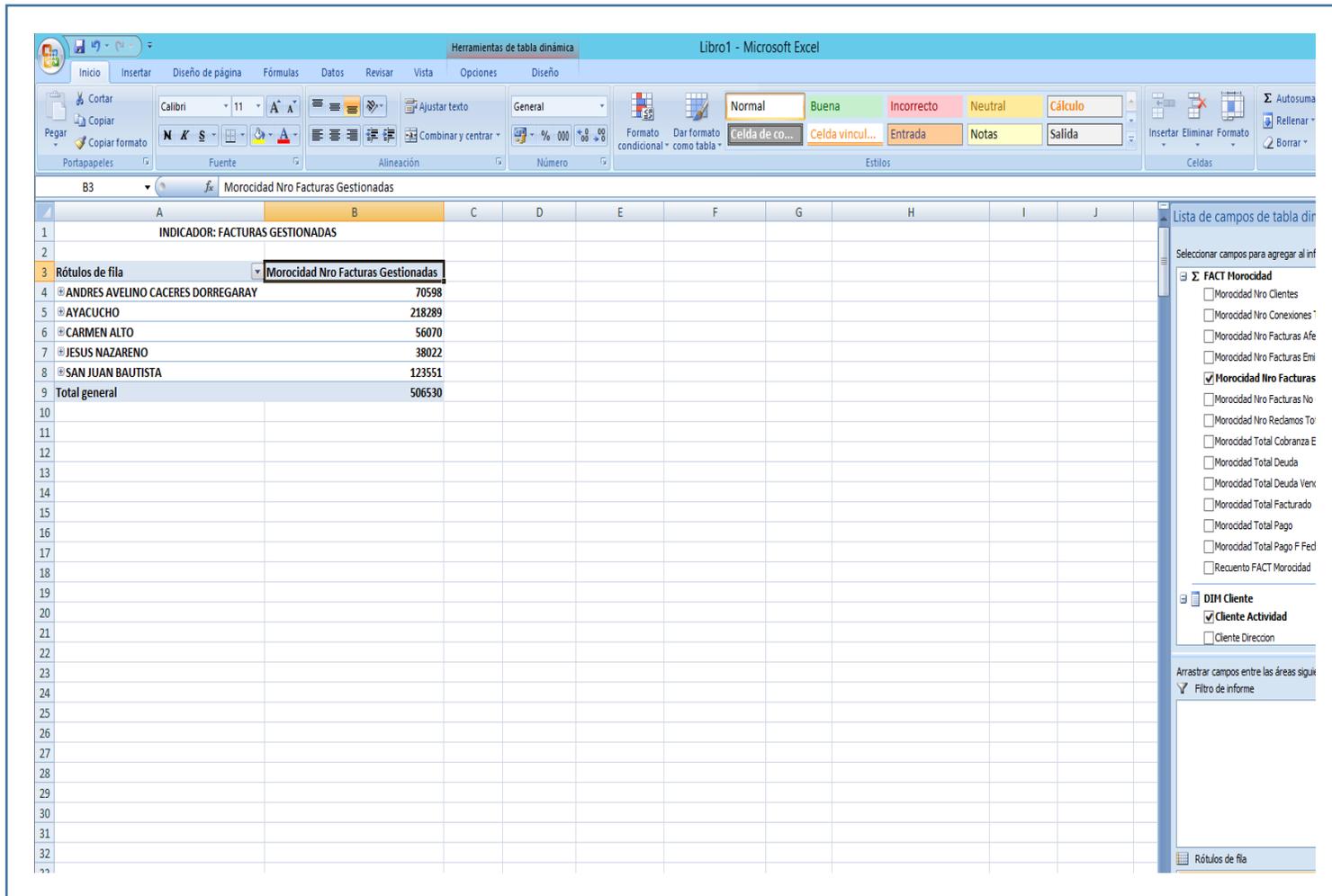


Figura 4.35. Cantidad de facturas gestionadas por distrito

**RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EXCEL POR EL INDICADOR CLIENTES GESTIONADOS POR DEUDA:**

Morocidad Nro Clientes	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	Total general
ANDRES AVELINO CACERES DORREGARAY	6609	6620	6608	6595	6728	6755	6794	6818	6808	6849	6866	6878	80928
AYACUCHO	20723	20722	20737	20686	20842	20914	20994	20988	21014	21122	21141	21151	251034
CARMEN ALTO	5273	5267	5292	5286	5395	5462	5444	5474	5474	5492	5517	5531	64907
JESUS NAZARENO	3622	3624	3633	3623	3653	3673	3678	3678	3680	3691	3699	3702	43956
SAN JUAN BAUTISTA	11711	11702	11705	11682	11822	11829	11883	11906	11937	11939	11962	11965	142043
<b>Total general</b>	<b>47938</b>	<b>47935</b>	<b>47975</b>	<b>47872</b>	<b>48440</b>	<b>48633</b>	<b>48793</b>	<b>48864</b>	<b>48913</b>	<b>49093</b>	<b>49185</b>	<b>49227</b>	<b>582868</b>

Figura 4.36. Cantidad de clientes facturados al mes por cada distrito.

**RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EXCEL POR EL INDICADOR PAGOS EFECTUADOS:**

The screenshot shows an Excel PivotTable with the following data:

Rótulos de fila	ACTIVO	Total general
ANDRES AVELINO CACERES DORREGARAY	4370505,93	4370505,93
AYACUCHO	14745355,56	14745355,56
CARMEN ALTO	2300048,59	2300048,59
JESUS NAZARENO	2132562,6	2132562,6
SAN JUAN BAUTISTA	5572701,81	5572701,81
<b>Total general</b>	<b>29121174,49</b>	<b>29121174,49</b>

Figura 4.37. Pago total anual por distrito respecto al estado de conexión.

## RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EXCEL POR EL INDICADOR DEUDA VENCIDA:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2		INDICADOR:MOROCIDAD TOTAL DEUDA												
3														
4		Morocidad Total Deuda												
5		Rótulos de columna												
5		Rótulos de fila												
5		ANDRES AVELINO	CACERES DORREGARAY	AYACUCHO	CARMEN ALTO	JESUS NAZARENO	SAN JUAN BAUTISTA	Total general						
6		PRIMER TRIMESTRE												
7	01	40925,34	111655,42	18040,33	24843,49	47072,8	242537,38							
8	02	1205,9	2269,2	304,4	596,9	1015,5	5391,9							
9	03	680,1	2881	394,3	639,6	719,4	5314,4							
10		SEGUNDO TRIMESTRE												
11	04	832,21	4006,6	366,8	886,2	2082,2	8174,01							
12	05	27303,25	99413,79	17938,65	15196,9	46205,65	206058,24							
13	06	6596,55	29631,77	5425,7	3125,2	12146,35	56925,57							
14		TERCER TRIMESTRE												
15	07	28390,21	108101,89	18994,67	19675,05	47628,28	222790,1							
16	08	2374,6	13902,96	2304,15	1339,8	5384,7	25306,21							
17	09	15866,6	69181,9	13390,35	6738,36	31106,64	136283,85							
18		CUARTO TRIMESTRE												
19	10	72,1	595	139,3	19	152,5	977,9							
20	11	60,1	422,9	131,5	84,1	301,4	1000							
21	12	2575,55	23363,3	3669,25	4492,1	5073,35	39173,55							
22		Total general												
22		126882,51	465425,73	81099,4	77636,7	198888,77	949933,11							

Figura 4.38. Morocidad de la deuda vencida por trimestre en cada distrito de la región.

**RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EXCEL POR EL INDICADOR RECLAMOS TOTALES:**

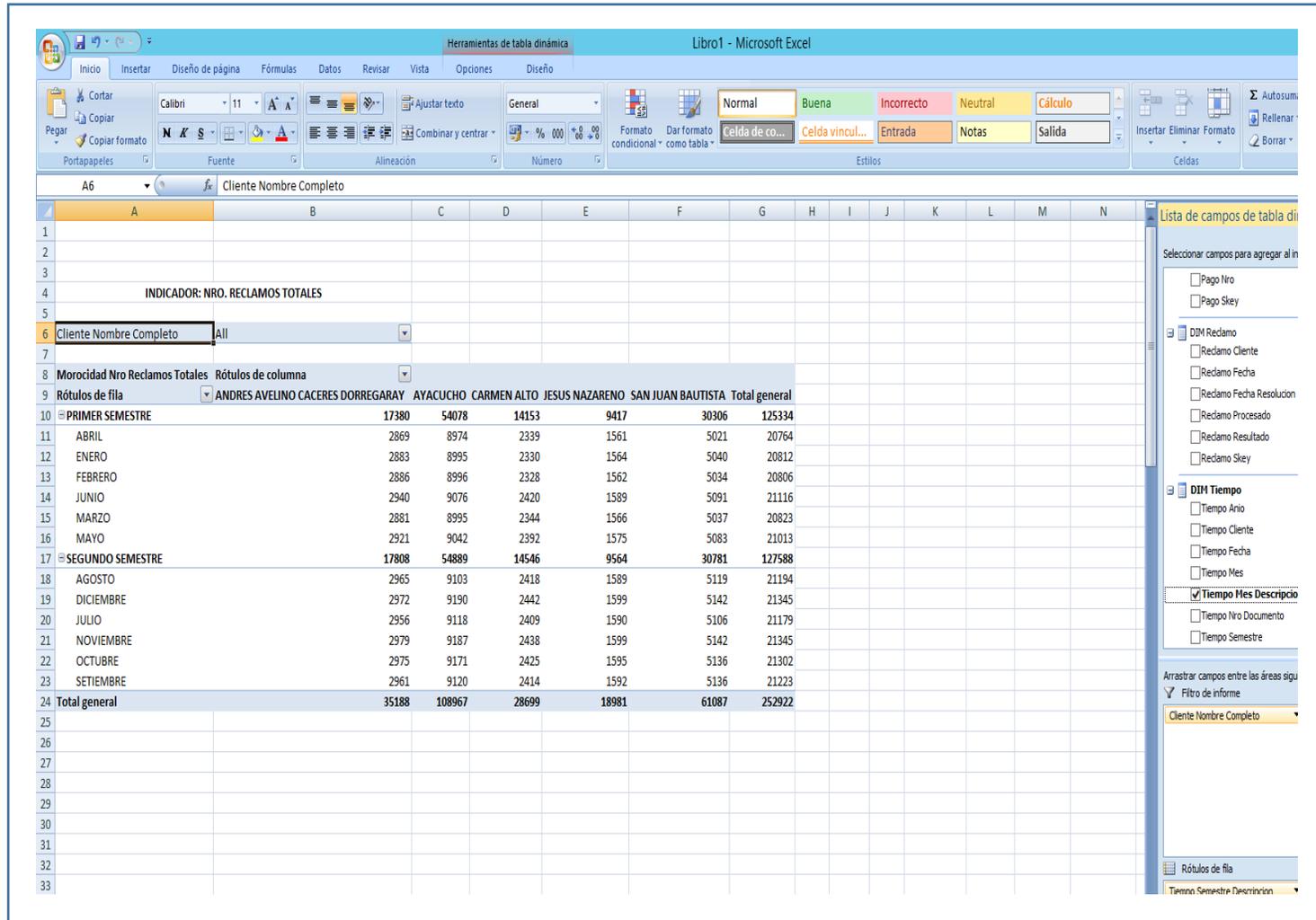


Figura 4.39. Cantidad de reclamos totales por distrito durante el año respecto a cada usuario

**RESULTADOS OBTENIDOS DEL CUBO MULTIDIMENSIONAL EN EXCEL POR EL INDICADOR CONEXIONES ACTIVAS:**

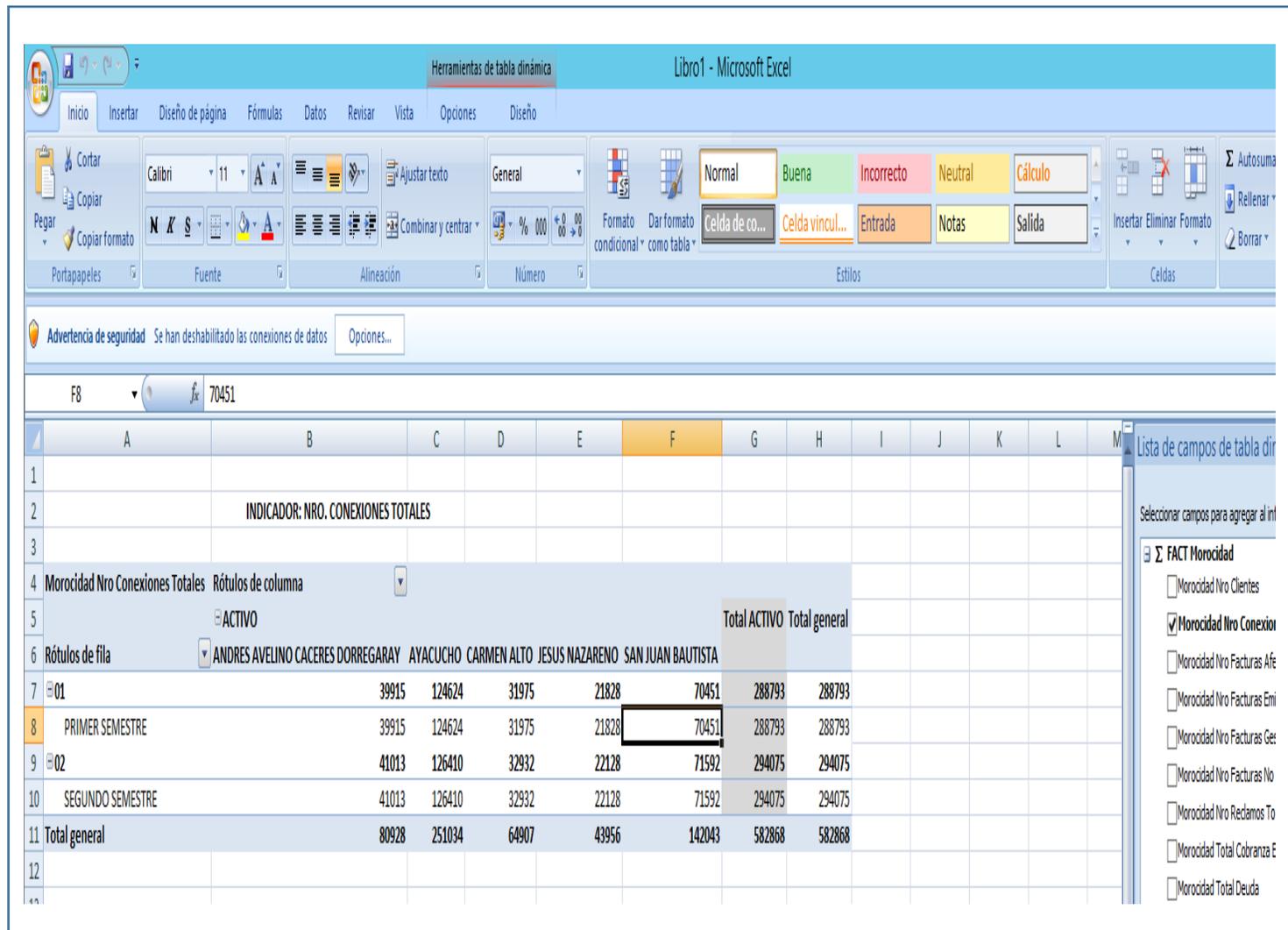


Figura 4.40. Cantidad de conexiones totales por distrito durante el año respecto al estado de conexion

## **CAPITULO V**

### **CONCLUSIONES Y RECOMENDACIONES**

#### **5.1. CONCLUSIONES**

- a) De acuerdo al marco teórico del capítulo II, sección 2.2.5, se construye el método para el DataMart mostrado en el capítulo III, sección 3.7; usando el método para desarrollar el DataMart, se ha logrado los resultados que se muestra en el capítulo IV, figuras del 4.33 al 4.35, que brinda información táctica respecto a la cobranza.
- b) De acuerdo al marco teórico del capítulo II, sección 2.2.5, se construye el método para el DataMart mostrado en el capítulo III, sección 3.7; usando el método para desarrollar el DataMart, se ha logrado los resultados que se muestra en el capítulo IV, figuras del 4.36 al 4.38, que brinda información táctica respecto a la deuda.
- c) De acuerdo al marco teórico del capítulo II, sección 2.2.5, se construye el método para el DataMart mostrado en el capítulo III, sección 3.7; usando el método para desarrollar el DataMart, se ha logrado los resultados que se muestra en el capítulo IV, figuras del 4.39 al 4.40, que brinda información táctica respecto al cliente.

#### **5.2. RECOMENDACIONES**

- a) Se recomienda realizar una investigación para obtener información táctica de todos las áreas de la entidad en estudio, de esta manera analizar indicadores que permitan describir deficiencias a nivel institucional.
- b) Se recomienda documentar bien el uso de las herramientas y realizar pruebas antes de iniciar el uso de producción de estas. Esto con la intencion de implementar una aplicación móvil que permita visualizar en tiempo real la morosidad de los usuarios de la entidad.
- c) Se recomienda realizar una aplicación web, que muestre los indicadores de morosidad de clientes, usando la metodología Hefesto.

## BIBLIOGRAFÍA

1. Bernabeu, R. D. (2010), Metodología Hefesto (5taEd.). Córdoba, Argentina: Editorial Tierra del sur.
2. Benites, J. D., & López, S. C. (2015). Solución de inteligencia de negocio para empresas de servicio de importación y exportación de calzado. Tesis de Pregrado. Universidad Privada Antenor Orrego, Trujillo.
3. Castañeda, E. E., & Tamayo, J. J. (2013). La morosidad y su impacto en el cumplimiento de los objetivos estratégicos de la agencia real plaza de la caja municipal de ahorro y crédito de Trujillo. Tesis de Pregrado. Universidad Privada Antenor Orrego, Trujillo, Perú.
4. Camps, R, Casillas, L. A., Costal, D., Gilbert, M., Martín, C. y Pérez, O. (2005) Bases de Datos (1ª Ed.) Catalunya, España. España: Eureka Media.
5. Cervera, S. N., & Lopez, R. E. (2017). Análisis y determinantes de la morosidad de los clientes de la caja municipal sullana. Tesis de Pregrado. Universidad Católico Santo Toribio de Mogrovejo, Chiclayo, Perú.
6. Cabrera, G. S. (2017). Metodologia para la construccion de un sistema de ayuda a la decision para Instituciones Financieras del sector de la economia popular y Solidaria. Tesis de Postgrado. Universidad de Cuenca, Cuenca.
7. Cobo, A. (s.f.). Base de datos relacionales: Teoría y práctica (1ª ed.). Madrid, España: Visión Libros.
8. Effy, O. (2001). Administración de Sistemas de Información (2da Ed.). México, DF. México: Tomson Learning.
9. Francois, D. (2004). Planificación Táctica y de Mediano Plazo (1era Ed.). Toronto, Canadá: CECADI.
10. Gómez, A. y De Abajo, N. (1998). Los sistemas de información en la empresa. Madrid, España: Servicio de publicaciones de la Universidad de Oviedo.
11. Gomez, J. T. (2016). Propuesta de estrategias de cobranza para disminuir la morosidad en la empresa electro oriente s.a. Tesis de Pregrado. Universidad Cesar

- Vallejo, Chiclayo, Perú.
12. Hernández Sampieri, R., Fernández, C. y Baptista, P. (2014) Metodología de la investigación (6ta Ed.). México, DF., México: McGraw Hill Interamericana.
  13. Inmon, B. (2002), Building the Data Warehouse (3ra Ed.). Toronto. Canadá: Wiley.
  14. Jiménez, C. J. (2017). Data mart para el soporte de la toma de decisiones en el área de planificación de tiendas. Tesis de Pregrado. Universidad Cesar Vallejo, Lima, Perú.
  15. Kimball, R. y Ross, M. (2013). The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling (2da edición). Toronto. Canadá: Wiley.
  16. Laudon, K. y Laudon, J. (2008). Sistemas de Información Gerencial (10ma Ed.). México, DF., México: Pearson Educación.
  17. Luque, I., Gómez, M., López, N. y Cerruela, G. (2002). Base de Datos. México D.F., México: Alfaomega Grupo Editor.
  18. Moss, L. y Atre, S. (2003) Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications. Boston. Inglaterra: Addison Wesley.
  19. Morales, M., & Vargas, N. A. (2017). Identificar los factores externos y su influencia en los índices de morosidad en una empresa comercializadora de productos de electricidad. Tesis de Pregrado. Pontificia Universidad Católica del Perú, Lima.
  20. Nevado, V. (s.f.). Introducción a las bases de datos relacionales. Madrid, España: Vision Libros.
  21. Osorio, F. (2008). Base de datos relacionales: Teoría y práctica (1ª ed.). Madrid, España: Thomson.
  22. Pons, O., Marín, N., Medina, J., Acid, S. y Vila, A (2005). Introducción a las bases de datos: El modelo relacional. Madrid, España: Thomson.
  23. Reina, O. I. (2012). Diseño de un modelo de gestión de cobranza para disminuir el índice de morosidad. Tesis de Pregrado. Universidad Señor de Sipán, Pimentel, Perú.
  24. Rojas, E. L. (2015). La gestión de la morosidad y su relación con el apalancamiento financiero de la empresa grupo unión wanca s.s.c. del distrito de puente piedra. Tesis de Pregrado. Universidad Cesar Vallejo, Lima, Perú.
  25. Ruiz, J. (2018). La morosidad en el pago de arbitrios en la municipalidad distrital

- de lince. Tesis de Pregrado. Universidad Nacional Mayor de San Marcos, Lima, Perú.
26. Silberschatz, A., Korth H. y Sudarshan S. (2005). Fundamentos de bases de datos (4ta Ed.) España, Madrid: McGraw Hill Interamericana.
  27. Tamayo y Tamayo, M. (1997) El Proceso de la Investigación científica. México DF., México: Editorial Limusa S.A.
  28. Tamayo, M. y Moreno, F. (2006, diciembre) Análisis del modelo de almacenamiento MOLAP frente al modelo de almacenamiento ROLAP. Ingeniería e Investigación. Recuperado el 16 de febrero del 2018, de <http://www.redalyc.org/html/643/64326317/>.
  29. Thierauf, R. (1994). Sistema de Información Gerencial para Control y Planificación. México, D.F., México: Limusa S.A.
  30. Torres, L. (2007). Business Intelligence. Recuperado de <http://www.gravitar.biz/index.php/bi/bi-terminologia-1>.
  31. Ullman, J.D. y Widom, J. (1999). Introducción a los Sistemas de Bases de Datos. México D.F., México: Editorial PHH.
  32. Valdiviezo, M. J., Herrera, I. Z. S. y Jáuregui, G. D. (2007). Análisis y Diseño de una Herramienta de Desarrollo de Soluciones para Inteligencia de Negocios – Análisis dimensional .Tesis de pregrado. Pontificia Universidad Católica del Perú. Lima. Perú.
  33. Vitt, E. y Luckevich, M. (2002). Business Intelligence Técnicas de análisis para la toma de decisiones estratégicas (1era Ed.). Madrid. España: McGraw-Hill/Interamericana.
  34. Yalán, J. y Palomino, L. (2013, enero - junio), Implementación de un Data mart como una solución de Inteligencia de Negocios para el área de logística de T-Impulso, Revista de investigación de sistemas e informática, Universidad Nacional Mayor de San Marcos, Lima, Perú.
  35. Zambrano, J. (2011). Análisis, diseño e implementación de un data mart para el área de mantenimiento y logística de una empresa de transporte público de pasajeros. Tesis de pregrado. Pontificia Universidad Católica del Perú, Lima, Perú.

## ANEXO A: OPERACIONALIZACIÓN DE VARIABLES

VARIABLE	VARIABLE DESCRIPTIVA	INDICADOR		ITEM
INFORMACION TACTICA	COBRANZA	1	Pagos fuera del plazo promedio de cobranza	¿Cuál es plazo promedio de cobranza de los clientes?
				¿Qué cantidad de clientes cumplen con el plazo promedio de cobranza?
		2	Porcentaje de Eficiencia de cobranza	¿Cuál es la cantidad neta facturada de los clientes?
				¿Cuál es la eficiencia de gestión en base a la factura cobrada?
		3	Porcentaje de cobranza gestionada	¿Cuál es la cantidad de la deuda gestionada?
		DEUDA	4	Porcentaje de clientes gestionados por deuda
	¿Qué cantidad de clientes realizaron el pago durante el año?			
	5		Porcentaje de pagos efectuados	¿Qué cantidad de pagos efectuados se realizó por distrito durante el año?
				¿Qué cantidad de la deuda vencida fue cubierto durante el año?
	CLIENTE	7	% de Reclamos totales	¿Cuál es la cantidad de clientes que realizaron su pago?
				¿Qué cantidad de clientes logran presentar sus reclamos?
		8	Porcentaje de Conexiones activas	¿Qué cantidad de reclamos son atendidos?
				¿Cuántos clientes presentan conexiones activas?
				¿Cuál es la cantidad de conexiones totales?

## ANEXO B: FICHA PARA ANÁLISIS DOCUMENTAL DEL CUBO MULTIDIMENSIONAL

### **FICHA NRO 1:**

¿Cuál es la deuda total durante el año por distrito?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Cliente	☐ Cliente Direccion Distrito	Igual	{ AYACUCHO }	<input type="checkbox"/>
<Seleccionar dimensión>				

Tiempo Semestre	Tiempo Semestre Descripción	Morosidad Total Deuda
01	PRIMER SEMESTRE	249857,78
02	SEGUNDO SEMESTRE	215567,95

### **FICHA NRO 2:**

¿Qué cantidad total se cobró durante el año por actividad económica en la región de ayacucho ?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Cliente	☐ Cliente Actividad	Igual	{ Agencia de viajes }	<input type="checkbox"/>
<Seleccionar dimensión>				

Tiempo Mes	Morosidad Total Cobranza Efectuada
01	82,8
02	116,6
03	85,3
04	150,3
05	151
06	275,9
07	105,9
08	166,5
09	189,4
10	387,8
11	177,9
12	291

### **FICHA NRO 3:**

¿Cuál es la cantidad de facturas gestionadas por cliente durante el año ?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Cliente	Cliente Nombre Completo	Igual	{ ABANTO LUJAN MARIA YSABEL }	<input type="checkbox"/>
<Seleccionar dimensión>				

Tiempo Mes	Cliente Key	Morocidad Nro Facturas Gestionadas
01	2365320	1
02	2365320	1
03	2365320	1
04	2365320	1
05	2365320	0
06	2365320	1
07	2365320	0
08	2365320	1
09	2365320	1
10	2365320	1
11	2365320	1
12	2365320	1

### **FICHA NRO 4:**

¿Cuál es la cantidad de clientes gestionados por deuda por distrito durante el año ?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Pago	PagoCodigo	Igual		<input type="checkbox"/>
<Seleccionar dimensión>				

Cliente Direccion Distrito	Morocidad Nro Clientes
ANDRES AVELINO CAC...	80928
AYACUCHO	251034
CARMEN ALTO	64907
JESUS NAZARENO	43956
SAN JUAN BAUTISTA	142043

## **FICHA NRO 5:**

¿Cuál es la cantidad de facturas no canceladas por distrito durante el segundo trimestre?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Tiempo	⌵ Tiempo Trimestre Descripción	Igual	{ SEGUNDO TRIMESTRE }	<input type="checkbox"/>
<Seleccionar dimensión>				

Cliente Direccion Distrito	Morocidad Nro Facturas No Cancela...
ANDRES AVELINO CAC...	2982
AYACUCHO	8560
CARMEN ALTO	2417
JESUS NAZARENO	1386
SAN JUAN BAUTISTA	4958

## **FICHA NRO 6:**

¿Cuál es la facturación total por distrito durante el año de los clientes?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Cliente	⌵ Cliente Direccion Distrito	Igual	{ CARMEN ALTO }	<input type="checkbox"/>
<Seleccionar dimensión>				

Factura Dia Pago	Morocidad Total Facturado
01/02/2018	226,3
01/03/2018	3333,6
01/05/2018	235,7
01/06/2018	5694,1
01/08/2018	13121,7
01/09/2018	1806,9
01/10/2018	3775,5
01/11/2018	197,9
01/12/2018	7916
02/01/2019	3064,4
02/02/2018	4259,91
02/03/2018	2224,4
02/04/2018	4095,4
02/05/2018	4646,3
02/07/2018	3129,9
02/08/2018	779,1
02/10/2018	2208,5
02/11/2018	49,6
03/01/2019	8661,1

## **FICHA NRO 7:**

¿Cuál es la cantidad de conexión total activa por mes de cada distrito durante el año?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Tiempo	Tiempo Mes Descripción	Igual	{ ENERO }	<input type="checkbox"/>
<Seleccionar dimensión>				

Factura Estado Coneccion	Cliente Direccion Distrito	Morocidad Nro Conexiones Totales
ACTIVO	ANDRES AVELINO CAC...	6609
ACTIVO	AYACUCHO	20723
ACTIVO	CARMEN ALTO	5273
ACTIVO	JESUS NAZARENO	3622
ACTIVO	SAN JUAN BAUTISTA	11711

## **FICHA NRO 8:**

¿Cuál es el monto total de la deuda por código de pago en cada distrito?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Cliente	Cliente Direccion Distrito	Igual	{ AYACUCHO }	<input type="checkbox"/>
<Seleccionar dimensión>				

Reclamo Procesado	Reclamo Fecha	Morocidad Nro Reclamos Totales
NO	2018-01-04	36
NO	2018-01-05	8
NO	2018-01-08	206
NO	2018-01-09	200
NO	2018-01-10	12
NO	2018-01-11	28
NO	2018-01-12	199
NO	2018-01-15	168
NO	2018-01-16	268
NO	2018-01-17	0
NO	2018-01-18	0
NO	2018-01-19	12
NO	2018-01-22	24
NO	2018-01-23	48
NO	2018-01-24	52
NO	2018-01-25	132
NO	2018-01-29	0
NO	2018-01-30	0
NO	2018-02-05	24
NO	2018-02-06	88
NO	2018-02-07	17

## **FICHA NRO 9:**

¿Cuál es el monto total del pago fuera de fecha por distrito durante el año?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
<Seleccionar dimensión>				

Tiempo Semestre	Cliente Direccion Distrito	Tiempo Semestre Descripción	Morocidad Total Pago F Fecha
01	ANDRES AVELINO CAC...	PRIMER SEMESTRE	77543,3499999999
01	AYACUCHO	PRIMER SEMESTRE	249857,78
01	CARMEN ALTO	PRIMER SEMESTRE	42470,18
01	JESUS NAZARENO	PRIMER SEMESTRE	45288,29
01	SAN JUAN BAUTISTA	PRIMER SEMESTRE	109241,9
02	ANDRES AVELINO CAC...	SEGUNDO SEMESTRE	49339,16
02	AYACUCHO	SEGUNDO SEMESTRE	215567,95
02	CARMEN ALTO	SEGUNDO SEMESTRE	38629,22
02	JESUS NAZARENO	SEGUNDO SEMESTRE	32348,41
02	SAN JUAN BAUTISTA	SEGUNDO SEMESTRE	89646,8699999999

## **FICHA NRO 10:**

¿Cuál es la cantidad de reclamos totales procesados durante el año?

Dimensión	Jerarquía	Operador	Expresión de filtro	Parámetros
DIM Cliente	☰ Cliente Direccion Distrito	Igual	{ SAN JUAN BAUTISTA }	<input type="checkbox"/>
<Seleccionar dimensión>				

Factura Estado Coneccion	Tiempo Cliente	Morocidad Total Pago
ACTIVO	11090590	94,3
ACTIVO	12103170	4432,3
ACTIVO	14060815	129
ACTIVO	14060817	953,1
ACTIVO	14060832	5088,5
ACTIVO	14060845	340,9
ACTIVO	14060853	2329,4
ACTIVO	14060856	83
ACTIVO	14060926	182,5
ACTIVO	14060929	5228,2
ACTIVO	14060931	125,1
ACTIVO	14060932	2863,7
ACTIVO	14060933	303,5
ACTIVO	14060934	580,1
ACTIVO	14060936	210,5
ACTIVO	14060937	334,5
ACTIVO	14060938	318