



**UNIVERSIDAD CÉSAR VALLEJO**

## **FACULTAD DE INGENIERÍA**

**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**DATA MART PARA EL PROCESO CREDITICIO EN EL ÁREA DE  
CAMPAÑAS COMERCIALES DE LA EMPRESA MIBANCO**

**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:  
INGENIERO DE SISTEMAS**

**AUTOR:**

**JEAN BRYANT CALLIRGOS CARDENAS**

**ASESOR:**

**DR. ADILIO CHRISTIAN ORDOÑEZ PÉREZ**

**LÍNEA DE INVESTIGACIÓN:**

**SISTEMAS DE INFORMACIÓN Y COMUNICACIONES**

**LIMA - PERÚ**

**2018**

# Página del jurado

 <b>UCV</b> UNIVERSIDAD CÉSAR VALLEJO	<b>ACTA DE APROBACIÓN DE LA TESIS</b>	Código : F07-PP-PR-02.02 Versión : 07 Fecha : 31-03-2017 Página : 1 de 106
--	---------------------------------------	---

El Jurado encargado de evaluar la tesis presentada por don (a):

**CALLRGOS CARDENAS JEAN BRYANT**

cuyo título es:

**DATAMART PARA EL PROCESO CREDITICIO EN EL ÁREA DE CAMPAÑAS  
COMERCIALES DE LA EMPRESA MIBANCO**

Reunido en la fecha, escuchó la sustentación y la resolución de preguntas por el estudiante, otorgándole el calificativo de: **15** (números) **QUINCE**(letras).

Lima, Miércoles 12 de Diciembre del 2018

  
.....  
PRESIDENTE  
Mgtr. CRUZADO PUENTE DE LA VEGA  
CARLOS FRANCISCO

  
.....  
SECRETARIO  
Mgtr. CHUMPE AGOSTO JUAN BRUES LEE

  
.....  
VOCAL  
Dr. ORDOÑEZ PEREZ ADILIO CHRISTIAN

Elaboró	Dirección de Investigación	Revisó	Representante de la Dirección / Vicerrectorado de Investigación y Calidad	Aprobó	Rectorado
---------	----------------------------	--------	---	--------	-----------

### **Dedicatoria**

Dedico esta tesis a mi familia por el gran apoyo, tiempo, cariño y paciencia brindado en mi etapa universitaria y sobre todo a mi madre Rosa, mi padre Jose Luis y mi hermano Bairon que son un ejemplo de esfuerzo, coraje, valentía que me inspiran día a día a seguir y a no rendirme ninguna ante cualquier adversidad.

### **Agradecimiento**

Agradezco a mi familia y amigos que me apoyaron que me alentaron para el desarrollo de esta tesis.

Al Dr. Adilio Christian Ordoñez Pérez, por apoyarme su asesoría permanente en la elaboración de esta tesis.

A la empresa Mibanco por permitirme desarrollarme profesionalmente realizando mis prácticas profesiones y permitiéndome implementar mi tesis como solución de negocio.

# Declaratoria de autenticidad

## DECLARACIÓN DE AUTENTICIDAD

Yo Jean Bryant Callirgos Cardenas con DNI N° 72481689 a efecto de cumplir con las disposiciones vigentes consideradas en el Reglamento de Grados y Títulos de la Universidad Cesar Vallejo, Escuela de ingeniería de Sistemas declaro bajo juramento que toda la documentación que acompaño es veraz y autentica.

Así mismo, declaro también bajo juramento que todos los datos e información que se presenta en la presente tesis son auténticos y veraces.

En tal sentido asumo la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión tanto de los documentos como de información aportada por la cual me someto a lo dispuesto en las normas académicas de la Universidad Cesar Vallejo.

Lima, 03 diciembre 2019



**Jean Bryant Callirgos Cardenas**

**DNI: 72481689**

## Presentación

Señores miembros del Jurado:

Dado el cumplimiento establecido a las normas en el Reglamento de Grados y Títulos sección de Pregrado de la Universidad César Vallejo para la experiencia curricular de Metodología de la Investigación Científica, presento el trabajo de investigación pre-experimental denominado: titulado “Data mart para la evaluación de los créditos en el área de campañas comerciales en la empresa Mibanco” en el año 2018.

La investigación, tiene como propósito fundamental: determinar cómo influye un Data mart en la evaluación de los créditos en el año 2018.

Esta investigación se divide en 7 capítulos:

Inicia con el primer capítulo presentando la problemática: se formula la problemática, se plantean los objetivos, se plantea la hipótesis, se plantea la justificación, se plantea los trabajos previos y la fundamentación científica. Así mismo en el capítulo dos, se muestra el marco metodológico sobre la investigación en la que se aplica el trabajo de campo de la variable de estudio, la selección del diseño, la selección de la población y la muestra, las técnicas e instrumentos de recolección de datos y los métodos de análisis. En el capítulo tres se interpretan los resultados. En el capítulo cuatro se realiza la discusión del trabajo. Las conclusiones se muestran en el quinto capítulo, las recomendaciones se muestran en el sexto capítulo y finalmente las referencias bibliográficas se muestran en el capítulo siete.

Señores miembros del jurado espero que esta investigación sea evaluada y merezca su aprobación.

## ÍNDICE

	página
Página del jurado	ii
Dedicatoria	iii
Agradecimiento	iv
Declaratoria de Autenticidad	v
Presentación	vi
Índice	vii
Índice de tablas	ix
Índice de figuras	x
Resumen	xi
Abstract	xii
I. INTRODUCCIÓN	13
1.1. Realidad Problemática	14
1.2. Trabajos previos	17
1.3. Limitaciones de la investigación	25
1.4. Teorías relacionadas al tema	26
1.5. Formulación del problema	39
1.6. Justificación del estudio	40
1.7. Hipótesis	42
1.8. Objetivos	42
II. MÉTODOS	44
2.1. Diseño de investigación	45
2.2. Variables Operacionalización	48
2.3. Población y Muestra	52
2.4. Técnicas e instrumentos de recolección de datos	54
2.5. Métodos de análisis de datos	59
2.6. Aspectos éticos	64
III. Resultados	66
3.1. Análisis Descriptivo	67
3.2. Análisis inferencial	69
3.3. Prueba de hipótesis	74
IV. DISCUSIÓN	80
V. CONCLUSIONES	82

VI.	RECOMENDACIONES	84
VII.	Referencias Bibliográficas	86
VIII.	ANEXOS	95
	Anexo 1: Matriz de Consistencia	96
	Anexo 2: Ficha técnica, instrumento de recolección de datos	97
	Anexo 3: Instrumento de investigación	98
	Anexo 4: Base de datos experimental	102
	Anexo N°6: Evaluacion del instrumento	109
	Anexo 7: Entrevista	118
	Anexo 8: Carta de aprobación de la empresa	121
	Anexo 9 DESARROLLO DE LA METODOLOGÍA	122



## ÍNDICE DE TABLAS

	Página
Tabla 1: Criterios de selección para las metodologías de un Data mart	34
Tabla 2: Expertos para la calificación de desarrollo de las Metodologías de un Data mart	35
Tabla 3: Operacionalización de las variables	50
Tabla 4: Indicadores de la evaluación de créditos	51
Tabla 5: Población	52
Tabla 6: Recolección de datos	55
Tabla 7: Validez para el índice de morosidad	56
Tabla 8: Validez para el índice de cobertura	56
Tabla 9: Confiabilidad para Índice de morosidad	58
Tabla 10: Confiabilidad para el índice de cobertura	59
Tabla 11: Medidas descriptivas del índice de morosidad en la evaluación de créditos después de la implementación del data mart	67
Tabla 12: Medidas descriptivas del índice de cobertura de la evaluación de créditos antes y después de implementar el Data mart	68
Tabla 13: Pruebas de Normalidad del índice de morosidad en la evaluación de créditos antes y después de implementar el Data mart	70
Tabla 14: Pruebas de Normalidad del índice de cobertura en la evaluación de créditos antes y después de implementar el Data mart	72
Tabla 15: Prueba de T-Student para el índice de morosidad en la evaluación de créditos antes y después de implementado el Data mart	75
Tabla 16: Prueba de T-Student para el índice de cobertura en la evaluación de créditos antes y después de implementado el Data mart	78

## ÍNDICE DE FIGURAS

	Página
Figura 1: Incremento de morosidad en el sistema financiero	14
Figura 2: Índice de morosidad en la evaluación de los créditos	16
Figura 3: Índice de cobertura en la evaluación de los créditos	17
Figura 4: Arquitectura de un Data mart	30
Figura 5: Extracción, transformación, carga	31
Figura 6: Fases de la metodología de Kimball	37
Figura 7: Cálculo de la confiabilidad o fiabilidad	58
Figura 8 Índice de morosidad en la evaluación de créditos antes y despues de la implementaión del Data mart	68
Figura 9 Índice de cobertua antes y despues de implementar el Data mart	69
Figura 10 Prueba de normalidad del índice de morosidad antes de implementar el Data mart	71
Figura 11 Prueba de normalidad del índice de morosidad después de implementar el Data mart	71
Figura 12 Prueba de normalidad del índice de cobertura antes de implementar el Data mart	73
Figura 13 Prueba de normalidad del índice de cobertura después de implementar el Data mart	73
Figura 14 índice de morosidad - Comparativa general	75
Figura 15 Prueba T-Student- Índice de morosidad	76
Figura 16 índice de cobertura - Comparativa General	78
Figura 17 Prueba T-Student - Índice de cobertura	79

## Resumen

La presente investigación muestra el detalle de desarrollo de un Data mart para la evaluación de los créditos en el área de campañas comerciales de la empresa Mibanco, debido a que el escenario de la empresa antes de implementación del data mart presentaba problemas en el indicador de índice de morosidad e índice de cobertura. El objetivo de esta tesis fue determinar cómo es que influencia de un Data mart para la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco en el año 2018.

Comienza describiendo previamente los aspectos teóricos de la evaluación de créditos, así mismo la metodología que se aplicó para la implementación del Data mart. Para la implementación del Data mart, se escogió mediante el juicio de expertos la metodología Ralph Kimball, por ser la que más se adecuaba a las etapas del proyecto y a las necesidades, ya que es una de las metodologías más aplicadas en bancos como se mostró en los trabajos previos.

Aplicada fue el tipo de investigación, preexperimental es el diseño de la investigación y el enfoque es cuantitativo. La población para el índice de morosidad y cobertura fue de 92 por desembolsos. El tamaño de la muestra se conformó por 75 lo cual fue estratificado por días resultando 25. Aleatorio probabilístico simple fue el tipo de muestreo. El fichaje es la técnica de recolección de datos y la ficha de registro es el instrumento, el cual fueron validados por expertos. La implementación del Data mart permitió disminuir el índice de morosidad del 78% al 14% del mismo modo, el índice de cobertura incremento del 12%, al 75%. Los resultados mencionados, nos llevó a la conclusión que el Data mart mejora la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

Palabras clave: Inteligencia de negocios, evaluación de créditos, Ralph Kimball

## **Abstract**

The present investigation shows the Data mart development detail to the evaluation of credits in the commercial campaigns field of MiBanco enterprise, because the scenario before of the data mart implementation had certain problems in the morosity index indicator and coverage index indicator. The main objective of this thesis was determine how is the influence of a data mart to credits evaluation in commercial campaigns field of MiBanco enterprise in 2018

For the beginning, the theoretic aspects of credits evaluation were described previously, also the methodology applied for the data mart implementation. To this, with experts judgement, the Ralph Kimball methodology was chosen, because it was the most adequate to the necessities of the current project, and is one of the most recommended methodologies to be applied in banks how was mentioned in previous documents.

Applied was the investigation type for this project, preexperimental is the design of the investigation and The focus was quantitative. The population to the morosity and coverage index was of 92 in disbursements. The sample size was about 75 which was stratified for days obtaining 25 finally. Simple random sample was the sample type. The signing is the technique of data collection and The register token is the instrument validated for experts. The data Mart implementation allows diminish the morosity index from 78% to 14%, in the same way allows an increment with the coverage index from 12% to 75%. The mentioned results, allows us to conclude that the Data Mart improve the credits evaluation in the commercial campaigns field of MiBanco enterprise.

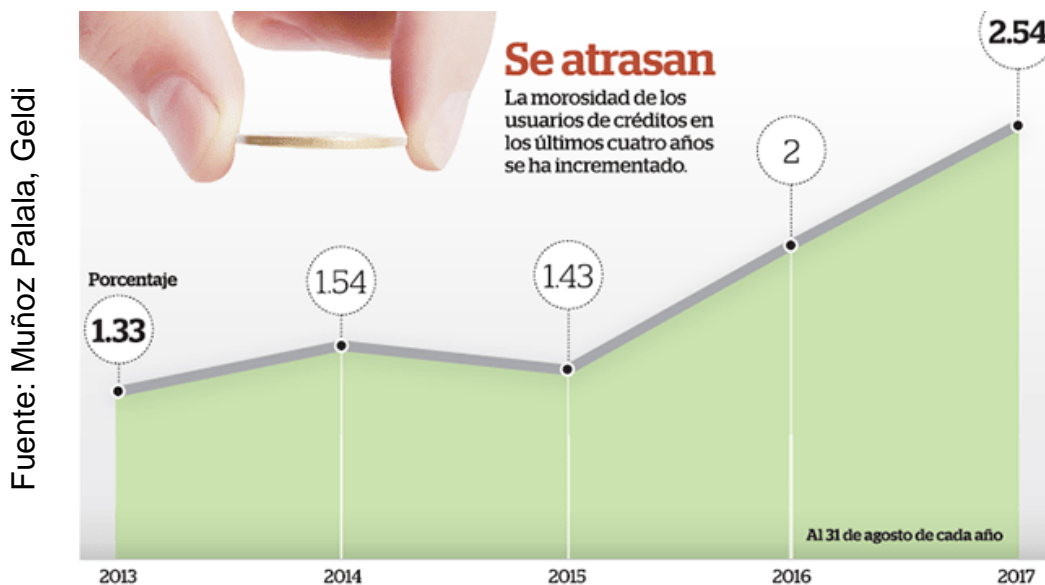
Keywords: Business intelligence, evaluation of the credits, Ralph Kimball

# **I. INTRODUCCIÓN**

## 1.1. Realidad Problemática

En el escenario internacional, según Muñoz Palala en su publicación en el diario el Periódico de Guatemala (2017) nos menciona que “una de las problemáticas que afronta el sistema financiero es el aumento de la morosidad de 2.54%, podría deberse a que hubo una merma en los ingresos de los deudores individuales, que quizá cuando accedieron al préstamo contaban con un trabajo, y para las organizaciones podría tratar de que a los negocios no les fue bien. Además, podría tratarse de que los bancos a la hora de realizar los análisis para ofrecer los créditos no evaluaron bien si los usuarios aplicaban. En general podría ser que de alguna manera el exceso de liquidez que hay en la banca podría haber relajado algunos requisitos para ofrecer los créditos y crear un programa de créditos más agresivos” (parr.7). Según Muñoz la mora viene incrementándose por años, como se muestra en la figura 1, el incremento desde el año 2013 al 2017 es de 1.21% con tendencia a seguir incrementándose.

**Figura 1: Incremento de morosidad en el sistema financiero**

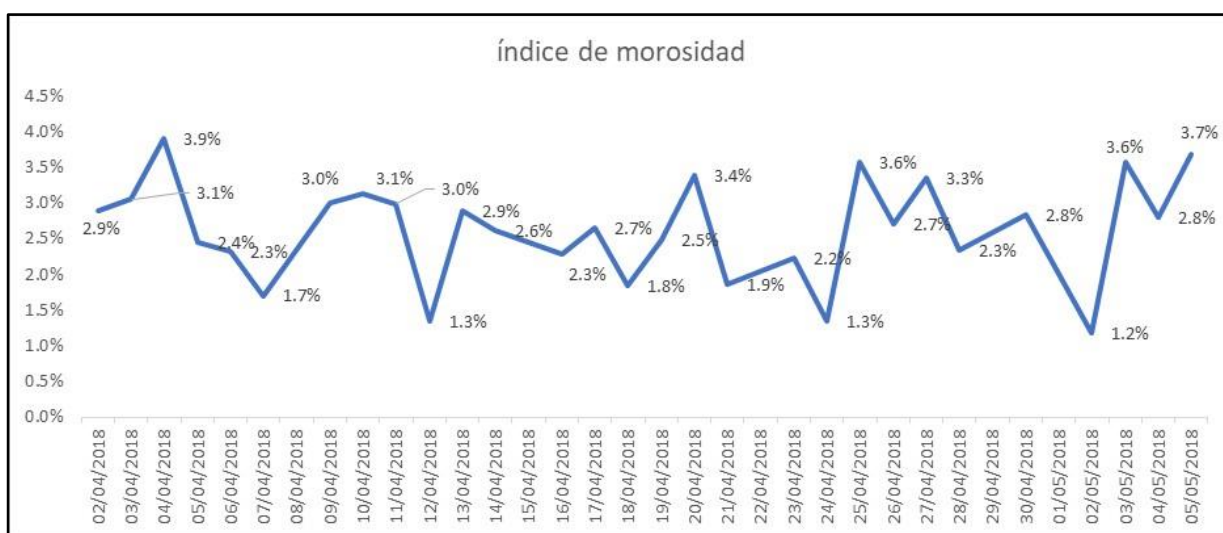


En el escenario nacional, según Castillo Arévalo en su publicación en el diario El Comercio nos menciona que “la mora en los créditos pymes creció más de 37%

por sobreendeudamiento, dicho deterioro del financiamiento pyme responde que se les ofreció una sobreoferta, relajando las políticas de financiamiento, y no evaluando si el cliente calificaba para esos préstamos, aumentaron el plazo de financiamiento y a la vez incrementando el prestamos de los clientes sin considerar su capacidad de endeudamiento” (2014, parr.6). Así mismo según Solis Espinoza en su publicación en el diario Gestión (2016) nos dice que “Existe un problema de sobreendeudamiento debido a que existen muchos clientes compartidos por las entidades acreedoras y ocasionan incremento en la morosidad” (parr.2). Por último, Mía Rios en la publicación realizada en el diario Gestión nos menciona que “los clientes compartidos con diferentes entidades representan el 78.3% de la distribución de deuda, generando una preocupación de un posible sobre endeudamiento y con más probabilidad de endeudarse. La mora en los clientes exclusivos es menor respecto del segmento de clientes compartidos” (2017, par.9). Esta investigación se realiza en Mibanco, es una entidad que se desempeña en el sector financiero, que tiene como misión transformar la vida de sus clientes y trabajadores con la inclusión financiera, promoviendo el crecimiento del Perú. La evaluación de los créditos empieza con la actualización de las fuentes de información de campañas históricas y del sistema financiero (RCC), luego se comienza a analizar la información consolidada de cosechas, efectividades, colocaciones y aportes de campañas. Luego se procede con la propuesta de perfil del cliente considerando los lineamientos de los documentos normativos, luego se comenzará a analizar el nuevo perfil propuesto para la nueva campaña, luego proponer el nuevo perfil del cliente en campaña al Gerente de área de inteligencia y desarrollo comercial, y finalmente generar la base de campañas para que sean gestionadas a nivel nacional en todas las agencias de Mibanco. El problema que está ocurriendo en Mibanco es el incremento de la cartera vencida (clientes morosos), uno de los motivos de este incremento es al momento de generar el perfilamiento de dichos clientes, la información obtenida de las fuentes de datos no es 100% confiable ya que algunos registros cuentan con información errónea, y se tienen que solucionar de manera manual, otro problema que presenta el área es que los análisis de información histórica se realiza en un Excel de forma muy manual y con una estructura no dimensional lo que genera mucho tiempo en realizar los análisis, consumiendo muchas horas hombre. Al no contar con una

herramienta para generar estrategias y una oportuna información para toma de decisiones que permitan generar campañas de mayor impacto y más efectivas ya que actualmente se están generando pérdidas por los clientes morosos. Tampoco se cuenta con un seguimiento a los encartes que se envían a los clientes ofreciéndoles un crédito, ya que se realiza una inversión para hacer posible dicho envío y no se cuenta con un seguimiento. Uno de los problemas que tiene la empresa es la complejidad de poder analizar el índice de morosidad, ya que actualmente se encuentra en 3.7% con tendencia a seguir incrementándose como se evidencia en la figura 2, debido a que se están otorgando créditos a personas que no cumplían ciertos requisitos como también se están desembolsando créditos muchos mayores al monto ofertado que fue calculado por parte del área de campañas además no se cuenta con un repositorio de datos histórica consolidada, ya que toda la data se encuentra guardada de manera temporal en diferentes fuentes de datos y no aprovecharse para su explotación. Los resultados esperados según la presenta investigación es que no haya clientes que caigan en mora o que sea el mínimo esperado.

**Figura 2: Índice de morosidad en la evaluación de los créditos**



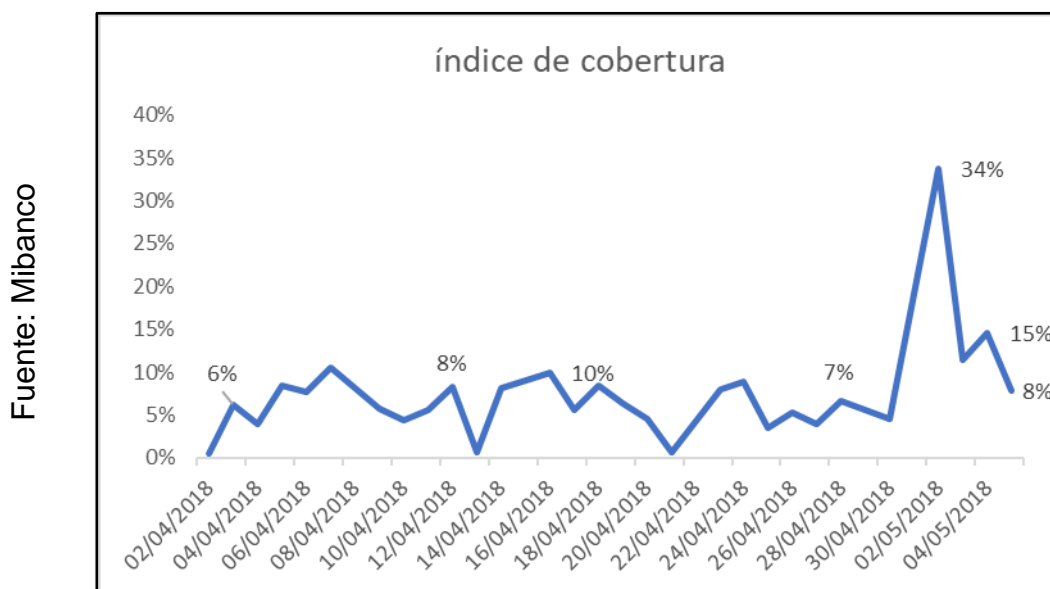
Fuente: Mibanco

Otro de los problemas cruciales que presenta en Mibanco es que no se lleva un control sobre las provisiones ya consumidas por el deterioro de los créditos, ya que



actualmente cuenta con un 8% de consumo de las provisiones de los créditos. Como se evidencia en la figura 3, la cartera vencida está afectando a las provisiones el banco, con la presente investigación tiene como uno de los objetivos es llegar a reducir los clientes morosos y hacer que en índice de cobertura disminuya.

**Figura 3: Índice de cobertura en la evaluación de los créditos**



A persistencia de los problemas planteado surge la interrogante: ¿Qué sucederá si en Mibanco continua con los mismos problemas?, en respuesta a dicha pregunta, su mala evaluación a clientes generará pérdidas económicas. Así mismo, el porcentaje de índice de morosidad se mantendrá o disminuirá. Mostrando que no habrá mejora en la reducción de morosidad.

## 1.2. Trabajos previos

Se reviso distintas fuentes primarias para el desarrollo de esta tesis, sirviendo a esta investigación base que sustente el problema mencionado.

En el año 2017, Hidalgo Ortiz en la tesis de maestría llamada "El análisis inteligente de datos y su incidencia en el otorgamiento de créditos financieros en la

cooperativa de ahorro y crédito Cushun Ltda” en la universidad técnica de Ambato, Ecuador. Trato la problemática de la inadecuada colocación de créditos, que se debió a la carencia de conocimiento de la administración de Riesgos, e información inoportuna generada por las herramientas informáticas que brindaron facilidades, para una inoportuna toma de decisiones, y por la necesidad de querer cumplir las metas de colocaciones, promovieron la ineficiencia de las operaciones. El objetivo de esta tesis de maestría es garantizar la solvencia y estabilidad a la organización, con una colocación adecuada de créditos financieros a través del análisis inteligente de datos. El resultado obtenido en esta tesis mostro que se solucionó el problema de análisis erróneos de información histórica para tomar de decisiones en el proceso de otorgamiento de créditos, automatización e incorporación de herramientas que ayuden a predecir, detectar patrones, contar con reportes actualizados, eficientes que permitieron tomar decisiones gerenciales de manera oportuna y fiable. En conclusión, se demostró que la implementación de herramientas de business Intelligence mejoró considerablemente la toma de decisiones mediante análisis estratégicos permitiendo la reducción del índice de morosidad de 74% al 4%, así mismo para el indicador de índice de cobertura incremento de 12% a 92%, así mismo incrementó el número de colocaciones a un 30%.

De esta investigación, se tomó en cuenta para obtener mayor conocimiento sobre el indicador de índice de morosidad e índice de cobertura, pues se reflejó en este proyesto la disminución de la morosidad al aplicar business Intelligence.

En el 2015, Avellaneda Rojas en la tesis “Implementación de un Data mart como herramienta de mejora en la toma de decisiones del servicio de colocaciones bancarias en una entidad financiera del estado”, en la Universidad San Martin, Perú. La problemática planteada en esta tesis nos dice que el banco de la nación carece de un software que facilite tener datos veloz, oportuna y confiable de sus colocaciones actuales, además carece de un sistema que le facilite la data histórica para realizar proyecciones, así mismo no se sabe el perfilamiento de sus clientes

por los prestamos realizados, no se tiene establecida una segmentación de para un análisis histórico. La presente investigación su objetivo es establecer como se relaciona la implementación de un data mart y como la toma de decisiones llega a mejorar con dicha implementación. Los resultados obtenidos en esta investigación muestran que se logró disminuir el tiempo de creación de reportes y se aumentó de manera importante el nivel de servicio de los requerimientos de datos. Finalmente se demostró que implementando un data mart mejora notablemente el proceso de tomar una decisión y que facilita realizar una medición estructurada de los beneficios que se adquieren al implementar un data mart. Se demuestra que se puede aplicar de manera exitosa la metodología de Ralph Kimball en una institución financiera. Las conclusiones de esta investigación se demuestran que al aplicar Data mart mejora notablemente en la generación de los reportes, reduciendo el tiempo en las ejecuciones. En esta investigación se determinó que el aumento fue del 41.46% y así mismo se mejoró el tiempo al momento de realizar decisiones.

De esta tesis se tomó la metodología de desarrollo llamada Ralph Kimball que demuestra la factibilidad de utilizarla en una institución financiera, también las reducciones de tiempo al momento de generar los reportes y además el aumento considerable de colocaciones que tuvo la entidad financiera y demostrando que al implementar esta solución tecnológica conlleva beneficios muy considerables y también se tomara como aporte el uso de la herramienta Visual Studio 2014 para la creación de los cubos OLAP y la realización del ETL.

En el año 2014, Sánchez Trujillo en la tesis titulada “Desarrollo e implementación de un Data mart para el requerimiento de capital por riesgo de crédito del banco Interbank”, en la Universidad nacional de Trujillo, Perú. Acelerar la toma de decisiones fue el objetivo de esta tesis para el requerimiento del capital de riesgos mediante la creación de un data mart bajo el método que usa la SBS con el uso de herramientas Oracle. La problemática mostrada en esta tesis en el almacenado de información y cálculo de para la creación de reportes, dashboard y cálculo de indicadores, donde genera gran tiempo de generación porque las consultas no son

óptimas, así mismo no existe un modelo lógico donde se pueda guardar los datos calculados; generando sobrecostos en los recursos de la entidad, además existió una baja satisfacción de los usuarios en la generación del reporte, ya que las tareas que realizan los usuarios son muchas y tediosas en recursos y tiempo en la infraestructura utilizada. Los resultados de esta tesis demuestran que implementando un Data mart se optimizó el tiempo de integración y se redujo el tiempo de acceso a la información y se solucionó el exceso de recursos que tenía asignado para la generación de la base ya que había un alto costo de almacén de datos, ocupando y disminuyendo la performance de otros procesos en ejecución. La metodología que se usó en el proyecto fue Ralph Kimball.

De esta investigación se tomó como aporte, el uso del lenguaje PL/SQL junto con la metodología Ralph Kimball ya que el tiempo de consolidación de los datos a en sus 3 capas del modelo optimizado por el uso del lenguaje PL/SQL para los cálculos de gran impacto.

En el 2013 Temitope Adeoye Oketunji, en la tesis de maestría denominada “Diseño de Data Warehouse y Sistema de Business Intelligence un estudio de caso de una industria minorista, en la Universidad de Tecnología de Blekinge, Suecia. El problema presentado en esta investigación fue que no se contaba con un almacenamiento de datos estructurado, y además no se almacenaba de manera histórica la información, ya que toda la información se guardaba en diferentes tablas generando pérdida de datos o información inconsistente y al querer crear o actualizar los reportes se realizaba de manera manual consumiendo muchas horas hombre y además realizando toma de decisiones erróneas generando grandes pérdidas a la organización. Se formuló como objetivo desarrollar un sistema de almacenamiento de datos e inteligencia empresarial para apoyar a los usuarios que toman decisiones y estrategias de negocios a partir de datos históricos. Podemos afirmar categóricamente que el Business Intelligence en última instancia, aumenta los ingresos para el entretenimiento de cristal mejorando; la capacidad de aprovechar los datos del cliente para entender los patrones de compra tiene una

idea de éxito de nuevos productos, características y servicios, programa de visibilidad de los clientes más rentables, promociones y campañas de marketing de destino hacia los productos que fallan. Para la implementación del DW se aplicó la metodología del método espiral. En esta tesis se mostró cómo los datos se pueden integrar de diferentes fuentes a un almacenamiento de datos llamada repositorio único que se utiliza para la entrega a los usuarios finales y ejecutivos con el fin de su análisis y explotación de los datos. Se tuvo éxito en el diseño de cómo los datos pueden ponerse a disposición de los negocios para el día a día.

De esta investigación se contribuyó al marco teórico afirmándonos con lo siguiente: Con la implementación de una herramienta de BI la manera en se genera la carga y a la vez la transformación de la información es más veloz y además de unificar y consolidar toda la información, lo que favorece a la generación de reportes de manera más eficiente y además automatizada para una toma de decisiones eficaz.

En el año 2016 Pintar Matej en la tesis denominada “Data Warehouse en la sociedad para el comercio de electricidad”, en la Universidad de Ljubljana. La problemática presentada en la investigación fue de que la empresa no contaba con una estrategia para consolidar grandes cantidades de información por lo que al querer explotar la información se realizaba de manera muy manual y al no tener un modelo dimensional se realizaba cada consulta de manera muy lenta generando malestar en el personal encargado de realizarlo, además se realizó una proyección a futuro, ya que la empresa venía expandiéndose por Europa y al no tener un plan estratégico para almacenar grandes cantidades de datos, iba a generar una gran pérdida económica a la empresa. Se formuló como objetivo consolidar y unificar los datos de todas las fuentes y los campos clave para proporcionar una única fuente para la presentación de análisis e informes de inteligencia empresarial, y establecer una visión única de ventas y comercio. Los resultados de esta tesis demuestran que al implementar un Data Warehouse se consolidó toda la información principal mejorando la calidad de datos ya teniendo los datos corregidos y además cuando la empresa de electricidad siga expandiéndose por

Europa no tendrá problemas al momento de analizar su data histórica además de agilizar los reportes. La metodología que se usó en el proyecto fue la del Método Espiral. En conclusión, de la investigación se llegó a corregir la complejidad de la gestión de almacenamiento de los datos y consolidándolo y optimizando toda la información de la empresa de electricidad y realizando nuevos análisis sacando ventaja competitiva de otras empresas.

De esta investigación se tomó como aporte la consolidación de toda la data histórica y además de corregir los problemas con las datas almacenadas en sus distintas tablas, ya que mejorando la calidad de los datos se podrá realizar nuevos análisis y nuevas conclusiones para así generar una ventaja competitiva en el sector financiero.

En el 2013, Mayo Claudio en la tesis denominada "Diseño de un Data mart para la gestión de cartera de clientes de la banca mayorista de una entidad bancaria, en la Universidad Nacional de Ingeniería, Perú. La problemática presentada en esta tesis nos dice que en el BCP contaba con muchas fuentes de información, con las que se realizaban procesos para obtener reportes consolidados. Los reportes eran entregados luego de 2 meses, no permitiendo a los funcionarios de negocios hacer una correcta gestión de su cartera para aumentar la participación de mercado y rentabilidad del banco. No permitiendo sus seguimientos de los indicadores de gestión, desempeño ni llegar al nivel de operaciones ni desembolsos de productos dentro de estos reportes. Así mismo en el proceso para obtener la información por parte del funcionario para su gestión no es óptima. El no contar con agilidad en los procesos representa un problema ante el dinamismo del mercado y los clientes actualmente en crecimiento, en particular para la Banca Mayorista que necesita un trato especial por la cartera que maneja con el Banco. Esta investigación tiene como objetivo generar reportes en un tiempo que le permita al funcionario de negocio gestionar su cartera con la información más actual y también la generación de mayor rentabilidad así mismo tener una mayor participación en el mercado generando ventajas competitivas. En conclusión, la creación del Data mart brinda

una solución estratégica alineada a los objetivos del área de planeamiento estratégico y desarrollo de negocios, que permita a los analistas de negocios tomar decisiones en forma ágil en base a informado confiable. Así mismo al contar con reportes automáticos, se tuvo más tiempo para enfocarse en el análisis de la información los analistas de negocios. Se tuvo la información del banco consolidada, convirtiéndose en un gran repositorio para ser explotado por diversas áreas de acuerdo con las relaciones y necesidades identificadas entre estas.

De esta tesis se tomó como aporte la automatización de los reportes, quitando las horas hombres en generarlas y enfocarse solo en el análisis de la información para obtener una visión de 360 grados del cliente para así poder generar varios tipos de campañas y generar mayores ingresos al banco con clientes de buen perfilamiento. Así mismo obtener toda la información histórica consolidada y explotarla de manera confiable.

En el año 2015, Guizado Verdezoto en la tesis denominada “Diseño e implementación de un Data mart OLAP para el análisis gerencial académico, que será implementado en la unidad educativa La Colina”, en la Universidad Central del Ecuador, Ecuador. La problemática presentada en la investigación nos dice que la unidad educativa la Colina carecía información histórica de personal docente , estudiantes y administrativo para realizar toma de decisiones lo cual hace difícil contar con una educación de calidad, otro problema presentado fue que al momento de la creación de reportes, ya que bien se utilizaba herramientas ofimáticas se ha podido aplacar pero no guarda una historia de lo que se ha generado y además que permita consolidar la información y así poder realizar una decisión correcta y sobre salir de las demás instituciones educativas. Esta tesis tiene como objetivo diseñar e implementar un data mart para poder realizar análisis gerencial académico donde ayude a tomar decisiones mediante indicadores que faciliten el ministerio de educación, así mismo de la verificación de la información actualizada de la institución, además de la elaboración de reportes mediante el Reporting Services. En conclusión, la utilización de una herramienta de BI, proporciona a las organizaciones a tener un mayor control de sus áreas que integran en el negocio; mostrando como resultado la facilidad en tomar decisiones,

los cubos OLAP sirven de gran ayuda para explotar los datos históricos. Dado que se automatizó a un 100% los reportes que se generaban antes de manera manual, facilitando el análisis de información de la gerencia.

De esta investigación, se tomó como aporte las herramientas Visual Studio, Data tools para la generación de cubos OLAP con el análisis Services, así mismo la factibilidad de la automatización de reportes a un 100%, aumentando la productividad y disminuyendo horas hombres.

En el año 2016, Espinoza Vargas en la tesis denominada “Desarrollo de un Data mart para optimizar la generación de información estratégica de apoyo a la toma de decisiones en la vicepresidencia de banca comercial de Interbank”, en la Universidad San Martín de Porres, Perú. Trato la problemática del exceso de recursos para la generación de reportes en el departamento de vicepresidencia de banca comercial, ya que no existía oportunidades de mejoras para el área, ni para toda la organización ya que muchas áreas replicaban los mismos reportes según la necesidad, generando redundancia e inconsistencia y pérdida de información al no tener un almacén de datos histórica consolidada. El objetivo de esta tesis es minimizar la utilización de recursos a través de la creación de un modelo dimensional (data mart) que mejore la creación de datos estratégicos para apoyar la toma de decisiones. En conclusión, implementando el data mart se consiguió reducir a 1 hora con 37 min la generación de informes, lo cual es igual a un 90.6%. Obtener los datos real y en un menor tiempo permitirá generar mayores propuestas por parte de los gerentes para nuevas estrategias de ventas para incrementar los ingresos de los productos de menores colocaciones; así mismo establecer pautas para agilizar la atención y desembolso de los productos; medir la eficiencia de los ejecutivos de negocio identificando que recursos generan mayores eficiencias como también los que no aportan a la organización con la finalidad de potenciarlos para que aumenten su desempeño y analizar tendencias de consumo y minería de datos gracias a la información histórica consolidada. Todo ello se refleja en los aumentos notorios de las utilidades de la entidad financiera.



De esta investigación, se tomó en cuenta la reducción en un 100% de las labores operaciones de los analistas encargados del proceso de generación estratégica, ya que permitió usar la capacidad y tiempo de los analistas para desarrollar su verdadera función en el área que es analizar la información, permitiendo el incremento de estrategias para la generación de campañas para los clientes. Así mismo se tomará como aporte la metodología Ralph Kimball el desarrollo de un data mart en un banco.

### **1.3. Limitaciones de la investigación**

Para Arturo Morales y Antonio Morales (2014, p.23) el proceso crediticio es “un préstamo de dinero, donde se hace un compromiso un cliente a pagar la cantidad que recibió en el plazo o tiempo establecido según las pautas establecidas para dichos préstamos, más los intereses costos, devengados y seguros asociados si los hubiere”. Así mismo Gómez Antonio (2017, p.23) nos menciona que “el proceso crediticio cuenta con las siguientes fases:

- Determinación de un mercado objetivo
- Evaluación del crédito
- Evaluación de condiciones en que se otorgan
- Aprobación
- Documentación y desembolso
- Administración del crédito en referencia
- Seguimiento de riesgo”

Debido a que el proceso crediticio es un tema muy amplio para un Data mart, a partir de este párrafo se considerara como la variable dependiente la evaluación de créditos.

## **1.4. Teorías relacionadas al tema**

### **A. Evaluación de créditos**

Según Pérez Candiotti (2013, p.75), define que la evaluación de créditos “es el conjunto de documentaciones normativas como; políticas de crédito, manuales, procesos y reglamentos para establecer el nivel de riesgo de crédito y concluir si se rechaza o entrega el crédito. La evaluación crediticia puede decirse que es la necesidad de establecer razonablemente la capacidad y voluntad de pago que tiene el cliente”.

Por otro lado, los autores Pacheco Ivana y Valencia Eddy (2016) nos definen la evaluación de créditos al “procedimiento mediante el cual una persona capacitada recolecta información respecto al cliente solicitante y de acuerdo con la teoría financiera o al reglamento de la entidad donde se solicite el crédito se determinara si cumple al otorgamiento del crédito” (p.10).

Finalmente Arturo Morales y Jose Morales nos definen que la evaluación de crédito “es el análisis de la información que facilitara decidir si se niega o se presta un crédito, lo que habitualmente se toma en cuenta son los hábitos de pago y su capacidad del cliente para atender sus compromisos. Como indicación orientativa se deberá revisar al menos: Estados financieros e Informes de crédito” (2014, p.100).

### **Fases de la evaluación de créditos**

Según Salvador Molina (2015) para la evaluación del crédito existen riesgos de créditos, es por ello que nos menciona que “existen 3 fases principales para el riesgo de crédito: en primer lugar, la admisión, fase crítica que culmina con la aceptación o denegación de nuevas operaciones; en segundo lugar, el seguimiento de clientes y operaciones con objetivo de evitar situaciones de impago; y, por último, una vez producida el impago, la fase de recuperación, con el establecimiento de políticas de recobro para dirigir los esfuerzos según la situación de morosidad” (p.35).

## **Fase de admisión de un crédito**

Según Salvador Molina (2015, p.37) sostiene que “el proceso de admisión o concesión de una operación de crédito se considera la fase más crítica dentro del ciclo del crédito, además de evaluar unas series de variables relevantes para la aprobación de un crédito comercial, que son tanto cualitativas (posición competitiva del cliente o potencial, fortalezas y debilidades) y cuantitativas (ratio de liquidez y solvencia). Variables financieras para poder medir la capacidad para la admisión de un crédito comercial:

- Niveles de liquidez (ratio de liquidez, fondo de maniobra, ratio de solvencia).
- Niveles de endeudamiento (ratio de amortización).
- Rentabilidad
- Productividad y eficiencia”.

## **Fase de control y seguimiento**

Según Salvador Molina (2015, p.40) “La importancia de esta fase es que la empresa que otorga un crédito disponga de un sistema de seguimiento, control y mitigación del riesgo asociado a la cartera de clientes actual con objeto de disminuir la probabilidad de ocurrencia del riesgo y reducir el posible impacto de impago de un cliente. En cuanto a la medición en la fase de control y seguimiento del riesgo crediticio asociado a la cartera de clientes, se encuentra el seguimiento de la calidad crediticia que permite la detección temprana de deterioros.

Indicadores para la medición de la calidad crediticia:

- Índice de la morosidad (Cartera Vencida/Cartera Total)
- Índice de Cobertura (Estimación preventiva/Cartera Vencida)” (p.40).

## **Fase de recuperación**

Según Salvador Molina sostiene que “la recuperación de riesgos es la fase durante la cual se intenta gestionar de forma activa la recuperación de los créditos comerciales dudosos bien por impago, bien por otras razones que supongan deterioro de la cartera de riesgo comercial. El objetivo de esta fase es la recuperación del importe del crédito por rehabilitación dl riesgo dudoso o por recuperación de las cantidades debidas, de manera que sea mínima la pérdida crediticia” (2015, p.40).

## **Dimensión e indicadores en la evaluación de créditos**

### **Dimensión: Control y Seguimiento**

Para los autores Antonio Morales y Arturo Morales (2014, p.146) nos menciona que el control y seguimiento de créditos “son las acciones realizadas para evitar el impago de un cliente, disminuir el riesgo de mora en los portafolios de las cuentas por cobrar y asegurar los elementos necesarios para actuar en caso de aumento de riesgo por el aumento de los clientes que no realizan sus pagos”.

### **Indicadores:**

#### **Indicador 1: Índice de morosidad**

Según Pampillon nos menciona que el índice de morosidad “nos indica que porcentaje de los créditos se hallan en dificultades de ser devueltos” (2012, p.224).

$$\text{Índice de morosidad} = \left( \frac{\text{Crédito vencido}}{\text{crédito total}} \right)$$

#### **Indicador 2: Índice de cobertura**

Según Díez nos menciona que “El índice de cobertura se interpreta como la protección que cuentan los bancos ante los créditos no pagados que poseen en sus respectivas carteras. En concreto, se entiende al porcentaje de créditos

morosos que están cubiertos por provisiones que son realizadas por las entidades financieras” (2010, p.203).

$$\text{Índice de cobertura} = \left( \frac{\text{Provisiones para créditos}}{\text{Cartera atrasada}} \right)$$

## B. Data mart

Según Lurillo Michele define que “Un Data mart es un grupo de datos con el propósito de generar toma de decisiones a un área en específica. La información existente en este contexto puede ser agrupados, propagados y explorados de diferentes maneras que el personal pueda explotar la información de manera deseada según sus necesidades dentro del negocio” (2016, p.14).

Por otro lado, Thomas Gauchet (2015) define que “Un Data mart es la agrupación de tablas de información organizadas en una estructura que favorece la lectura para realizar un reporte analítico sobre información histórica más amplio que el que se conserva en producción. Los Data mart están completamente orientados al usuario final según a sus necesidades” (p.27).

Finalmente, los autores Conesa y Curto (2016) definen que un Data mart “una parte del Data warehouse que tiene como objetivo atender un análisis, función o necesidad determinada, la información está estructurada de forma de estrella o como de copo de nieve así mismo un data mart puede estar incluido dentro de un data warehouse o también puede trabajarse de manera independiente. El Data mart se implementa para satisfacer los requerimientos de un área de organización o de un grupo de trabajo” (p.31).

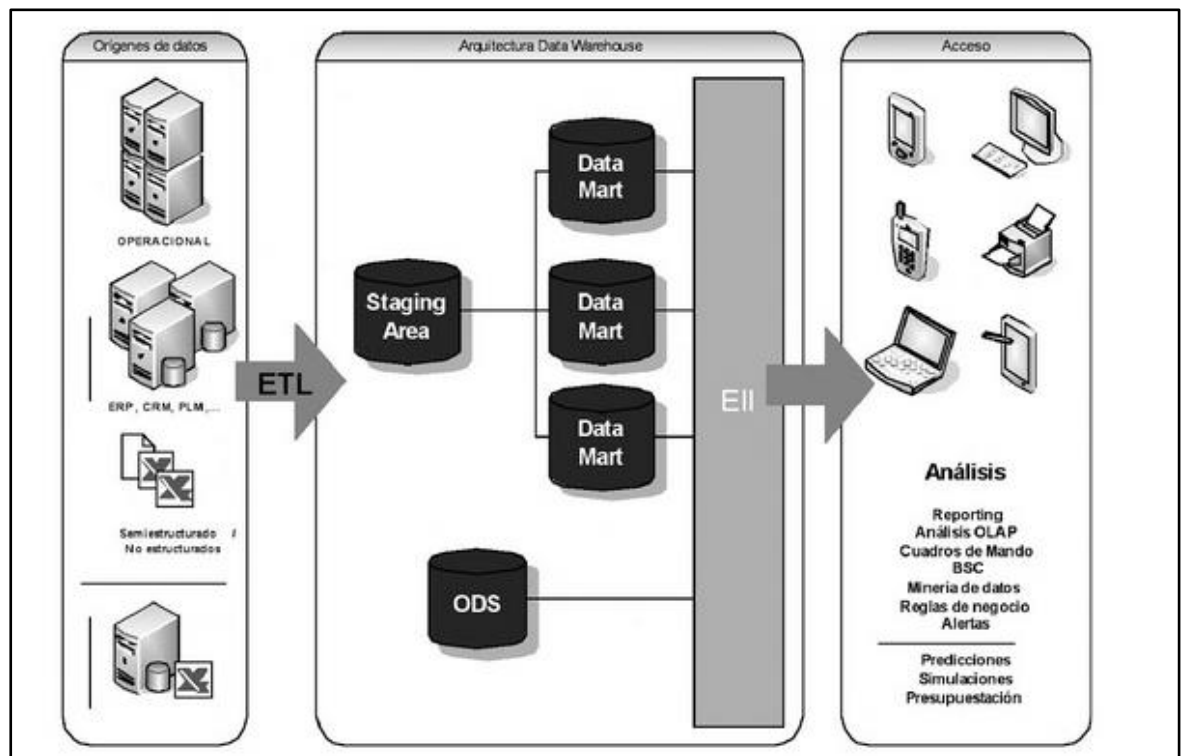
### **Arquitectura de un Data mart**

Los autores Conesa Caralt y Curto Díaz definen que “para la arquitectura de un data mart existen principalmente 3 enfoques:

- Enterprise Bus Architecture: Es una arquitectura que se basa en data mart de manera independiente que pueden hacer una staging área en el caso sea necesario.
- Corporate Information Factory: Existe un data warehouse organizacional en esta arquitectura y un conjunto de data mart dependientes.
- Enterprise Data warehouse 2.0: Es la metodología Bill Inmon para aplicar la experiencia por más de 20 años” (2016, p.39).

**Figura 4: Arquitectura de un Data mart**

Fuente: Conesa Caralt y Curto Díaz



## Componentes de un Data mart

### Tabla de hecho

Para Conesa y Curto, nos definen que una tabla de hecho “representa el proceso de la organización en el Data Warehouse. Un claro ejemplo sería el proceso de ventas, se consideraría una tabla con el nombre de hechos ventas” (2016, p.32).

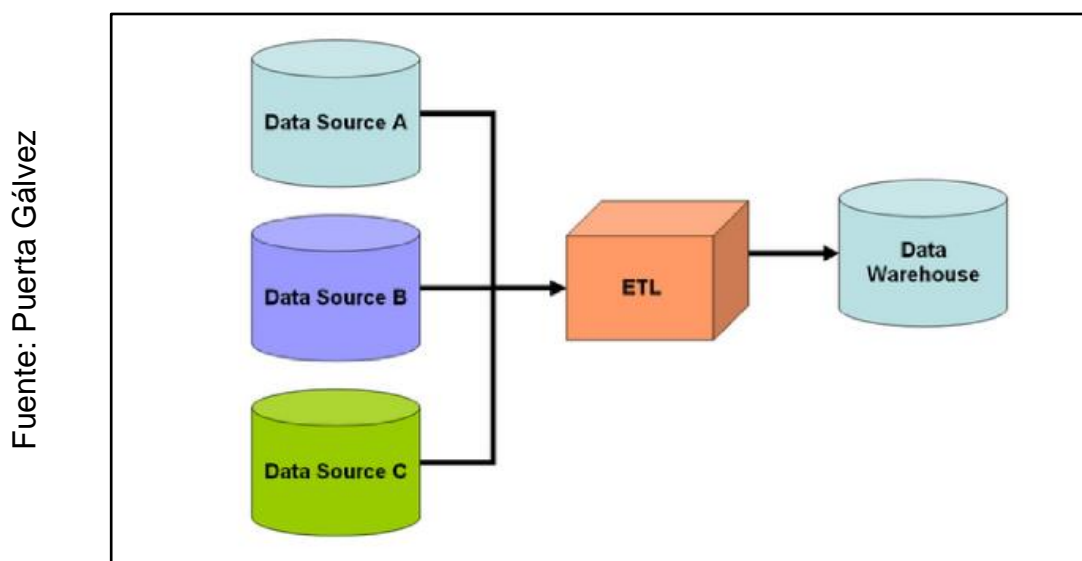
### Dimensión

Conesa y Curto nos dicen que “una dimensión podría considerarse un proceso de una organización y esta sería representando en el Data Warehouse como una dimensión” (2016, p.33).

### ETL

Para Puerta Gálvez (2016) un ETL (Extract, Transform and Load) es “un proceso destinado a la extracción, transformación y carga de información procedentes de un repositorio transaccional para un Data Warehouse. La extracción, transformación y carga es el más riguroso y el más largo de esta fase, pues es la extracción de la información de un repositorio de origen transaccional, puesto que se tiene que transformar y limpiar dicha información y cargarlo al Data Warehouse” (p.87).

**Figura 5: Extracción, transformación, carga**



## **Cubo**

Para Thomas Gauchet (2015, p.27) “un cubo es como un Data mart, solo que además contiene datos que han sido agregados previamente sobre los distintos niveles de los ejes de análisis. Estos agregados pre calculados permiten reducir considerablemente el tiempo de consulta a las peticiones. Hay varios tipos de agregación; el más habitual es la suma, Cuando se tratan de volúmenes de millones de registros, el uso de un cubo supone una enorme diferencia para todo dato que deba sumarse sobre cualquiera de los ejes de análisis”.

## **Esquema Estrella**

Según Carlos, Norberto y Pardillo “en un esquema estrella los hechos y las dimensiones se corresponden con tablas de hecho y dimensión. Un esquema estrella tiene en una tabla en el centro (tabla hecho) y alrededor dimensiones” (2013, p.34).

## **Esquema de constelaciones**

Según Carlos, Norberto y Pardillo (2013) “Este esquema es parecido al modelo estrella. La diferencia principal del modelo estrella y del modelo constelación es que consta de tablas de hechos para cada nivel de agregación. Por lo tanto, contiene más de una tabla de hechos (de ahí su nombre)” (p.41).

## **Esquema Copo de nieve**

Los autores Carlos, Norberto y Pardillo (2013) nos mencionan que “el esquema se diferencia del esquema estrella en que alguna de las dimensiones se normaliza según los niveles de jerarquía. En la tabla dimensión se encuentran los valores del mínimo nivel de jerarquía” (p.42).



## **Metodologías para el desarrollo de un Data mart**

### **Metodología Hefesto**

Los autores Bernabeau y Garcia nos dicen que “La metodología Hefesto es única a comparación de las metodologías actuales, experiencias únicas en las fases de construcción de fuentes de información. Cabe resaltar que HEFESTO siempre evolucionando y mejorando ya que se han tomado como aportes los diversos feedbacks que han aportado en diferentes países al momento de implementar la metodología” (2017, p.124).

### **Metodología Ralph Kimball**

Según Gustavo Rivadera menciona que “la metodología Kimball se enfoca en el ciclo de vida dimensional del Negocio. Dicho ciclo se basa en 4 principios, ofrecer la solución completa, realizar entregas en incrementos significativos, centrarse en el negocio y armar una infraestructura que se adecue” (2013, p.58).

### **Metodología Bill Inmon**

Según nos dice Bustos Barrera (2013) “la metodología fue desarrollada por Bill Inmon en su libro "Building the Data Warehouse" en el año 1992. Esta metodología del DW es conocido como una agrupación de información integrado con cambios no volátiles y en el tiempo, que su objetivo es dar apoyo a las toma de decisiones estratégicas de una organización (p.38).

### **Selección de la Metodología de desarrollo de un Data mart**

En la presente tesis se hizo una comparación de las tres metodologías para poder seleccionar cuál de esas metodologías se adapte mejor para la creación de un data mart. Se realizo un cuadro para las comparaciones de las metodologías (ver tabla 1) donde se explica los criterios, el cual será evaluador por expertos de ingeniería.

**Tabla 1: Criterios de selección para las metodologías de un Data mart**

Ítems	Criterios	Descripción
1	Los resultados y objetivos son fácil de comprender y se distinguen fácilmente.	Los objetivos de cada fase al momento de desarrollar la metodología para el desarrollo del cubo son fáciles de distinguirse y sencillos de realizarse.
2	Aplica modelos lógicos y conceptuales lo cual es sencillo de analizar e interpretar.	Al momento de desarrollarse la metodología, al crear los modelos conceptuales de la base de datos es de manera fácil de realizar, interpretar y analizar.
3	La involucración que tiene la metodología con los usuarios en todas las fases del proyecto	Al desarrollar la metodología siempre tiene que estar el usuario que lidera el proyecto en todas las etapas que se va desarrollando el data mart.
4	La metodología después de aplicarse es de fácil mantenimiento.	Al terminar el proyecto del data mart, se podrá realizar mantenimiento como modificaciones a lo desarrollado.
5	Tienen etapas completamente detalladas, que sumándolas dan como resultado la completitud	Las etapas de desarrollo son muy detallada y fácil de comprender en la metodología.
6	Genera menor tiempo de consultas.	Según la metodología, al momento de que ya se haya implementado, que tan difícil será consultar la información almacenada.

**Fuente: Elaboración Propia**

## Evaluación para el seleccionamiento de la metodología para la elaboración de un data mart

En la presente investigación se consultó a expertos de ingeniería para la validación de selección de la metodología que más se adapte a esta investigación, colocando un puntajes tal como se concibe en la tabla N°2

**Tabla 2: Expertos para la calificación de desarrollo de las Metodologías de un Data mart**

Expertos	Grado	Ralph Kimball	Hefesto	Bill Inmon
Flores Masías, Edward José	Magister	35	15	13
Cueva Villavicencio, Juanita Isabel	Magister	33	31	11
Saavedra Jiménez, Roy	Magister	35	21	7
<b>Promedio</b>		<b>34.3</b>	22.3	10.3

**Fuente: Elaboración Propia**

En la tabla N°2, se recurrió a 3 expertos y de acuerdo con los criterios propuestos, se colocó puntajes a la metodología que se adecua a la investigación. El mayor promedio como se observa en la tabla 2, la Metodología Ralph Kimball obtuvo el puntaje de 34.3, es por ello, para el desarrollo del Data mart se aplicó la Metodología Ralph Kimball.

### **Metodología seleccionada: Ralph Kimball**

#### **Metodología Ralph Kimball**

Para Gustavo Rivadera menciona que “la metodología Kimball se enfoca en el ciclo de vida dimensional del negocio. Dicho ciclo se basa en 4 principios” (2013, p.58).

### **Centrarse en el negocio:**

Según Gustavo Rivadera menciona que “se debe centrar en identificar los requerimientos del negocios, y usar esos requerimientos para implementar una relación con el negocio, para facilitar el desarrollo y un mejor análisis de los implementadores” (2013, p.58).

Según Gustavo Rivadera elaborar una infraestructura de datos apropiada: “Diseñar una fuente de datos integrada, única y de que tenga la facilidad de utilizarse, con un alto rendimiento donde se mostrara todos los requerimientos solicitados por la organización” (2013, p.58).

### **Realizar entregas en incrementos significativos:**

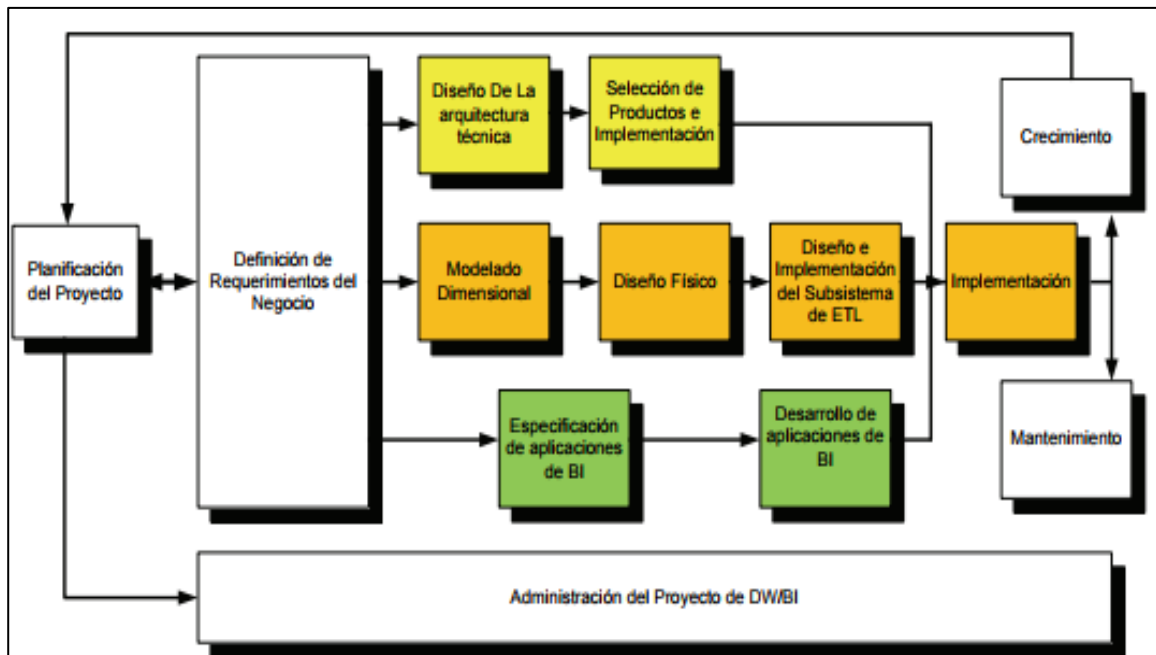
Según Gustavo Rivadera menciona que “crear una fuente de información (Data Warehouse) con un aumento de entregables en plazos de seis a doce meses. Se utilizará el valor de la organización en cada elemento hallado para encontrar el orden de ejecución. En esto hay una similitud con las metodologías de elaboración de software” (2013, p.58).

### **Ofrecer la solución completa:**

Según Gustavo Rivadera menciona que “proveer todo lo necesario para poder entregar a los usuarios de la organización un valor. Es decir que se contara con una información sólida, confeccionada y con una calidad accesible y probada. Así mismo se proporcionará herramientas de consulta ad hoc, alguna aplicación para realizar análisis avanzados e informes y una documentación” (2013, p.58).

**Figura 6: Fases de la metodología Ralph Kimball**

Fuente: The Kimball Group Reader



### **Pasos para desarrollar la metodología Kimball**

#### **Planificación:**

Según Gustavo Rivadera menciona que “Se determina el propósito de la implementación del DW, su alcance y objetivos específicos, los principales peligros y un acercamiento inicial a las necesidades de los datos” (2013, p.60).

#### **Análisis de requerimientos:**

Según Gustavo Rivadera menciona que “los requerimientos se obtienen en un proceso de entrevistar al trabajador del negocio, siempre tiene se tiene que ir con una preparación. Se debe observar cómo es que funciona el negocio, como funciona la competencia, la industria y hasta los mismos clientes. Se debe entender cómo funciona a detalle la organización, leer la estrategia interna y sobre todo entender la terminología y términos del negocio” (2013, p.60).

## **Modelo Dimensional**

Para Rivadera (2013), nos dice que “la implementación de un modelo dimensional es un proceso alternante iterativo y dinámico. Cuenta con 4 pasos el proceso iterativo:

- Escoger el proceso de negocio.
- Seleccionar el nivel del grano.
- Seleccionar las dimensiones.
- Hallar la tabla hechos y sus medidas” (p.62).

## **Diseño Físico**

Para Rivadera (2013), nos dice que “en esta fase, se deben intentar aclarar las siguientes preguntas.

¿Qué tan grande será el Data Warehouse/BI?

¿Cuál de los factores tendrán una configuración más compleja y grande?

¿Cuánta capacidad de memoria y la cantidad de servidores se necesitarán? ¿Qué marca de procesadores o almacenamiento se necesitarán?” (p.67).

## **Extracción, Transformación y carga (ETL)**

Para Rivadera (2013), nos dice que “Es la fuente que nutre el DW. Si el sistema ETL se desarrolla correctamente, se extrae la información del sistema transaccional, aplicar distintas reglas para incrementar la consistencia y calidad, consolidar los datos que vienen de diferentes sistemas, y al final cargar los datos en el Data Warehouse en un formato adecuado para la explotación con las herramientas para realizar análisis” (p.68).

## **Especificación y desarrollo de aplicaciones de BI**

Para Rivadera (2013), nos dice que “El fundamento de todo proyecto de BI es proporcionar una forma estructurada y de fácil acceso al almacén de información. Este acceso se da con las herramientas de explotación de BI” (p.68).

## **Herramienta para el desarrollo de un Data mart**

### **Sistema gestor de base de datos**

#### **ORACLE**

Para los autores Jiménez, Santander, Meléndez, Suarez y Calabria nos mencionan que Oracle “es una base de datos de próxima generación diseñada para satisfacer estas necesidades, proporcionando una nueva arquitectura multiusuario en la parte superior de una plataforma de base de datos rápida, escalable, confiable y seguro” (2017, p.88).

#### **Visual Studio**

Según Thomas Gauchet (2015, p.85) “Visual Studio permite desarrollar proyectos. Un proyecto es un componente aplicativo de un cierto tipo: un almacén de información, una aplicación de Windows o un grupo de flujos ETL. Visual studio permite agrupar los proyectos en soluciones. Una solución puede agrupar los diferentes componentes de la arquitectura BI”.

## **1.5. Formulación del problema**

### **Problema Principal**

¿De qué manera influye un Data mart en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco?

### **Problema Secundario**

¿De qué manera influye un Data mart en el índice de morosidad en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco?

¿De qué manera influye un Data mart en el índice de cobertura en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco?

## **1.6. Justificación del estudio**

### **Justificación Institucional**

Josep Curto (2016) nos dice que, “un Data mart proporciona una visión conformada, persistente, calidad, histórica y única toda la información relevante para la organización” (p.24).

La creación de un Data mart en la empresa Mibanco va a beneficiar considerablemente en la evaluación de créditos generando un mejor perfilamiento de clientes que no caerán en mora y así el área de campañas tendrá un prestigio de generar buenos perfiles de clientes, y así cuando otras áreas de Mibanco deseen impulsar un nuevo beneficio de desembolso a clientes, darán a escoger al área de campañas el seleccionamiento de dichos clientes.

### **Justificación económica**

Según Avellaneda Rojas (2015) nos menciona que “a través del desarrollo de un Data mart, realizando un análisis de sus indicadores concluyo que tuvo un aumento de colocaciones bancarias de 41.46% evidenciando la factibilidad de inversión ya que hubo una ganancia de cerca de \$100,000 mostrando un beneficio importante”

Con la implementación de esta tesis el usuario podrá contar con información optima y centralizada, la cual permitirá sacar mejores perfilamientos de clientes y disminuir la mora en campañas además de aumentar los desembolsos de clientes y la fidelización de los mejores clientes generando ganancias importantes a la organización y sobre todo seguir siendo el banco pionero de los clientes mype. Actualmente en el área de campañas comerciales hay un saldo deteriorado de 5, 989,099 de soles, el cual se plantea que implementando el Data mart esa pérdida



se disminuirá considerablemente y se obtendrá ganancias de 10, 000,000 de soles en 2 meses.

### **Justificación Operativa**

Según Josep Curto (2016) nos plantea que, “La implantación de un sistema dimensional genera muchos beneficios, de lo cual podemos destacar:

- Permite una visión única, histórica, calidad y persistente de la información.
- Crear indicadores representativos para la empresa.
- Entregar información al más mínimo detalle.
- Mejora la competencia dentro de la organización y como resultado de ser capaces de diferenciar lo relevante sobre lo superfluo, acceder más rápido a la información, tener mayor agilidad en la toma de decisiones” (p.20).

Implementar un Data mart permitió emitir reportes más eficientes y de manera automatizada, también permitió realizar consultas a bases pesadas de manera muy rápida, y a la vez permitió realizar un mejor análisis para la captación de los mejores clientes de Mibanco, mediante una interfaz amigable del sistema, garantizando la operabilidad del Data mart.

### **Justificación tecnológica**

Los autores Marzón, Pardillo y Trujillo (2013) plantean que, “Las tecnologías de la información, facilitan reducir las barreras físicas entre mercados y empresas. Las empresas se enfrentan a entornos más rigurosos con un gran índice de competencia. Dado esta situación es necesario aplicar decisiones estratégicas que ayuden a una organización generar ventajas competitivas respecto a otras entidades” (p.2).

Mientras existan muchos sistemas transaccionales en Mibanco, más compleja es la forma de extraer información crucial para analizarla y explotarla, lo cual cambiara implementando el Data mart en Campañas Comerciales, para obtener una información más consolidada y centralizada para poder explotarla de manera más sencilla. De esta manera implementar el Data mart permitió cumplir con los objetivos estratégicos de Mibanco.

## **1.7. Hipótesis**

### **Hipótesis General**

**HG:** El Data mart mejora la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

### **Hipótesis Específicas**

**HE1:** El Data mart disminuye el índice de morosidad en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

**HE2:** El Data mart incrementa en el índice de cobertura en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco

## **1.8. Objetivos**

### **General**

Desarrollar un Data mart para la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

### **Específicos**

**O1:** Determinar la influencia de un Data mart en el índice de morosidad en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

**O2:** Determinar la influencia de un Data mart el índice de cobertura en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

## **II.MÉTODOS**

## **2.1. Diseño de investigación**

### **Método de investigación: Hipotético Deductivo**

Según los autores Hernández, Ramos, Placencia, Indacochea, Quimis y Moreno nos mencionan que “a partir de determinados principios, teorías o leyes se derivan respuestas que explican el fenómeno y que vuelven a ser confirmadas en la práctica” (2018, p.95).

En esta tesis fue seleccionado el método Hipotético – Deductivo, porque según la problemática descrita se plantea una hipótesis y se demuestra realizando uso de datos.

### **Tipo de estudio**

#### **Explicativa**

Los autores Baptista, Hernández y Fernández (2014, p.108) plantean lo siguiente “el estudio explicativo se centraliza en revelar por qué sucede un acontecimiento y en qué situaciones se revela, o por dos o más variables se relacionan”.

Para esta investigación se seleccionó el estudio de tipo Explicativo ya que se busca explicar el motivo que originaron la situación analizada y establecer conclusiones en la investigación.

#### **Experimental**

Según Baena (2014) nos menciona lo siguiente “se presenta mediante la manipulación de una variable experimental no comprobada la investigación experimental, en condiciones rigurosamente controladas, con el fin de describir de qué modo o por qué motivo se realiza un acontecimiento o situación” (p.14).

En la presente investigación se seleccionó el estudio de tipo Experimental porque ya que se buscará describir de qué manera o porque causa se realiza un acontecimiento o una situación particular.

## **Aplicada**

Según Baena (2014) nos dice que “Su objetivo es el estudio de una problemática destinada a la realización. Esta investigación puede aportar nuevos hechos si lo aplicamos de manera adecuada en la investigación, de tal manera que se pueda confiar en los resultados obtenidos, dicha información podría ser estimable y útil para aplicarlo en la teoría” (p.11).

La presente investigación es una investigación aplicada, ya que su objetivo es saber el impacto que tiene un Data mart para la evaluación de créditos en el área de campañas, así mismo si se requiere realizar una intervención en la investigación tipo aplicada deben de realizarse con calidad, en consecuencia, son analizados en diferentes etapas del resultado y proceso esta investigación.

## **Diseño de estudio: Pre-Experimental**

Los autores Baptista, Hernández y Fernández (2014) nos explican que “este diseño es llamado así porque su nivel de intervención es pequeño. Diseño de preprueba-posprueba con un solo grupo. A un grupo se le aplica una prueba previa al estímulo o tratamiento experimental, después se le administra el tratamiento y finalmente se le aplica una prueba posterior al estímulo. Este diseño ofrece una ventaja sobre el anterior: existe un punto de referencia inicial para ver qué nivel tenía el grupo en las variables dependientes antes del estímulo; es decir, hay un seguimiento del grupo” (p.345).

Se diagrama así:



Pre Test y Post Test diseños de medición

G: Grupo experimental: Es medir la muestra que se seleccionó para la evaluación de los créditos para medir el índice de cobertura y índice de morosidad.

X: Data mart: Es la aplicación del Data mart en la evaluación de créditos en la empresa Mibanco. A través de 2 evaluaciones (Pre Test y Post Test) se realizara la medición si el Data mart realiza modificaciones en la evaluación de créditos de la empresa Mibanco.

O1: Pre test: Medición antes de aplicar el data mart a un grupo experimental en la evaluación de los créditos en la empresa Mibanco. El post test servirá como medición con este pre test.

O2: Post Test: Medición de un grupo experimental después de aplicar el Data mart en la evaluación de los créditos en la empresa Mibanco. Las 2 mediciones del pre y post test serán comparadas y ayudarán a validar la disminución de la mora en Mibanco antes y después del desarrollo del Data mart.

En la investigación se aplicará el diseño pre- experimental, ya que valdrá como una primera aproximación a la problemática de la investigación, así mismo sirve como estudio exploratorio y pretenda administrar la evaluación de créditos en la forma de pre-prueba y pos-prueba.

## **2.2. Variables Operacionalización**

### **Definición Conceptual**

Las 2 variables de la presente tesis son la siguiente:

#### **Variable Independiente (VI): Data mart**

Para Conesa Jordi y Curto Josep (2016) nos mencionan que un Data mart es “una parte del Data warehouse que tiene como objetivo responder a una necesidad determinada, los datos están se encuentran de forma estructurada con un modelo de forma de estrella o como de copo de nieve así mismo un Data mart puede ser parte de un Data warehouse o también puede trabajarse de manera independiente. El Data mart se implementa para solucionar problemas en distintas áreas” (p.31).

#### **Variable Dependiente (VD): Evaluación de créditos**

Según Perez Candiotti (2013) nos menciona que la evaluación de créditos “es el conjunto de documentaciones normativas como; políticas de crédito, manuales, procesos y reglamentos para establecer el nivel de riesgo de crédito y concluir si se rechaza o entrega el crédito. La evaluación crediticia puede decirse que es la necesidad de establecer razonablemente la capacidad y voluntad de pago que tiene el cliente” (p.75).



## **Definición Operacional**

### **Variable Independiente (VI): Data mart**

El data mart es un repositorio histórico de datos de forma estructura que se alimenta de diferentes fuentes de datos, centralizando toda la información y además cuenta con una estructura dimensional que permite el análisis de información de manera óptima y centralizada, agilizando las consultas a través de las dimensiones o tablas hechos y facilitando la generación de reportes.

### **Variable Dependiente (VD): Evaluación de créditos**

La evaluación de créditos consiste en la selección del perfil de clientes que accederán a una campaña para poder desembolsar un préstamo a través de un análisis de información, posteriormente ese perfil seleccionado será presentado a los gerentes para dar su conformidad o rechazo. Finalmente, una vez aceptado el perfil del cliente se procederá a generar la base de clientes para desplegarse a nivel nacional y así mismo gestionarse por los asesores de negocios.

En la tabla número 3 se encuentra la operacionalización de las variables.

**Tabla 3: Operacionalización de las variables**

TIPO	VARIABLE	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	Dimensión	INDICADORES	ESCALA DE MEDICIÓN
VARIABLE INDEPENDIENTE	Data mart	Según Gauchet (2015) "Un Data mart es un conjunto de tablas de datos organizadas en una estructura que favorece la lectura para realizar un reporte analítico sobre información histórica más amplio que el que se conserva en producción. Los Data mart están completamente orientados al usuario final según a sus necesidades" (p.27)	Un Data mart es un almacén de datos histórico que se alimenta de diferentes fuentes de datos, centralizando toda la información y además cuenta con una estructura dimensional que permite el análisis de información de manera óptima y centralizada, agilizando las consultas a través de las dimensiones o tablas hechos y facilitando la generación de reportes.			
VARIABLE DEPENDIENTE	Evaluación de créditos	"Es el análisis de la información que facilitara decidir si se niega o se presta un crédito, lo que habitualmente se toma en cuenta son los hábitos de pago y su capacidad del cliente para atender sus compromisos. Como indicación orientativa se deberá revisar al menos: Estados financieros e Informes de crédito" (Arturo Morales y Jose Morales, 2014, p.100).	La evaluación de créditos consiste en la selección del perfil de clientes que accederán a una campaña para poder desembolsar un préstamo a través de un análisis de información, posteriormente ese perfil seleccionado será presentado a los gerentes para dar su conformidad o rechazo. Finalmente, una vez aceptado el perfil del cliente se procederá a generar la base de clientes para desplegarse a nivel nacional y así mismo gestionarse por los asesores de negocios.	Control y seguimiento	Índice de cobertura  índice de morosidad	Razón  Razón

**Fuente: elaboración propia**

## Indicadores

En la tabla número 4 se visualiza los indicadores del proceso de evaluación de créditos:

**Tabla 4: Indicadores de la evaluación de créditos**

Indicador	Descripción	Técnica	Instrumento	Unidad de medida	Formula
Índice de Morosidad	“Nos indica que porcentaje de los créditos se hallan en dificultades de ser devueltos” (Pampillon, 2012, p.224).	Fichaje	Ficha de Registro	Unidad	$\text{Índice de morosidad} = \left( \frac{\text{Crédito vencido}}{\text{crédito total}} \right)$
Índice de Cobertura	“Este Índice, entendido como fórmula de protección antes problemas para el futuro, es uno de los indicadores más utilizados para evaluar la calidad crediticia de las empresas” (Diez, 2010, p.203).	Fichaje	Ficha de registro	Unidad	$\text{Índice de cobertura} = \left( \frac{\text{Provisiones para créditos}}{\text{Cartera atrasada}} \right)$

**Fuente: Elaboración propia**

## 2.3. Población y Muestra

### Población

En esta investigación se conformará la población por 2 indicadores que se presenta para la evaluación de créditos en el área de campañas comerciales respecto a los desembolsos.

Los autores Diez y Juez (2015) nos indican que “Población: Se designa con este término a cualquier conjunto de elementos que tienen unas características comunes. Cada uno de los elementos que integran tal conjunto recibe el nombre de individuo” (p.5).

La población para los indicadores de índice de morosidad y índice de cobertura fueron todos los desembolsos que se realizaron en el mes de abril durante el periodo de lunes a sábados.

**Tabla 5: Población**

Población Total
92 desembolsos

Fuente: Elaboración propia

### Muestra

Los autores Diez y Juez nos menciona que “Muestra: Subgrupo de la población de donde se recolectan la información que debe ser representativa de la población” (2015, p.6).

La población es conocido para los 2 indicadores que son:

Para encontrar la población se aplicará la siguiente formula:

$$n = \frac{Z^2 pq}{e^2 + \frac{Z^2 pq}{N}}$$

Donde:

n: Tamaño de la muestra

Z: Nivel de confianza deseado

E: Error estándar

N: Tamaño de la población

p: Proporción de la población con la característica deseada

q: Proporción de la población sin la característica deseada

$$n = \frac{1.96^2 * 0.5 * 0.5}{0.05^2 + \frac{1.96^2 * 0.5 * 0.5}{92}}$$
$$n = \frac{0.9604}{0.01282414} = 74.89 \rightarrow n \cong 75$$

Para la presente investigación se aplicó el muestreo probabilístico y de tipo aleatorio simple, siendo recomendable ya que, Valderrama Mendoza nos menciona “se selecciona de manera aleatoria a los individuos de la población, hasta contar con la muestra deseada, así mismo en lo teórico se cuentan con anticipación los elementos y después se seleccionan de acuerdo con los números aleatorios” (2013, p.72).

## **2.4. Técnicas e instrumentos de recolección de datos**

### **Técnica**

Según Gil Pascual nos menciona que “Engloba todos los medios técnicos que se utilizan para facilitar el tratamiento o registrar las observaciones” (2016, p.9).

### **Técnica: Fichaje**

Los autores Parraguez, Chunga, Flores y Romero (2017, p.150) nos definen que “el fichaje es la técnica que permite el registro de información seleccionada para el proceso de investigación. Su aplicación requiere el uso de fichas para ayudarnos a recoger y organizar la información extraídas de diversas fuentes de interés, según la investigación”.

### **Instrumento**

Para Gil Pascual nos menciona que el instrumento es el “objeto con entidad independiente y externa y con recursos medios utilizados para obtener y registrar la información” (2016, p.9).

### **Instrumento: Ficha de registro**

Carlos Muñoz nos dice que una ficha de registro “Sirve para anotar los datos y registrar o identificar la fuente o fuentes documentales consultadas. Las fichas de registro se corresponden con cada una de las fuentes documentales y dan origen a la bibliografía, del informe de investigación” (2015, p.80).

El investigador pedirá los datos que serán extraídas del sistema transaccional de Mibanco para evaluar los desembolsos que cayeron en mora, para medir el Pre Test.

2 fichas de registros son las que se usaran:

Ficha de registro “Cartera Morosa”

Ficha de registro “Cartera Clientes”

Se elaboró una ficha de registro para el indicador índice de morosidad donde se registró el desembolso y la mora (ver Anexo 5) para obtener el resultado en unidades durante un mes.

Se elaboró una ficha de registro para el indicador índice de cobertura donde se registró la provisión de un crédito y el monto en mora (ver anexo 5) para obtener el resultado en unidades durante 1 mes.

En la tabla 6 indica las técnicas e instrumentos para la recolección de datos

**Tabla 6: Recolección de datos**

Indicador	Descripción	Técnica	Instrumento	Formula
índice de morosidad	Se evaluará el índice de morosidad de Mibanco	Fichaje	Fichaje de registro	Crédito dudoso/Crédito total
Índice de cobertura	Se evaluará el índice que cuenta Mibanco para los créditos morosos	Fichaje	Ficha de registro	(Provisiones crédito/Cartera atrasada) *100

Fuente: Elaboración propia

**Validez:**

Según Fontes de Gracia explica que “la validez de una investigación como «el grado de confianza que puede adoptarse respecto a la veracidad o falsedad de una determinada investigación”. (2015, p 68).

Las fichas de registro son los instrumentos a utilizarse, se validó con 3 expertos cada indicador, tal cual se observan en las tablas 7 y 8.

**Indicadores:**

**Índice de morosidad**

**Tabla 7: Validez para el índice de morosidad**

Nº	Expertos	Grado Académico	Puntaje
1	Gálvez Tapia Orleans	Magister	80.0
2	Chumpe A, Juan	Magister	71.0
3	Ordoñez Perez, Adilio	Doctor	91.0

Fuente: Elaboración Propia

Las fichas de registro se presentaron para que sean validado por 3 expertos (ver Anexo 6), la evaluación obtuvo un puntaje promedio de 80% mostrando una confianza alta demostrando que estos instrumentos son lo correcto para capturar la información de los indicadores.

**Índice de cobertura**

**Tabla 8: Validez para el índice de cobertura**

Nº	Expertos	Grado Académico	Puntaje
1	Gálvez Tapia Orleans	Magister	80.0
2	Chumpe A, Juan	Magister	71.5
3	Ordoñez Perez, Adilio	Doctor	91.0

Fuente: Elaboración Propia

Se presentaron fichas de registro para que sean validado por 3 expertos (ver Anexo 6), la evaluación obtuvo un puntaje promedio de 80% mostrando una confianza alta de que los instrumentos que son los correctos para capturar la información de los indicadores.



## **Confiabilidad**

Según Vilalta Perdomo nos menciona que “la confiabilidad radica en que la medición este bien realizada que al realizarse más de una vez siempre se obtenga el mismo resultado” (2016, p.30).

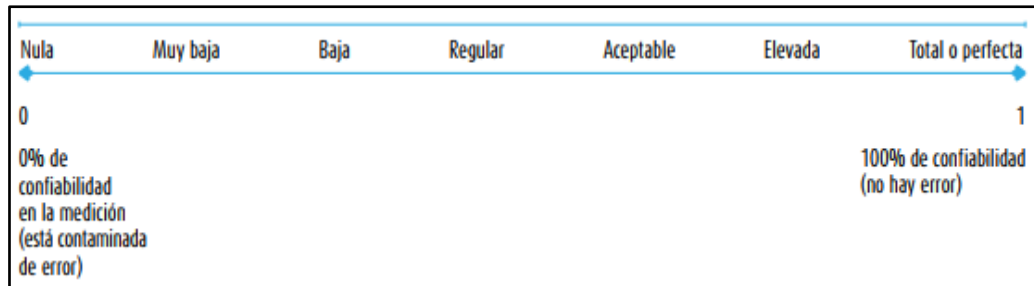
Los autores Betanzos, Escoto y Chávez (2017), menciona que “la confiabilidad mide la consistencia de un instrumento. Existen varios tipos de confiabilidad, por ejemplo, la temporal o test-retest, que consiste en que si una escala da un resultado en este momento y tiempo después se aplica la misma escala, entonces los resultados deben ser muy parecidos” (p.102).

Según Martinez Mediano (2014) define que “test-retest es la aplicación del instrumento en dos o más ocasiones dejando un intervalo de tiempo adecuado entre las aplicaciones”. (p.12)

Según Hernandez, Fernandez, y Baptista (2014) dicen que “hay diversos procedimientos para calcular la confiabilidad de un instrumento de medición. Todos utilizan procedimientos y fórmulas que producen coeficientes de fiabilidad. La mayoría oscilan entre cero y uno, donde un coeficiente de cero significa nula confiabilidad y uno representa un máximo de confiabilidad (fiabilidad total, perfecta). Cuanto más se acerque el coeficiente a cero, mayor error habrá en la medición.” (p.85).

Fuente: Hernández, Roberto  
Fernández, Carlos, Baptista, Pilar.

**Figura 7: Cálculo de la confiabilidad o fiabilidad**



Según Hernández, Fernández, y Baptista (2014) mencionan que el coeficiente de correlación de Pearson “es una prueba estadística para analizar la relación entre dos variables medidas en un nivel por intervalos o de razón. Se le conoce también como “coeficiente producto-momento”. El coeficiente de correlación de Pearson se calcula a partir de las puntuaciones obtenidas en una muestra en dos variables. Se relacionan las puntuaciones recolectadas de una variable con las puntuaciones obtenidas de la otra, con los mismos participantes o casos”.

**Tabla 9: Confiabilidad para Índice de morosidad**

		Correlaciones	
		TEST_INDICE_MOROSIDAD	RE_TEST_INDICE_MOROSIDAD
TEST_INDICE_MOROSIDAD	Correlación de Pearson	1	,741**
	Sig. (bilateral)		0.000
	N	24	24
RE_TEST_INDICE_MOROSIDAD	Correlación de Pearson	,741**	1
	Sig. (bilateral)	0.000	
	N	24	24

\*\* La correlación es significativa en el nivel 0,01 (bilateral).

Fuente: Elaboración propia.

Como se observa en el anexo 5 y tabla 9 el análisis de la confiabilidad del indicador índice de morosidad según el SPSS muestra el resultado de 0.74 que indica un nivel aceptable de confiabilidad.

**Tabla 10: Confiabilidad para el índice de cobertura**

Correlaciones

		IND_COBER_TEST	IND_COBER_RE_TEST
IND_COBER_TEST	Correlación de Pearson	1	,854**
	Sig. (bilateral)		0.000
	N	24	24
IND_COBER_RE_TEST	Correlación de Pearson	,854**	1
	Sig. (bilateral)	0.000	
	N	24	24

\*\* La correlación es significativa en el nivel 0,01 (bilateral).

Fuente: Elaboración propia.

Se muestra en el anexo 5 y tabla 10 el análisis de la confiabilidad del indicador del índice de cobertura de la evaluación de crédito según el SPSS muestra el resultado de 0,85 que indica un nivel aceptable.

## 2.5. Métodos de análisis de datos

En esta investigación se trabaja el procesamiento y análisis de la información que se tuvo como resultado con los instrumentos a través de la recolección de datos, para evaluar los resultados se utiliza la estadística a partir de procesar la información y realizar la validación de la hipótesis general, así mismo la hipótesis específica.

### Prueba de Normalidad

Para Vilalta Perdomo “Las pruebas de normalidad tienen por objetivo probar la hipótesis de que los valores de una variable aleatoria continua es una muestra representativa proviene de una población que sigue un comportamiento normal” (2016, p.81). Se presentan las siguientes pruebas estadísticas: Shapiro- Wilks, AndersonDarling y Kolgomorov-Smirnov.

**Anderson-Darling:** Según Quiñones (2011) indica que para para saber si la población de donde se extrajo la muestra es normal. Se tiene como valor crítico más utilizado de superarlo la hipótesis de normalidad que es rechazada (aceptándose H1), a un nivel de confianza o significancia:

- Si supera 1.029 entonces se acepta H1 para un nivel de prueba de 1% (99% de confianza).

- Si supera 0.870 entonces se acepta H1 para un nivel de prueba de 2.5% (97.5% de confianza).

- Si supera 0.751 entonces se acepta H1 para un nivel de prueba de 5% (95% de confianza).

- Si supera 0.632 entonces se acepta H1 para un nivel de prueba de 10% (90% de confianza).

Calcular el valor de P no siempre es posible, y si ya está calculado, el criterio anterior puede ser interpretado de manera práctica, si ese valor supera el nivel de prueba, se acepta la hipótesis nula ya que los datos seguirían una distribución normal.

**Prueba de normalidad de Shapiro Wilks:** Según García, Vaquer y Gomis (2010) indica que “es una prueba estadística que se utiliza para verificar la hipótesis de que una muestra procede de una distribución normal. Se acompaña de un valor de probabilidad que para valores igual o inferior a 0.05 indicaría diferencias con respecto a una distribución normal”. (p.18)

**Prueba de normalidad de Kolgomorov Smirnov:** Quiñones nos indica que interpretar el valor de P es más sencillo de interpretar, si dicho valor es mayor al nivel de prueba se confirma que los datos son normales ( $p > 0.05$  para un nivel de confianza del 95%).

## **Hipótesis estadística**

Según Valderrama Mendoza (2013) nos indica que “las hipótesis alternativas y nulas se presentan cuando se transforman en símbolos estadísticos, además se pueden analizar y recolectar para comprobar o rechazar las hipótesis” (p.122).

### **Hipótesis Específicas:**

**Hipótesis HE1:** El Data mart disminuye el índice de morosidad en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

### **Variables:**

**IIMa:** El índice de morosidad antes de utilizar el Data mart.

**IIMd:** El índice de morosidad después de utilizar el Data mart.

**H0:** El Data mart no disminuye el índice de morosidad en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

$$\text{H0: IIMa} - \text{IIMd} < 0$$

**Ha:** El Data mart disminuye el índice de morosidad en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

$$\text{Ha: IIMa} > \text{IIMd}$$

**Hipótesis HE2:** El Data mart disminuye en el índice de cobertura en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

**IICa:** El índice de cobertura antes de utilizar el Data mart.

**IICd:** El índice de cobertura después de utilizar el Data mart.

**H0:** El Data mart no disminuye el índice de cobertura en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

$$H_0: IICa - IICd < 0$$

**Ha:** El Data mart disminuye en el índice de cobertura en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

$$H_a: IICa < IICd$$

**Nivel de significancia:**

La significancia tuvo un nivel utilizado de  $x = 5\%$  (error), equivalente a 0.05, esto permite realizar una equivalencia para que se decida aceptar o rechazar la hipótesis. Nivel de confiabilidad:  $(1-x) = 0.95$

Margen de error:  $x = 0.05$

Nivel de confiabilidad:  $(1-x) = 0.95 = 95\%$

### Estadística de prueba:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

Donde:

N= Tamaño de la muestra pre test

M= Tamaño de la muestra post test

S<sub>1</sub>= Varianza pre test

S<sub>2</sub>= Varianza post test

X= Media pre test

Y= Media post test

Región de rechazo

La región de rechazo es  $t = t_x$

Donde  $t_x$  es tal que:

$P [t > t_x] = 0.05$ , donde  $t_x$  = Valor Tabular

Luego Región de Rechazo:  $t > t_x$

Cálculo de la Media

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

## Cálculo de la Varianza

$$\delta^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$$

## Desviación estándar

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

Dónde:

$\bar{x}$  = Media

$\delta^2$  = Varianza

$S^2$  = Desviación Estándar

$X_i$  = Dato i que está entre (0,n)

$\bar{X}$  = Promedio de los datos

$n$  = Número de datos

## 2.6. Aspectos éticos

La información del grupo de control y experimental se procesó de manera correcta y realizarse alguna adulteración ya que los datos se encuentran cimentados en el instrumento que se aplicó a los grupos de estudio.

Se aseguro que toda la información obtenida para realizar la parte estadística haya sido completamente confidencial con respecto a los créditos y clientes de Mibanco.

Los lineamientos y reglamentos de la Universidad Cesar vallejo fueron aplicados en la presente tesis.



Se garantizo a base de criterios de transparencia y prudencia la confidencialidad de toda la información.

El proyecto es completamente original es por ello por lo que no existe ningún trabajo similar en la institución donde se realiza el estudio.

Los resultados obtenidos por ningún motivo han sido adulterados o plagiados de alguna otra investigación de tesis y se realizó un correcto uso de la tesis en beneficio de todos.

Se conservo toda la data de los desembolsos de Mibanco que participo en esta tesis y los resultados que se obtuvo de manera muy confidencial.

### **III. Resultados**

### 3.1 Análisis Descriptivo

En el estudio se aplicó un Data mart para evaluar el índice de morosidad y el índice de cobertura en la evaluación de créditos; para ello se aplicó un Pre Test que permitió conocer las condiciones iniciales de los indicadores y Pos Test que permitió conocer las condiciones finales de los indicadores. Los resultados descriptivos de estas medidas se observan en las tablas 11 y 12.

#### Indicador: Índice de morosidad

Los resultados descriptivos del índice de morosidad en la evaluación de créditos antes y después de implementado el Data mart.

**Tabla 11: Medidas descriptivas del índice de morosidad en la evaluación de créditos después de la implementación del data mart**

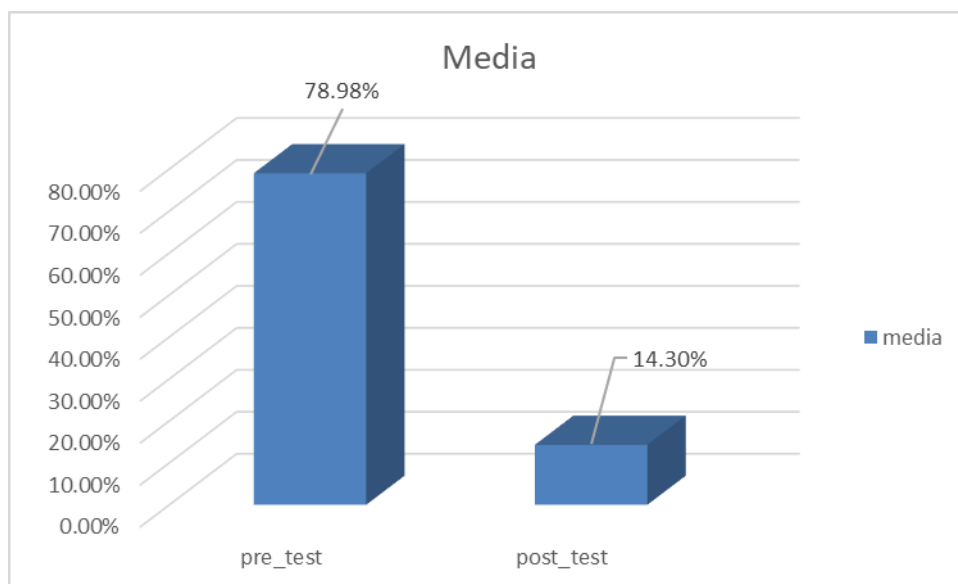
Estadísticos descriptivos					
	N	Mínimo	Máximo	Media	Desviación estándar
ind_morosidad_pre	25	,69	,92	,7898	,06835
ind_morosidad_post	25	,06	,19	,1430	,04252
N válido (por lista)	25				

Fuente: Elaboración propia

En la tabla 11, se puede ver para el caso del índice de morosidad en la evaluación de créditos, en el Pre-test se obtuvo un valor de 78.98% mientras que en el post-test fue de 14.30% tal como se aprecia en la figura 8; esto indica una gran diferencia antes y después de la implementación del Data mart; así mismo, el índice de morosidad mínima fue del 69% antes, y 6% después de la implementación del Data mart.

En cuanto a la dispersión del índice de morosidad, en el Pre-test se tuvo una variabilidad de 6.8%; sin embargo, en el post – test se tuvo un valor de 43%.

**Figura 8 Índice de morosidad en la evaluación de créditos antes y después de la implementación del Data mart**



Fuente: Elaboración propia

**Indicador: Índice de cobertura**

Los resultados descriptivos del índice de cobertura de la evaluación de créditos se observan en la tabla 12.

**Tabla 12: Medidas descriptivas del índice de cobertura de la evaluación de créditos antes y después de implementar el Data mart**

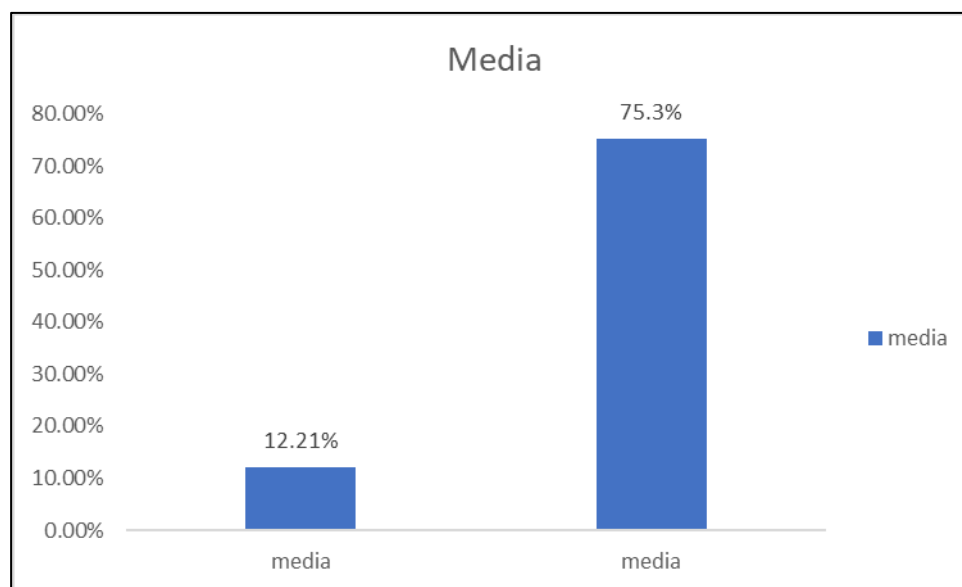
Estadísticos descriptivos					
	N	Mínimo	Máximo	Media	Desviación estándar
ind_cobertura_pre	25	,05	,18	,1221	,03372
ind_cobertura_post	25	,60	,99	,7532	,09860
N válido (por lista)	25				

Fuente: Elaboración propia

En el caso del índice de cobertura de la evaluación de créditos, en el pre-test se obtuvo un valor de 12.21% mientras que en el post-test fue de 75.3% tal como se aprecia en la figura 9; esto indica una gran diferencia antes y después de implementar el data mart; así mismo, el índice de cobertura de la evaluación de créditos mínima fue de 5% antes, y 60% (ver tabla 10) después de la implementación del Data mart.

En cuanto a la dispersión del índice de cobertura en la evaluación de créditos, en el pre-test se tuvo una variabilidad de 3.37%; sin embargo, en el post-test se tuvo un valor de 10%.

**Figura 9 Índice de cobertura antes y después de implementar el Data mart**



Fuente: Elaboración propia

### **3.2 Análisis inferencial**

#### **Prueba de Normalidad**

Se procedió a realizar la prueba de normalidad para los indicadores de índice de morosidad y índice de cobertura a través del método del Shapiro-Wilk, debido a que el tamaño de la muestra está conformado por 25 documentos de estratificados por día durante un mes y es menor a 50.

Si:

Sig. < 0.05 adopta una distribución no normal.

Sig.  $\geq$  0.05 adopta una distribución normal.

Los resultados fueron los siguientes:

**Indicador: Índice de morosidad**

Con el objetivo de seleccionar la prueba de hipótesis; los datos fueron sometidos a la comprobación de su distribución, específicamente si los datos del índice de morosidad contaban con distribución normal.

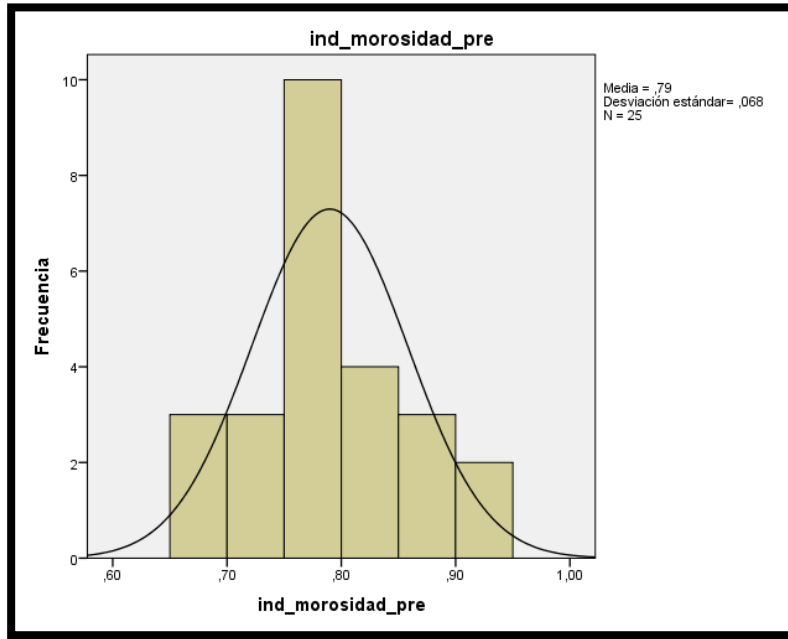
**Tabla 13: Pruebas de Normalidad del índice de morosidad en la evaluación de créditos antes y después de implementar el Data mart**

Pruebas de normalidad			
	Shapiro-Wilk		
	Estadístico	gl	Sig.
ind_morosidad_pre	,944	25	,184
ind_morosidad_post	,923	25	,060
a. Corrección de significación de Lilliefors			

Fuente: Elaboración propia

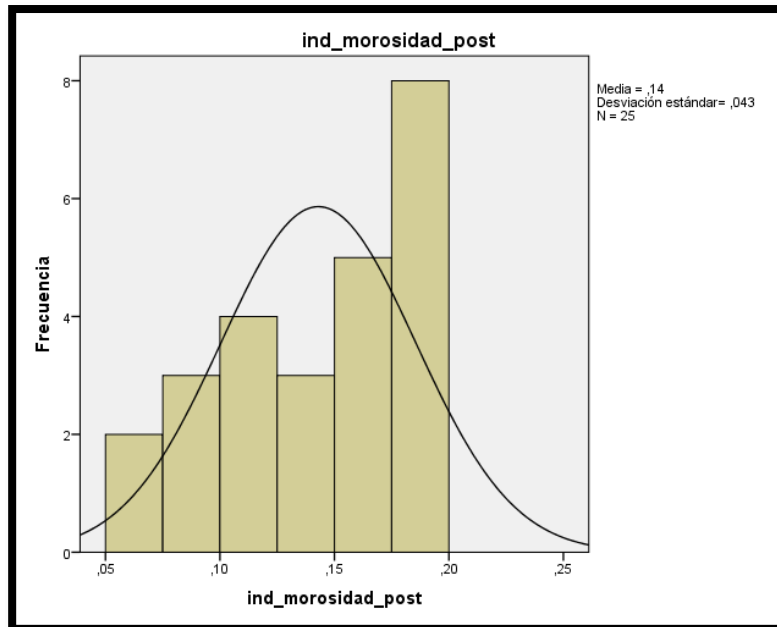
Como se muestra en la tabla 13 los resultados de la prueba indican que el índice de morosidad en la evaluación de créditos fue de 0.184, cuyo valor es mayor que 0.05, por lo que indica que el índice de morosidad se distribuye normal. Los resultados de la prueba del Post Test indican que el índice de morosidad fue de 0.060, cuyo valor es mayor que 0.05, por lo que indica que el índice de morosidad se distribuye normalmente. Lo que confirma la distribución normal de ambos datos de la muestra, se puede apreciar en las figuras 10 y 11.

**Figura 10 Prueba de normalidad del índice de morosidad antes de implementar el Data mart**



Fuente: Elaboración propia

**Figura 11 Prueba de normalidad del índice de morosidad después de implementar el Data mart**



Fuente: Elaboración propia

### Indicador: Índice de cobertura

Con el objetivo de seleccionar la prueba de hipótesis; los datos fueron sometidos a la comprobación de su distribución, específicamente si los datos del nivel de eficacia del índice de cobertura de la evaluación de créditos contaban con una distribución normal.

**Tabla 14: Pruebas de Normalidad del índice de cobertura en la evaluación de créditos antes y después de implementar el Data mart**

	Pruebas de normalidad		
	Shapiro-Wilk		
	Estadístico	gl	Sig.
ind_cobertura_pre	0.939	25	0.142
ind_cobertura_post	0.953	25	0.293

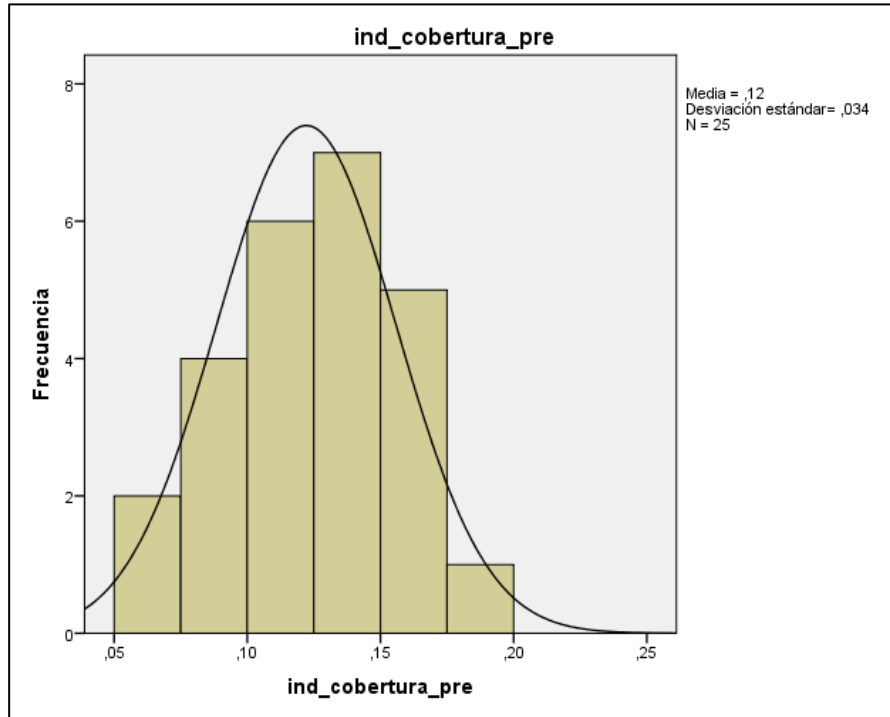
a. Corrección de significación de Lilliefors

Fuente: Elaboración propia

En la tabla 14 se muestra los resultados de la prueba que muestran que el índice de cobertura de la evaluación de créditos en el Pre test fue de 0.142 cuyo valor es mayor que 0.05, por lo que indica que el índice de cobertura en la evaluación de créditos se distribuye de manera normal. Los resultados de la prueba del Post Test indican que el índice de cobertura de la evaluación de créditos fue de 0.293, cuyo valor es mayor que 0.05, por lo que el índice de cobertura se distribuye normalmente. Lo que confirma la distribución normal de ambos datos de la muestra, se puede apreciar en las figuras 12 y 13.

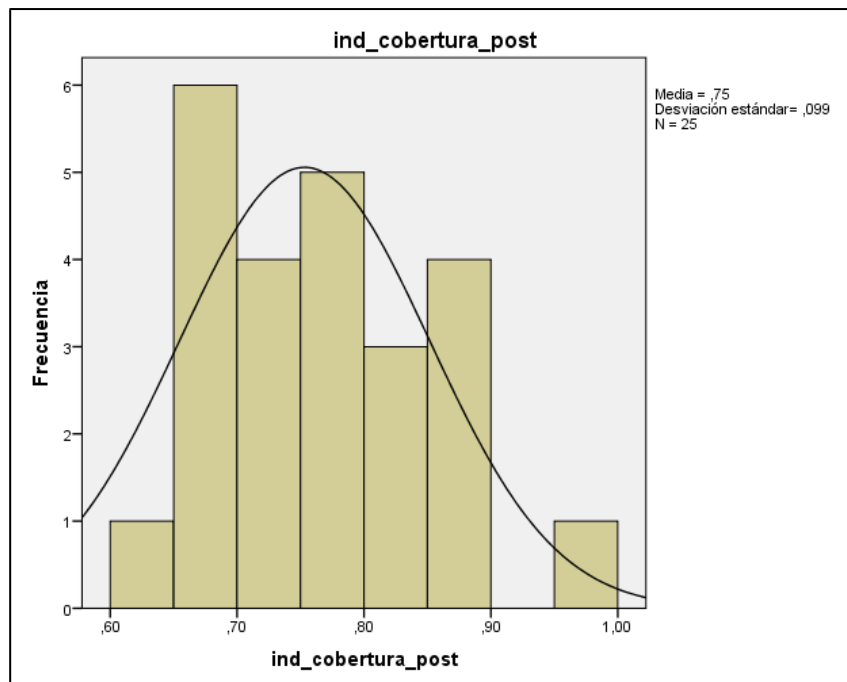


**Figura 12 Prueba de normalidad del índice de cobertura antes de implementar el Data mart**



Fuente: Elaboración propia

**Figura 13 Prueba de normalidad del índice de cobertura después de implementar el Data mart**



Fuente: Elaboración propia

### 3.3 Prueba de hipótesis

#### Hipótesis de investigación 1:

- **H1:** El Data mart disminuye el índice de morosidad en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.
- **Indicador:** índice de morosidad.

#### Hipótesis estadísticas

#### Definiciones de variables:

**TEa:** Índice de morosidad antes de utilizar el Data mart.

**Ted:** Índice de morosidad después de utilizar el Data mart.

**H0:** Data mart no disminuye el índice de morosidad en la evaluación de créditos en la empresa Mibanco.

$$H_0: TEa \geq TEd$$

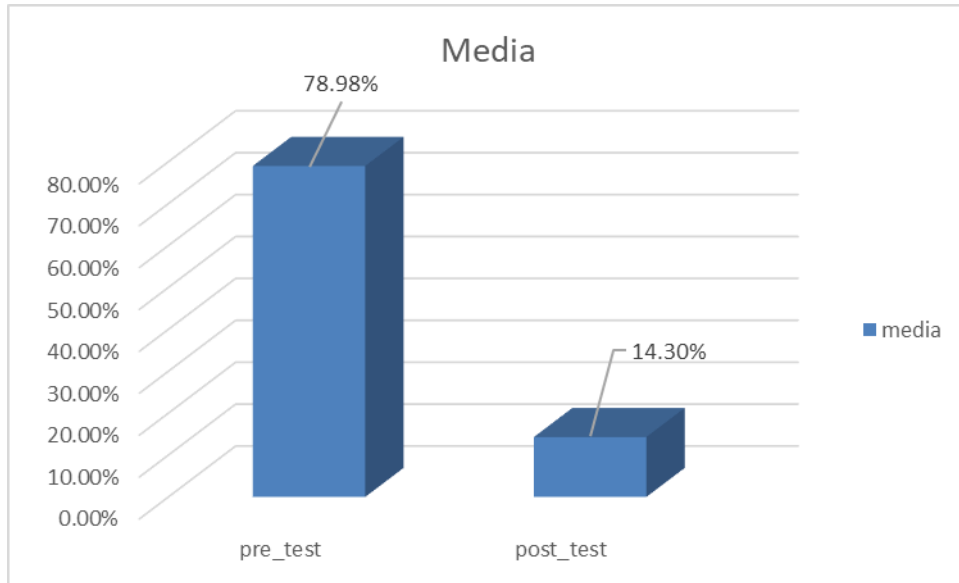
**Ha:** Data mart disminuye el índice de morosidad en la evaluación de créditos en la empresa Mibanco.

$$H_a: TEa < TEd$$

El indicador con el Data mart es mejor que el indicador sin el Datamart

En la figura 14, el índice de morosidad (Pre test), es de 78.98% y el Pos test es 14.3%.

**Figura 14 índice de morosidad - Comparativa general**



Fuente: Elaboración propia

En la figura 14 se llega a la conclusión que existe una disminución en el índice de morosidad, se puede verificar con una comparación de las medias, que disminuyo de 78.98% a 14.30%.

En cuanto al resultado del contraste de hipótesis se aplicó la Prueba de T-Student, debido a que los datos obtenidos durante la investigación (Pre test y Post test) se distribuyen normalmente. El valor de T contraste es de 38,694

El cual es claramente mayor que 1.7109 (ver tabla 15).

**Tabla 15: Prueba de T-Student para el índice de morosidad en la evaluación de créditos antes y después de implementado el Data mart**

z		prueba t de Student		
	Media	t	GI	Sig. (bilateral)
Antes	0.7898	38,694	24	0.000
Después	0.1430			

Fuente: Elaboración propia

Reemplazando entonces en T:

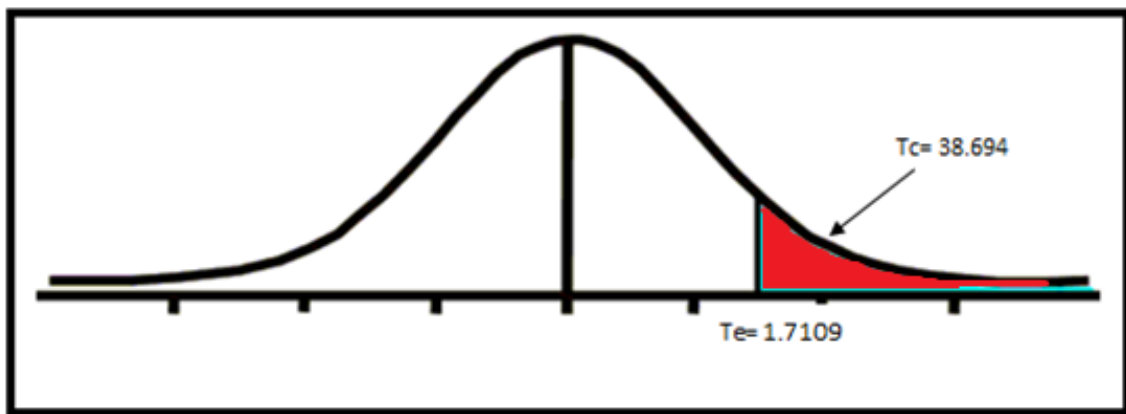
$$T_c = \frac{0.6468}{\sqrt{0.0002 + 0.00008}}$$

$$T_c = \frac{0.6468}{0.0167}$$

$$T_c = 38.694$$

Se rechaza la hipótesis nula, y tomando la hipótesis alterna con una confianza del 95%. Así mismo el valor de T que se obtuvo tal como se ve en la figura 15, se encuentra ubicada en la zona de rechazo. Por lo tanto, el Data mart disminuye el índice de morosidad en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco en el año 2018.

**Figura 15 Prueba T-Student- Índice de morosidad**



Fuente: Elaboración propia

### Hipótesis de investigación 2:

- **H2:** El Data mart incrementa el índice de cobertura en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco.

- **Indicador:** índice de cobertura.

### **Hipótesis estadísticas**

#### **Definiciones de variables:**

**TBa:** Índice de cobertura antes de utilizar el Data mart.

**TBd:** Índice de cobertura después de utilizar el Data mart.

**H0:** Data mart no incrementa el índice de cobertura en la evaluación de créditos en la empresa Mibanco.

$$\mathbf{H0:TBa-TBd < 0}$$

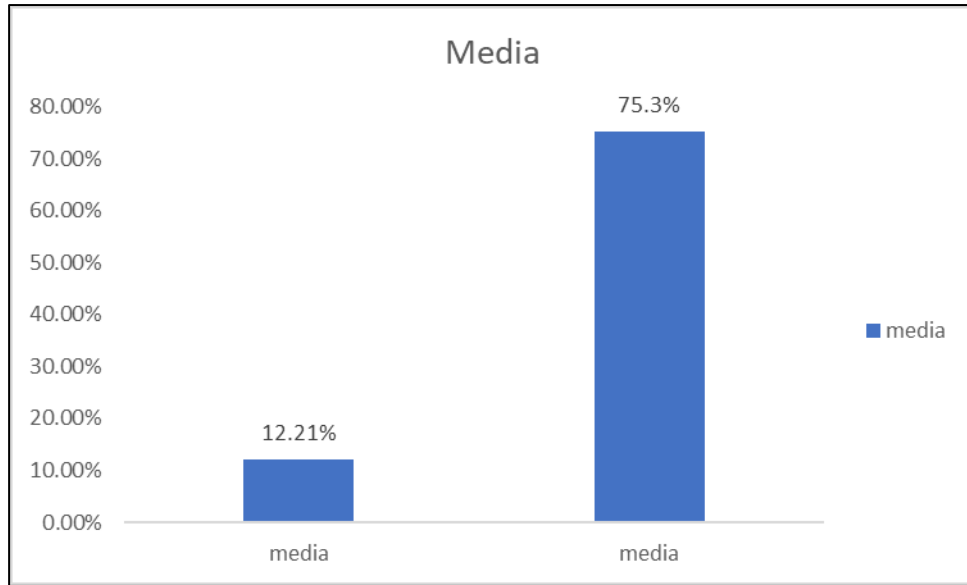
**Ha:** Data mart incrementa el índice de cobertura en la evaluación de créditos en la empresa Mibanco.

$$\mathbf{Ha:TBa-TBd > 0}$$

El indicador con el Data mart es mejor que el indicador sin el Datamart

En la figura 16, el índice de morosidad (Pre test), es de 0.1221 y el Pos test es 32.52.

**Figura 16 índice de cobertura - Comparativa General**



Fuente: Elaboración propia

En la figura 16 se concluye que existe un crecimiento en el índice de cobertura de la evaluación de créditos, lo cual se puede validar observando las medias, que asciende de 12.21% al valor de 75.3%.

En cuanto al resultado del contraste de hipótesis se aplicó la Prueba T-Student debido a que los datos obtenidos durante la investigación (Pre test y Post test)

Se distribuyen normalmente. El valor de T contraste es de -37.160 el cual es claramente menor que -1.7109. (ver tabla 16)

**Tabla 16: Prueba de T-Student para el índice de cobertura en la evaluación de créditos antes y después de implementado el Data mart**

z		prueba t de Student		
	Media	t	Gl	Sig. (bilateral)
Antes	0.1221	-37.16	24	0.000
Después	0.7530			

Fuente: Elaboración propia

Reemplazando entonces en T:

$$Tc = \frac{-0.6311}{\sqrt{\frac{0.001}{25} + \frac{0.10}{25}}}$$

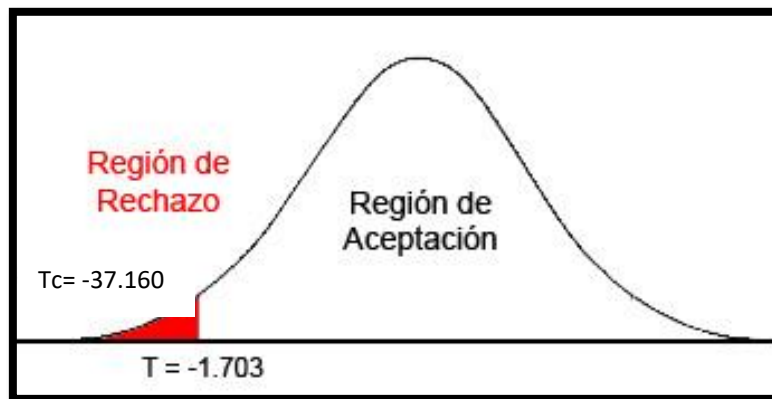
$$Tc = \frac{-0.6311}{0.08491/\sqrt{25}}$$

$$Tc = \frac{-0.6311}{0.016982}$$

$$Tc = -37.160$$

Se rechaza la hipótesis nula, y tomando la hipótesis alterna con una confianza del 95%. Así mismo el valor de T que se obtuvo tal como se ve en la figura 17, se encuentra ubicada en la zona de rechazo. Entonces el data mart incrementa el índice de provisiones en la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco en el año 2018

**Figura 17 Prueba T-Student - Índice de cobertura**



Fuente: Elaboración propia

## **IV. DISCUSIÓN**



## DISCUSIÓN

En esta tesis como resultado de obtuvo que el data mart logró disminuir el índice de morosidad en la evaluación de créditos de 78% a 14%, lo que equivale a una disminución de 64%. De la misma manera Hidalgo Ortiz en la tesis de maestría llamada “El análisis inteligente de datos y su incidencia en el otorgamiento de créditos financieros en la cooperativa de ahorro y crédito Cushun Ltda”, llegó a la conclusión que análisis de información influyen en la disminución de la morosidad, en su investigación disminuyó el índice de morosidad a 70%.

También se tuvo como resultado que el Data mart aumento el índice de cobertura de la evaluación de créditos de 12% a 75%, lo que es igual a un aumento promedio de 63%, de la misma manera Hidalgo Ortiz en la tesis de maestría denominada “El análisis inteligente de datos y su incidencia en el otorgamiento de créditos financieros en la cooperativa de ahorro y crédito Cushun Ltda”, llegó a la conclusión que con el data mart se incrementa el índice de cobertura, en su investigación incremento el índice de cobertura a 80%.

En la presente investigación el resultado mencionado demuestra que la utilización de una herramienta de BI facilita la información de manera oportuna y de fácil acceso para los procesos confirmando así que el Data mart para la evaluación de créditos disminuye el índice de morosidad en un 64% e incrementa el índice de cobertura en la evaluación de créditos en 63%; de los resultados que se obtuvo determina que el Data mart mejora la evaluación de créditos.

## **V. CONCLUSIONES**

## CONCLUSIONES

Para esta tesis se tiene como conclusión que el Data mart mejora la evaluación de los créditos en campañas comerciales de la empresa Mibanco pues permitió la disminución del índice de morosidad e incrementar el índice de cobertura, con lo que se llegó a alcanzar el objetivo de esta tesis.

Para esta tesis se tiene como conclusión que el Data mart disminuyo el índice de morosidad en un 64%. Por lo tanto, se afirma que el Data mart disminuye el índice de morosidad en la evaluación de créditos.

Para esta tesis se tiene como conclusión que el Data mart incremento el índice de cobertura en la evaluación de créditos en 63%. Se confirma que el Data mart agranda el índice de cobertura en la evaluación de créditos.

## **vi. RECOMENDACIONES**

## Recomendaciones

Es recomendable ampliar esta investigación con el propósito de seguir mejorando la evaluación de los créditos y otros procesos que se relacionan a este, de esta manera Mibanco podrá mantener de forma continua una mejor evaluación de créditos y generar ventaja la competitiva.

Se recomienda concientizar al área de campañas de usar esta solución de negocio y seguir creando diversos indicadores para explotarse a través del data mart con la finalidad de seguir mejorando en la evaluación de créditos.

Posteriormente se puede construir un data warehouse corporativo, siendo un tema de una posible ampliación de esta tesis, abarcando más áreas como productos activos o pasivos para contar con una visión global de la organización

Se sugiere seguir ampliando esta investigación con un módulo que permita predecir los próximos clientes propensos a caer en mora.

## **VII. Referencias Bibliográficas**

## REFERENCIAS BIBLIOGRÁFICAS

AVELLANEDA Rojas, Ramón. Implementación de un Data mart como herramienta de mejora en la toma de decisiones del servicio de colocaciones bancarias en una entidad financiera del estado. Tesis (Ingeniero de sistemas). Lima: Universidad San Martín, 2015.

BAENA, Guillermina. Metodología de la investigación. México: Grupo editorial Patria, 2014. 11pp. ISBN: 978-607-744-003-1.

BAENA, Guillermina. Metodología de la investigación. México: Grupo editorial Patria, 2014. 14pp. ISBN: 978-607-744-003-1

BETANZOS, Fabiola; ESCOTO, María.; CHAVEZ, Joana. Estadística aplicada en Psicología y Ciencias de la salud. Editorial El Manual Moderno, 2017. 96pp. ISBN: 978-607-448-640-7

BUSTOS Sofía y MOSQUERA Verónica, Análisis, diseño e implementación de una solución de business Intelligence para la generación de indicadores y control de desempeño, en la empresa Otecel S.A, utilizando la metodología Hefesto v2.0. 2012. Tesis (Ingeniero de sistemas). Sangolquí: Escuela politécnica del ejército, 2013. 38pp.

CARLOS, Juan, NORBERTO, José, PARDILLO, Jesús. Diseño y explotación de almacenes de datos. Conceptos de modelado multidimensional. España: Club universitario, 2013. 34pp. ISBN 9788499485461.

CARLOS, Juan, NORBERTO, José, PARDILLO, Jesús. Diseño y explotación de almacenes de datos. Conceptos de modelado multidimensional. España: Club universitario, 2013. 41pp. ISBN 9788499485461.

CARLOS, Juan, NORBERTO, José, PARDILLO, Jesús. Diseño y explotación de almacenes de datos. Conceptos de modelado multidimensional. España: Club universitario, 2013. 42pp. ISBN 9788499485461.

Castillo, Nicolás. Mora de pymes creció más de 37% por sobreendeudamiento [En línea]. El Comercio. Perú. 14 de enero del 2014. [Fecha de consulta: 10 de setiembre del 2018]. Disponible en: <https://elcomercio.pe/economia/peru/mora-pymes-crecio-37-sobreendeudamiento-165016>.

CONESA, Jordi y CURTO, Josep. Introducción al Business Intelligence. Rambla del Poblenou. Ed. UOC, 2016. 24pp. ISBN 9788497888868.

CONESA, Jordi y CURTO, Josep. Introducción al Business Intelligence. Rambla del Poblenou. Ed. UOC, 2016. 31 pp. ISBN 9788497888868.

CONESA, Jordi y CURTO, Josep. Introducción al Business Intelligence. Rambla del Poblenou. Ed. UOC, 2016. 32pp. ISBN 9788497888868.

CONESA, Jordi y CURTO, Josep. Introducción al Business Intelligence. Rambla del Poblenou. Ed. UOC, 2016. 33pp. ISBN 9788497888868.



CONESA, Jordi y CURTO, Josep. Introducción al Business Intelligence. Rambla del Poblenou. Ed. UOC, 2016. 39pp. ISBN 9788497888868.

DÍEZ, Ignacio. Cómo entender las finanzas de hoy. Barcelona-España: Antoni Bosch, editor, S.A., 2010. 203pp. ISBN:978-84-95348-54-8

El estudio y la investigación documental: Estrategias metodológicas y herramientas TIC por Parraguez María [et al.]. Perú: editorial Gerardo chungu Chinguel, 2017. 150pp. ISBN: 978-612-00-2603-8

Espinoza, Jenny. Desarrollo de un data mart para optimizar la generación de información estratégica de apoyo a la toma de decisiones en la vicepresidencia de banca comercial de interbank. Tesis (Ingeniero de sistemas). Lima: Universidad san Martín de Porres, 2016.

Fundamentos de investigación en psicología por Fontes de Gracia Sofia [et al.]. España: Editorial UNED, 2015. 115pp. ISBN: 9788436260557

GAUCHET, Tomas. SQL server 2014 implementación de una solución de Business Intelligence. España: Ediciones ENI, 2015. 27pp. ISBN: 978-2-7460-9458-1

GAUCHET, Tomas. SQL server 2014 implementación de una solución de Business Intelligence. España: Ediciones ENI, 2015. 27pp. ISBN: 978-2-7460-9458-1

GAUCHET, Tomas. SQL server 2014 implementación de una solución de Business Intelligence. España: Ediciones ENI, 2015. 85pp. ISBN: 978-2-7460-9458-1

GIL, Juan. Técnicas e instrumentos para la recogida de información. Madrid: Editorial UNED, 2016. 9pp. ISBN: 978-84-362-7128-7

GÓMEZ Antonio, Noemi. La evaluación crediticia y la morosidad de créditos comerciales en la financiera confianza S.A. Huánuco. Tesis (Contador público). Perú: Universidad de Huánuco, 2017. 23pp.

Guizado Verdezoto. Diseño e implementación de un Data mart OLAP para el análisis gerencial académico, que será implementado en la unidad educativa La Colina. Tesis (ingeniero de sistemas). Ecuador: Universidad Central del Ecuador, 2015. 39pp.

HERNÁNDEZ, Roberto, FERNÁNDEZ, Carlos y BAPTISTA, Pilar. Metodología de Investigación. México: Universidad de Cuenca, 2014. 99pp. ISBN: 9684229313

HERNÁNDEZ, Roberto, FERNÁNDEZ, Carlos y BAPTISTA, Pilar. Metodología de Investigación. México: Universidad de Cuenca, 2014. 100pp. ISBN: 9684229313

HIDALGO, Aydee. El análisis inteligente de datos y su incidencia en el otorgamiento de créditos financieros en la cooperativa de ahorro y crédito Cushun Ltda. Tesis (ingeniero de sistemas), Perú: Universidad técnica de Ambato, 2017.

JUEZ, Pedro y Díez, Francisco. Probabilidad y estadística en medicina aplicaciones en la práctica clínica y en la gestión sanitaria. Madrid: Días de Santos, S.A., 2015. 5pp. ISBN: 84-7978-278-1

JUEZ, Pedro y Díez, Francisco. Probabilidad y estadística en medicina aplicaciones en la práctica clínica y en la gestión sanitaria. Madrid: Días de Santos, S.A., 2015. 6pp. ISBN: 84-7978-278-1

LURILLO, Michele. Como implementar la inteligencia empresarial en su negocio en nueve preguntas más una. 3a edición. México, 2016. 14pp. ISBN: 978-14-25328-648

MARTÍNEZ, Catalina. Técnicas e instrumentos de recogida y análisis de datos. Madrid: Editorial UNED, 2014. 120pp. ISBN: 9788436268225

Mayo Claudio. Diseño de un Data mart para la gestión de cartera de clientes de la banca mayorista de una entidad bancaria. Tesis (ingeniero de sistemas). Perú: Universidad Nacional de Ingeniería, 2013.

Metodología de la investigación científica por Hernández Arturo [*et al.*]. México: editorial 3ciencias, 2018. 95pp. ISBN: 978-84-948257-0-5.

Mía, Ríos. Créditos: 1.4 millones de personas tienen deudas en más de dos entidades financieras. Gestión. 08 de junio del 2016. [Fecha de consulta: 25 de setiembre del 2018]. Disponible en: <https://gestion.pe/tu-dinero/creditos-1-4-millones-personas-deudas-dos-entidades-financieras-136843>.

MORALES, Arturo y MORALES, José. Crédito y cobranza. México: grupo editorial patria, S.A. DE C.V, 2014. 100pp. ISBN: 978-607-438-840-4.

MORALES, Arturo y MORALES, José. Crédito y cobranza. México: grupo editorial patria, S.A. DE C.V, 2014. 146pp.ISBN: 978-607-438-840-4.

MORALES, Arturo y Morales, José. Crédito y cobranza. México: grupo editorial patria, S.A. DE C.V, 2014. 23pp.ISBN: 978-607-438-840-4.

MUÑOZ, Carlos. Metodología de la investigación. México: Oxford University Press, 2015.80pp. ISBN: 9786074265422.

Muñoz, Geldi. Crece en mora en créditos del sistema financiero [En línea]. elPeriodico. Guatemala. 17 de octubre del 2017. [Fecha de consulta: 03 de setiembre del 2018]. Disponible en:  
<https://elperiodico.com.gt/inversion/2017/10/17/crece-mora-en-creditos-del-sistema-financiero/>.

PACHECO Ivana y VALENCIA Eddy. Evaluación crediticia y su incidencia moratoria en la agencia portal espinar de la caja municipal de ahorro y crédito cusco. Tesis (Contador público). Perú: Universidad Andina del Cusco, 2016. 10pp.

PAMPILLON Fernando, DE LA CUESTA Marta y RUZA Cristina, Introducción al sistema financiero.Madrid: Universidad Nacional de Educación a Distancia Madrid, 2012, 224pp. ISBN: 9788436258240.

PÉREZ, Candiotti. La evaluación crediticia y su relación con el riesgo crediticio, en la Cooperativa de Ahorro y Crédito San Viator Ltda. Tesis (Magíster en Banca y Finanzas). Perú: Universidad mayor de San Marcos, 2013. 75pp.

PINTAR, Matej. DATA WAREHOUSE EN LA SOCIEDAD PARA EL COMERCIO DE ELECTRICIDAD. Tesis (Magister en informática).Ljubliana: Universidad de Ljubljana,2016.

PUERTA, Alejandro. Business Intelligence y la tecnología de la información. 2.a ed. IT campus Academy, 2016. 87pp. ISBN 9781537596020.

RIVADERA, Gustavo. La metodología de Kimball para el diseño de almacenes de datos (Data warehouses).2010. Salta: Universidad Católica de Salta. 58pp.

RIVADERA, Gustavo. La metodología de Kimball para el diseño de almacenes de datos (Data warehouses).2010. Salta: Universidad Católica de Salta. 60pp.

SALVADOR, Molina. El ciclo del crédito. España: LID Editorial, 2015. 33pp. ISBN: 9788483569986.

SALVADOR, Molina. El ciclo del crédito. España: LID Editorial, 2015. 35pp. ISBN: 9788483569986.

SALVADOR, Molina. El ciclo del crédito. España: LID Editorial, 2015. 40pp. ISBN: 9788483569986.

SÁNCHEZ, Víctor. Desarrollo e implementación de un Data mart para el requerimiento de capital por riesgo de crédito del banco InterBank. Tesis (ingeniero de sistemas). Perú: Universidad nacional de Trujillo,2014.

Solís, Jorge. Fepcmac alerta riesgo de sobreendeudamiento por clientes compartidos entre operadores. Gestión. 21 de junio del 2016. [Fecha de consulta: 05 de setiembre del 2018]. Disponible en:

<https://gestion.pe/economia/mercados/fepcmac-alerta-riesgo-sobreendeudamiento-clientes-compartidos-operadores-124775>.

TEMITOPE, Oketunji, Diseño de Data Warehouse y Sistema de Business Intelligence un estudio de caso de una industria minorista. Tesis (Magister de sistemas). Suecia: Universidad de Tecnología de Blekinge,2013.

VALDERRAMA, Santiago. Pasos para elaborar Proyectos de Investigación Científica. Cuantitativa, Cualitativa y Mixta. Lima: San Marcos, 2013. ISBN:978-612-302- 878-7.

VILALTA, Carlos. Análisis de datos. México: CIDE, 2016. ISBN: 978 607 9367 91 6.

## **VIII. ANEXOS**

## Anexo 1: Matriz de Consistencia

Problemas	Objetivos	Hipótesis	Variables	Dimensiones	Indicadores	Metodología
General	General	General	Independiente			<b>Tipo de Investigación:</b> Aplicada <b>Diseño de Investigación:</b> Experimental - Pre-experimental <b>Población:</b> 92 desembolsos <b>Muestra:</b> 75 Desembolsos agrupados en 25 fichas de registro. <b>Método de Investigación:</b> Deductivo <b>Técnica e instrumentos de recolección de datos:</b> Fichaje <b>Instrumento:</b> Ficha de Registro Técnicas para el análisis e interpretación de datos: -Tablas de frecuencias -Diagramas de barras y sectores Unidad de medida: <b>unidades</b>
¿De qué manera influye un Data mart para la evaluación de un crédito en el área de campañas comerciales de la empresa Mibanco?	Desarrollar un Data mart para el análisis de otorgamiento de los créditos comerciales en el área de Desarrollo Comercial de la empresa Mibanco.	El Data mart influye en el análisis de otorgamiento de los créditos comerciales en el área de Desarrollo Comercial de la empresa Mibanco	Data mart			
Específicos	Específicos	Específicos	Dependiente			
<b>P1:</b> ¿De qué manera influye un Data mart en el índice de morosidad en el área de Desarrollo Comercial en la empresa Mibanco?	<b>O1:</b> Determinar la influencia de un Data mart en el índice de morosidad en el área de Desarrollo Comercial de la empresa Mibanco.	<b>H1:</b> El Data mart disminuye el índice de morosidad en el área de Desarrollo Comercial de la empresa Mibanco.	Evaluación de créditos	Seguimiento de riesgo	Índice de morosidad	
<b>P2:</b> ¿De qué manera influye un Data mart en el índice de cobertura en el área de Desarrollo comercial en la empresa Mibanco?	<b>O2:</b> Determinar la influencia de un Data mart en el índice de cobertura en el área de Desarrollo Comercial de la empresa Mibanco.	<b>H2:</b> El Data mart influye en el índice de cobertura en el área de Desarrollo Comercial de la empresa Mibanco.			Índice de cobertura	



## Anexo 2: Ficha técnica, instrumento de recolección de datos

<b>Autor</b>	Callirgos Cardenas Jean Bryant	
<b>Nombre del Instrumento</b>	Ficha de Registro	
<b>Lugar</b>	Mibanco	
<b>Fecha de Aplicación</b>	6 de febrero 2018	
<b>Objetivo</b>	Determinar la influencia de un data mart para la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco	
<b>Tiempo de duración</b>	20 días (lunes a viernes)	
<b>Elección de técnica de instrumento</b>		
<b>Variable</b>	<b>Técnica</b>	<b>Instrumento</b>
<b>Variable Dependiente</b>		
Proceso publicitario	Fichaje	Ficha de registro
<b>Variable Independiente</b>		
Data mart		
Fuente: Elaboración Propia		

### Anexo 3: Instrumento de investigación

Ficha de Registro			
Investigador	Callirgos Cardenas Bryant	Tipo de Prueba	Pre-Test
Empresa Investigada	Mibanco		
Variable	Indicador	Medida	Formula
Evaluación de créditos	índice de Morosidad	Unidad	$\text{Índice de morosidad} = \left( \frac{\text{Crédito vencido}}{\text{crédito total}} \right)$
Consideraciones			
*El rango de tiempo de morosidad es de 6 meses			
*Los desembolsos realizados solo en la provincia de Lima			
*Se considera solo desembolsos de lunes a sábado			

orden	fecha	cantidad	Crédito vencido	Crédito total	Índice de morosidad
1	2/04/2018	3	981054.6	1260000	0.78
2	3/04/2018	3	353477.1	408000	0.87
3	4/04/2018	3	2944467	3783030	0.78
4	5/04/2018	3	1470504.6	1824210	0.81
5	6/04/2018	3	144924.9	207000	0.70
6	7/04/2018	3	239219.7	281880	0.85
7	9/04/2018	3	656774.4	778590	0.84
8	10/04/2018	3	1417462.5	2056500	0.69
9	11/04/2018	3	2533771.2	3255000	0.78
10	12/04/2018	3	303978	390000	0.78
11	13/04/2018	3	1092474.3	1200030	0.91
12	14/04/2018	3	1365833.4	1695000	0.81
13	16/04/2018	3	1258698	1440000	0.87
14	17/04/2018	3	1513867.2	2089740	0.72
15	18/04/2018	3	298824.6	375000	0.80
16	19/04/2018	3	884478	1140000	0.78
17	20/04/2018	3	866710.8	975000	0.89
18	21/04/2018	3	75299.7	108030	0.70
19	23/04/2018	3	3649586.1	4770000	0.77
20	24/04/2018	3	232935.6	300000	0.78
21	25/04/2018	3	726556.5	1050000	0.69
22	26/04/2018	3	545541.9	690000	0.79
23	27/04/2018	3	1941471	2577000	0.75
24	28/04/2018	3	1697957.4	1843200	0.92
25	30/04/2018	3	306784.8	435120	0.71

Total

75



*[Handwritten Signature]*  
44890285

## Instrumento de investigación en el índice de cobertura

Ficha de Registro			
Investigador	Callirgos Cardenas Bryant	Tipo de Prueba	Pre-Test
Empresa Investigada	Mibanco		
Variable	Indicador	Medida	Formula
Evaluación de créditos	índice de Cobertura	Unidad	$\text{Índice de cobertura} = \left( \frac{\text{Provisiones para créditos}}{\text{Cartera atrasada}} \right)$
Consideraciones			
*El rango de tiempo de morosidad es de 6 meses			
*Los desembolsos realizados solo en la provincia de Lima			
*Se considera solo desembolsos de lunes a sábado			

Orden	fecha	cantidad	Cartera atrasada	Provisiones para créditos	Índice de cobertura
1	2/04/2018	3	981054.6	78484.368	0.08
2	3/04/2018	3	353477.1	49486.794	0.14
3	4/04/2018	3	2944467	353336.04	0.12
4	5/04/2018	3	1470504.6	176460.552	0.12
5	6/04/2018	3	144924.9	23187.984	0.16
6	7/04/2018	3	239219.7	35882.955	0.15
7	9/04/2018	3	656774.4	118219.392	0.18
8	10/04/2018	3	1417462.5	170095.5	0.12
9	11/04/2018	3	2533771.2	354727.968	0.14
10	12/04/2018	3	303978	15198.9	0.05
11	13/04/2018	3	1092474.3	142021.659	0.13
12	14/04/2018	3	1365833.4	204875.01	0.15
13	16/04/2018	3	1258698	100695.84	0.08
14	17/04/2018	3	1513867.2	242218.752	0.16
15	18/04/2018	3	298824.6	41835.444	0.14
16	19/04/2018	3	884478	53068.68	0.06
17	20/04/2018	3	866710.8	104005.296	0.12
18	21/04/2018	3	75299.7	9788.961	0.13
19	23/04/2018	3	3649586.1	291966.888	0.08
20	24/04/2018	3	232935.6	18634.848	0.08
21	25/04/2018	3	726556.5	101717.91	0.14
22	26/04/2018	3	545541.9	65465.028	0.12
23	27/04/2018	3	1941471	213561.81	0.11
24	28/04/2018	3	1697957.4	237714.036	0.14
25	30/04/2018	3	306784.8	49085.568	0.16
	total	75			



44890285

Ficha de Registro			
Investigador	Callirgos Cardenas Bryant	Tipo de Prueba	Post-Test
Empresa Investigada	Mibanco		
Variable	Indicador	Medida	Formula
Evaluación de créditos	índice de Morosidad	Unidad	$\text{Índice de morosidad} = \left( \frac{\text{Crédito vencido}}{\text{crédito total}} \right)$
Consideraciones			
*El rango de tiempo de morosidad es de 6 meses			
*Los desembolsos realizados solo en la provincia de Lima			
*Se considera solo desembolsos de lunes a sábado			

Orden	fecha	cantidad	Crédito vencido	Crédito total	Índice de morosidad
1	15/10/2018	3	70801.575	1260000	0.06
2	16/10/2018	3	78865.125	408000	0.19
3	17/10/2018	3	438365.938	3783030	0.12
4	18/10/2018	3	215238.425	1824210	0.12
5	19/10/2018	3	36,231	207000	0.18
6	20/10/2018	3	48,230	281880	0.17
7	22/10/2018	3	150033.825	778590	0.19
8	23/10/2018	3	327105.244	2056500	0.16
9	24/10/2018	3	633442.8	3255000	0.19
10	25/10/2018	3	65204.625	390000	0.17
11	26/10/2018	3	224,513	1200030	0.19
12	27/10/2018	3	300197.525	1695000	0.18
13	29/10/2018	3	114705.806	1440000	0.08
14	30/10/2018	3	378466.8	2089740	0.18
15	31/10/2018	3	61524.7375	375000	0.16
16	1/11/2018	3	221119.5	1140000	0.19
17	2/11/2018	3	120922.888	975000	0.12
18	3/11/2018	3	13627.4375	108030	0.13
19	5/11/2018	3	462159.788	4770000	0.10
20	6/11/2018	3	29959.125	300000	0.10
21	7/11/2018	3	171217.913	1050000	0.16
22	8/11/2018	3	93456.65	690000	0.14
23	9/11/2018	3	337116.863	2577000	0.13
24	10/11/2018	3	120514.163	1843200	0.07
25	12/11/2018	3	46583.7375	435120	0.11
	TOTAL	75			



*[Handwritten signature]*  
44890285

Ficha de Registro			
Investigador	Callirgos Cardenas Bryant	Tipo de Prueba	Post-Test
Empresa Investigada	Mibanco		
Variable	Indicador	Medida	Formula
Evaluación de créditos	índice de Cobertura	Unidad	$\text{Índice de cobertura} = \left( \frac{\text{Provisiones para créditos}}{\text{Cartera atrasada}} \right)$
Consideraciones			
*El rango de tiempo de morosidad es de 6 meses			
*Los desembolsos realizados solo en la provincia de Lima			
*Se considera solo desembolsos de lunes a sábado			

Orden	fecha	cantidad	Cartera atrasada	Provisiones para créditos	Índice de cobertura
1	15/10/2018	3	70801.575	42,480.9	0.6
2	16/10/2018	3	78865.125	60,947.0	0.8
3	17/10/2018	3	438365.9375	290,373.6	0.7
4	18/10/2018	3	215238.425	142,573.9	0.7
5	19/10/2018	3	36231.225	31,999.4	0.9
6	20/10/2018	3	48230.175	39,934.6	0.8
7	22/10/2018	3	150033.825	149,073.6	1.0
8	23/10/2018	3	327105.2438	216,674.5	0.7
9	24/10/2018	3	633442.8	489,524.6	0.8
10	25/10/2018	3	65204.625	46,947.3	0.7
11	26/10/2018	3	224512.5	161,110.2	0.7
12	27/10/2018	3	300197.525	248,563.6	0.8
13	29/10/2018	3	114705.8063	97,499.9	0.8
14	30/10/2018	3	378466.8	334,261.9	0.9
15	31/10/2018	3	61524.7375	47,546.3	0.8
16	1/11/2018	3	221119.5	179,106.8	0.8
17	2/11/2018	3	120922.8875	80,099.3	0.7
18	3/11/2018	3	13627.4375	9,779.0	0.7
19	5/11/2018	3	462159.7875	300,403.9	0.7
20	6/11/2018	3	29959.125	21,271.0	0.7
21	7/11/2018	3	171217.9125	132,317.2	0.8
22	8/11/2018	3	93456.65	61,905.7	0.7
23	9/11/2018	3	337116.8625	204,697.4	0.6
24	10/11/2018	3	120514.1625	93,133.3	0.8
25	12/11/2018	3	46583.7375	41,142.8	0.9
	TOTAL	75			



*[Handwritten signature]*  
44990285

### Anexo 4: Base de datos experimental

ORDEN	ÍNDICE DE MOROSIDAD		ÍNDICE DE COBERTURA	
1	0.78	0.06	0.08	0.6
2	0.87	0.19	0.14	0.8
3	0.78	0.12	0.12	0.7
4	0.81	0.12	0.12	0.7
5	0.70	0.18	0.16	0.9
6	0.85	0.17	0.15	0.8
7	0.84	0.19	0.18	1.0
8	0.69	0.16	0.12	0.7
9	0.78	0.19	0.14	0.8
10	0.78	0.17	0.05	0.7
11	0.91	0.19	0.13	0.7
12	0.81	0.18	0.15	0.8
13	0.87	0.08	0.08	0.8
14	0.72	0.18	0.16	0.9
15	0.80	0.16	0.14	0.8
16	0.78	0.19	0.06	0.8
17	0.89	0.12	0.12	0.7
18	0.70	0.13	0.13	0.7
19	0.77	0.10	0.08	0.7
20	0.78	0.10	0.08	0.7
21	0.69	0.16	0.14	0.8
22	0.79	0.14	0.12	0.7
23	0.75	0.13	0.11	0.6
24	0.92	0.07	0.14	0.8
25	0.71	0.11	0.16	0.9

## Anexo 5: Resultados de la Confiabilidad del Instrumento

Ficha de Registro			
Investigador	Callirgos Cardenas Bryant	Tipo de Prueba	Test
Empresa Investigada	Mibanco		
Variable	Indicador	Medida	Formula
Evaluación de créditos	índice de Morosidad	Unidad	$\text{Índice de morosidad} = \left( \frac{\text{Crédito vencido}}{\text{crédito total}} \right)$
Consideraciones			
*El rango de tiempo de morosidad es de 6 meses			
*Los desembolsos realizados solo en la provincia de Lima			
*Se considera solo desembolsos de lunes a sábado			

Orden	fecha	Crédito vencido	Crédito total	Índice de morosidad
1	1/02/2018	2700138	3584000	0.75
2	2/02/2018	864095.12	1288000	0.67
3	3/02/2018	867279	1610000	0.54
4	5/02/2018	181140.68	193200	0.94
5	6/02/2018	344865.92	560000	0.62
6	7/02/2018	793075.36	868000	0.91
7	8/02/2018	552954.36	1033200	0.54
8	9/02/2018	906577.84	1232000	0.74
9	10/02/2018	197354.36	344288	0.57
10	12/02/2018	573974.24	968800	0.59
11	13/02/2018	995746.92	1279600	0.78
12	14/02/2018	347745.16	378000	0.92
13	15/02/2018	378927.92	688492	0.55
14	16/02/2018	256427.64	336000	0.76
15	17/02/2018	172709.88	217980	0.79
16	19/02/2018	1227767.8	1824424	0.67
17	20/02/2018	329826.56	462000	0.71
18	21/02/2018	2220016.12	2436000	0.91
19	22/02/2018	102942.56	350000	0.29
20	23/02/2018	295184.68	490000	0.60
21	24/02/2018	1093627.36	1765400	0.62
22	26/02/2018	1260707.56	1324876	0.95
23	27/02/2018	2064476.4	3444000	0.60
24	28/02/2018	2601296.04	4564000	0.57
total	24			

Ficha de Registro			
Investigador	Callirgos Cardenas Bryant	Tipo de Prueba	RE Test
Empresa Investigada	Mibanco		
Variable	Indicador	Medida	Formula
Evaluación de créditos	índice de Morosidad	Unidad	$\text{Índice de morosidad} = \left( \frac{\text{Crédito vencido}}{\text{crédito total}} \right)$
Consideraciones			
*El rango de tiempo de morosidad es de 6 meses			
*Los desembolsos realizados solo en la provincia de Lima			
*Se considera solo desembolsos de lunes a sábado			

Orden	fecha	Crédito vencido	Crédito total
1/03/2018	2989438.5	3968000	0.75
2/03/2018	1743581.05	3782000	0.46
3/03/2018	937462.94	1782500	0.53
5/03/2018	200548.61	213900	0.94
6/03/2018	430163.13	711543	0.60
7/03/2018	1801431.39	2356000	0.76
8/03/2018	746140.24	1298900	0.57
9/03/2018	100913.99	124000	0.81
10/03/2018	217414.16	381176	0.57
12/03/2018	635471.48	1072600	0.59
13/03/2018	1157529.77	1630600	0.71
14/03/2018	1359828.95	1410500	0.96
15/03/2018	397108.45	762259	0.52
16/03/2018	283902.03	372000	0.76
17/03/2018	191214.51	241335	0.79
19/03/2018	1342461.82	2019898	0.66
20/03/2018	353909.64	511500	0.69
21/03/2018	2457874.99	2697000	0.91
22/03/2018	13092.23	18600	0.70
23/03/2018	326811.61	542500	0.60
24/03/2018	1181856.4	1954550	0.60
26/03/2018	1390190.35	1466827	0.95
27/03/2018	2263482.98	3813000	0.59
28/03/2018	2181267.57	2883000	0.76
total	24		



*[Handwritten signature]*  
44890285



## Correlaciones

Correlaciones

		TEST_INDICE_MOROSIDAD	RE_TEST_INDICE_MOROSIDAD
TEST_INDICE_MOROSIDAD	Correlación de Pearson	1	,741**
	Sig. (bilateral)		0.000
	N	24	24
RE_TEST_INDICE_MOROSIDAD	Correlación de Pearson	,741**	1
	Sig. (bilateral)	0.000	
	N	24	24

\*\* . La correlación es significativa en el nivel 0,01 (bilateral).

Ficha de Registro			
Investigador	Callirgos Cardenas Bryant	Tipo de Prueba	Test
Empresa Investigada	Mibanco		
Variable	Indicador	Medida	Formula
Evaluación de créditos	índice de Cobertura	Unidad	$\text{Índice de cobertura} = \left( \frac{\text{Provisiones para créditos}}{\text{Cartera atrasada}} \right)$
Consideraciones			
*El rango de tiempo de morosidad es de 6 meses			
*Los desembolsos realizados solo en la provincia de Lima			
*Se considera solo desembolsos de lunes a sábado			

Orden	fecha	Cartera atrasada	Provisiones para créditos	Índice de cobertura
1	1/02/2018	2,700,138	307,734	0.11
2	2/02/2018	864,095	307,734	0.36
3	3/02/2018	867279	307734	0.35
4	5/02/2018	181140.68	307734	1.70
5	6/02/2018	344865.92	307734	0.89
6	7/02/2018	793075.36	307734	0.39
7	8/02/2018	552954.36	307734	0.56
8	9/02/2018	906577.84	307734	0.34
9	10/02/2018	197354.36	307734	1.56
10	12/02/2018	573974.24	307734	0.54
11	13/02/2018	995746.92	307734	0.31
12	14/02/2018	347745.16	307734	0.88
13	15/02/2018	378927.92	307734	0.81
14	16/02/2018	256427.64	307734	1.20
15	17/02/2018	172709.88	307734	1.78
16	19/02/2018	1227767.8	307734	0.25
17	20/02/2018	329826.56	307734	0.93
18	21/02/2018	2,220,016	307,734	0.14
19	22/02/2018	102,943	307,734	2.99
20	23/02/2018	295184.68	307734	1.04
21	24/02/2018	1093627.36	307734	0.28
22	26/02/2018	1260707.56	307734	0.24
23	27/02/2018	2064476.4	307734	0.15
24	28/02/2018	2601296.04	307734	0.12
	total	24		

  
44890285



Ficha de Registro			
Investigador	Callirgos Cardenas Bryant	Tipo de Prueba	Re Test
Empresa Investigada	Mibanco		
Variable	Indicador	Medida	Formula
Evaluación de créditos	índice de Cobertura	Unidad	$\text{Índice de cobertura} = \left( \frac{\text{Provisiones para créditos}}{\text{Cartera atrasada}} \right)$
Consideraciones			
*El rango de tiempo de morosidad es de 6 meses			
*Los desembolsos realizados solo en la provincia de Lima			
*Se considera solo desembolsos de lunes a sábado			

Orden	fecha	Cartera atrasada	Provisiones para créditos	Índice de cobertura
1	1/03/2018	2989438.5	307734	0.10
2	2/03/2018	1743581.05	512890	0.29
3	3/03/2018	937462.94	307734	0.33
4	5/03/2018	200548.61	307734	1.53
5	6/03/2018	430163.13	410312	0.95
6	7/03/2018	1801431.39	512890	0.28
7	8/03/2018	746140.24	410312	0.55
8	9/03/2018	100913.99	205156	2.03
9	10/03/2018	217414.16	307734	1.42
10	12/03/2018	635471.48	307734	0.48
11	13/03/2018	1157529.77	410312	0.35
12	14/03/2018	1359828.95	410312	0.30
13	15/03/2018	397108.45	307734	0.77
14	16/03/2018	283902.03	307734	1.08
15	17/03/2018	191214.51	307734	1.61
16	19/03/2018	1342461.82	307734	0.23
17	20/03/2018	353909.64	307734	0.87
18	21/03/2018	2457874.99	307734	0.13
19	22/03/2018	13092.23	102578	7.84
20	23/03/2018	326811.61	307734	0.94
21	24/03/2018	1181856.4	307734	0.26
22	26/03/2018	1390190.35	307734	0.22
23	27/03/2018	2263482.98	307734	0.14
24	28/03/2018	2181267.57	205156	0.09
	Total		24	



*[Handwritten signature]*  
44890285

## Correlaciones

Correlaciones

		IND_COBER_TEST	IND_COBER_RE_TEST
IND_COBER_TEST	Correlación de Pearson	1	,854**
	Sig. (bilateral)		0.000
	N	24	24
IND_COBER_RE_TEST	Correlación de Pearson	,854**	1
	Sig. (bilateral)	0.000	
	N	24	24

\*\* . La correlación es significativa en el nivel 0,01 (bilateral).

## Anexo N°6: Evaluación del instrumento

TABLA DE EVALUACION DE EXPERTOS

Apellidos y nombres del experto: Cueva Villavicencio Juanita Isabel

Título y/o Grado:

Ph.D ( )	Doctor ( )	Magister ( <input checked="" type="checkbox"/> )	Licenciado ( )	Otros ( ) Especifique: _____
----------	------------	--	----------------	------------------------------

Universidad que labora: Universidad Cesar Vallejo

Fecha: 19-10-17

**TESIS**

Data Mart para el proceso crediticio en el área de Campañas Comerciales de la empresa  
Mibanco

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar las metodologías involucradas, mediante una serie de preguntas con puntuaciones especificadas al final de la tabla. Asimismo, le exhortamos en la corrección de los ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de las preguntas.

ITEMS	PREGUNTAS	METODOLOGIA			
		RALPH KIMBALL	HEFESTO	INMON	OBSERVACIONES
1	Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.	5	3	1	
2	Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.	5	5	1	
3	La metodología involucra al usuario durante las etapas del proyecto.	3	5	3	
4	La metodología después de aplicarse es de fácil mantenimiento.	5	5	1	
5	Tiene etapas completamente detalladas, que sumándolas dan como resultado la completitud del proyecto de BI.	5	5	1	
6	Utiliza una mejor adaptabilidad de desarrollo en las organizaciones.	5	3	1	
7	Genera menor tiempo en consultas	5	5	3	
TOTAL		33	31	11	

Evaluar con la siguiente puntuación:

1: Malo – 3: Regular – 5: Bueno

SUGERENCIAS

Firma del experto:


  
 \_\_\_\_\_

TABLA DE EVALUACION DE EXPERTOS

Apellidos y nombres del experto: Flores Masías, Edward José

Título y/o Grado:

Ph.D ( )	Doctor (X)	Magister ( )	Licenciado ( )	Otros ( ) Especifique:
----------	------------	--------------	----------------	------------------------

Universidad que labora: Universidad Cesar Vallejo

Fecha: 19-10-17

TESIS

Data Mart para el proceso crediticio en el área de Campañas Comerciales de la empresa Mibanco

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar las metodologías involucradas, mediante una serie de preguntas con puntuaciones especificadas al final de la tabla. Asimismo, le exhortamos en la corrección de los ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de las preguntas.

ITEMS	PREGUNTAS	METODOLOGÍA			
		RALPH KIMBALL	HEFESTO	INMON	OBSERVACIONES
1	Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.	5	3	1	
2	Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.	5	3	1	
3	La metodología involucra al usuario durante las etapas del proyecto.	5	1	3	
4	La metodología después de aplicarse es de fácil mantenimiento.	5	3	1	
5	Tiene etapas completamente detalladas, que sumándolas dan como resultado la completitud del proyecto de BI.	5	1	3	
6	Utiliza una mejor adaptabilidad de desarrollo en las organizaciones.	5	3	1	
7	Genera menor tiempo en consultas	5	1	3	
TOTAL		35	15	13	

Evaluar con la siguiente puntuación:

1: Malo - 3: Regular - 5: Bueno

SUGERENCIAS

.....  
 .....

Firma del experto:

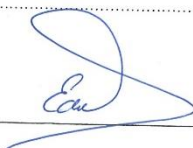


TABLA DE EVALUACION DE EXPERTOS

Apellidos y nombres del experto: Santana, Juan R.

Título y/o Grado: Magister

Ph.D ( )	Doctor ( )	Magister (x)	Licenciado ( )	Otros ( ) Especifique:
----------	------------	--------------	----------------	------------------------

Universidad que labora: Universidad Cesar Vallejo

Fecha: 16-10-17

TESIS

Data Mart para el proceso crediticio en el área de Campañas Comerciales de la empresa Mibanco

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar las metodologías involucradas, mediante una serie de preguntas con puntuaciones especificadas al final de la tabla. Asimismo, le exhortamos en la corrección de los ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de las preguntas.

ITEMS	PREGUNTAS	METODOLOGIA			
		RALPH KIMBALL	HEFESTO	INMON	OBSERVACIONES
1	Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.	5	3	1	
2	Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.	5	3	1	
3	La metodología involucra al usuario durante las etapas del proyecto.	5	3	1	
4	La metodología después de aplicarse es de fácil mantenimiento.	5	3	1	
5	Tiene etapas completamente detalladas, que sumándolas dan como resultado la completitud del proyecto de BI.	5	3	1	
6	Utiliza una mejor adaptabilidad de desarrollo en las organizaciones.	5	3	1	
7	Genera menor tiempo en consultas	5	3	1	
	TOTAL	35	21	7	

Evaluar con la siguiente puntuación:

1: Malo – 3: Regular – 5: Bueno

SUGERENCIAS

Firma del experto:



### VALIDACIÓN DE INSTRUMENTO

Apellidos y nombres del experto: *Gálvez Tapia Orleans*

Doctor ( )	Magister ( <input checked="" type="checkbox"/> )	Ingeniero ( )	Otros... especifique
------------	--	---------------	----------------------

Título y/o Grado:

Universidad que labora: Universidad Cesar vallejo

Fecha:

Título de Tesis: Data mart para la evaluación de créditos en el área de Campañas Comerciales de la empresa Mibanco

Tabla de evaluación de expertos para el indicador: Índice de morosidad

Ítems	Preguntas	Deficiente 0% - 19%	Regular 20% - 39%	Bueno 40% - 60%	Muy Bueno 61% - 80%	Excelente 81% - 100%
1	¿El instrumento es adecuado al tipo de investigación?				80	
2	¿Facilitara el estado de la variable dependiente				80	
3	¿el instrumento de recolección de datos facilitara el logro de los objetivos de la investigación?				80	
4	¿Es confiable, es decir, debe medir lo mismo en diferentes mediciones?				80	
5	¿Esta basado en aspectos teóricos científicos?				80	
6	¿El diseño del instrumento de recolección de datos facilita el análisis y procesamiento de datos?				80	
7	¿El instrumento de recolección de datos facilita el logro de los objetivos de la investigación?				80	
8	¿El instrumento de recolección de datos tiene relación con el título de la investigación?				80	
Promedio						

FIRMA DEL EXPERTO





### VALIDACIÓN DE INSTRUMENTO

Apellidos y nombres del experto: *Galvez Tapra Orlean*

Doctor ( )	Magister ( <input checked="" type="checkbox"/> )	Ingeniero ( )	Otros... especifique
------------	--	---------------	----------------------

Título y/o Grado:

Universidad que labora: Universidad Cesar vallejo

Fecha:

Título de Tesis: *Data mart para la evaluación de créditos en el área de Campañas Comerciales de la empresa Mibanco*

Tabla de evaluación de expertos para el indicador: Índice de cobertura

Ítems	Preguntas	Deficiente 0% - 19%	Regular 20% - 39%	Bueno 40% - 60%	Muy Bueno 61% - 80%	Excelente 81% - 100%
1	¿El instrumento es adecuado al tipo de investigación?				80	
2	¿Facilitara el estado de la variable dependiente				80	
3	¿el instrumento de recolección de datos facilitara el logro de los objetivos de la investigación?				80	
4	¿Es confiable, es decir, debe medir lo mismo en diferentes mediciones?				80	
5	¿Esta basado en aspectos teóricos científicos?				80	
6	¿El diseño del instrumento de recolección de datos facilita el análisis y procesamiento de datos?				80	
7	¿El instrumento de recolección de datos facilita el logro de los objetivos de la investigación?				80	
8	¿El instrumento de recolección de datos tiene relación con el título de la investigación?				80	
Promedio						

FIRMA DEL EXPERTO



**VALIDACIÓN DE INSTRUMENTO**

Apellidos y nombres del experto: *Chumpe Agosto, S.- B.*

Doctor ( )	Magister (x)	Ingeniero ( )	Otros... especifique
------------	--------------	---------------	----------------------

Título y/o Grado:

Universidad que labora: Universidad Cesar vallejo

Fecha: *06/06/2018*

Título de Tesis: Data mart para la evaluación de créditos en el área de Campañas Comerciales de la empresa Mibanco

Tabla de evaluación de expertos para el indicador: Índice de morosidad

Ítems	Preguntas	Deficiente 0% - 19%	Regular 20% - 39%	Bueno 40% - 60%	Muy Bueno 61% - 80%	Excelente 81% - 100%
1	¿El instrumento es adecuado al tipo de investigación?				71	
2	¿Facilitara el estado de la variable dependiente				71	
3	¿el instrumento de recolección de datos facilitara el logro de los objetivos de la investigación?				71	
4	¿Es confiable, es decir, debe medir lo mismo en diferentes mediciones?				71	
5	¿Esta basado en aspectos teóricos científicos?				71	
6	¿El diseño del instrumento de recolección de datos facilita el análisis y procesamiento de datos?				71	
7	¿El instrumento de recolección de datos facilita el logro de los objetivos de la investigación?				71	
8	¿El instrumento de recolección de datos tiene relación con el título de la investigación?				71	
<b>Promedio</b>					71	

*[Handwritten Signature]*  
**FIRMA DEL EXPERTO**  
*06/06/2018*



**VALIDACIÓN DE INSTRUMENTO**

Apellidos y nombres del experto:

*Chumpu Aguirre J. G.*

Doctor ( )	Magister ( <input checked="" type="checkbox"/> )	Ingeniero ( )	Otros... especifique
------------	--	---------------	----------------------

Título y/o Grado:

Universidad que labora: Universidad Cesar vallejo

Fecha: *06/06/2018*

Título de Tesis: Data mart para la evaluación de créditos en el área de Campañas Comerciales de la empresa Mibanco

Tabla de evaluación de expertos para el indicador: Índice de cobertura

Ítems	Preguntas	Deficiente 0% - 19%	Regular 20% - 39%	Bueno 40% - 60%	Muy Bueno 61% - 80%	Excelente 81% - 100%
1	¿El instrumento es adecuado al tipo de investigación?				71	
2	¿Facilitara el estado de la variable dependiente				71	
3	¿el instrumento de recolección de datos facilitara el logro de los objetivos de la investigación?				71	
4	¿Es confiable, es decir, debe medir lo mismo en diferentes mediciones?				71	
5	¿Esta basado en aspectos teóricos científicos?				72	
6	¿El diseño del instrumento de recolección de datos facilita el análisis y procesamiento de datos?				72	
7	¿El instrumento de recolección de datos facilita el logro de los objetivos de la investigación?				72	
8	¿El instrumento de recolección de datos tiene relación con el título de la investigación?				72	
<b>Promedio</b>					71.5	

*[Handwritten Signature]*  
**FIRMA DEL EXPERTO**  
*06/11/2018*



### VALIDACIÓN DE INSTRUMENTO

Apellidos y nombres del experto: *ORDÓÑEZ PÉREZ, DAVIDO CARISMAN*

Doctor ( <input checked="" type="checkbox"/> )	Magister ( <input type="checkbox"/> )	Ingeniero ( <input type="checkbox"/> )	Otros... especifique
--	---------------------------------------	--	----------------------

Título y/o Grado:

Universidad que labora: Universidad Cesar vallejo

Fecha:

Título de Tesis: *Data mart para la evaluación de créditos en el área de Campañas Comerciales de la empresa Mibanco*

Tabla de evaluación de expertos para el indicador: Índice de morosidad

Ítems	Preguntas	Deficiente 0% - 19%	Regular 20% - 39%	Bueno 40% - 60%	Muy Bueno 61% - 80%	Excelente 81% - 100%
1	¿El instrumento es adecuado al tipo de investigación?					90
2	¿Facilitara el estado de la variable dependiente					90
3	¿el instrumento de recolección de datos facilitara el logro de los objetivos de la investigación?					95
4	¿Es confiable, es decir, debe medir lo mismo en diferentes mediciones?					95
5	¿Esta basado en aspectos teóricos científicos?				80	
6	¿El diseño del instrumento de recolección de datos facilita el análisis y procesamiento de datos?					90
7	¿El instrumento de recolección de datos facilita el logro de los objetivos de la investigación?					95
8	¿El instrumento de recolección de datos tiene relación con el título de la investigación?					90
<b>Promedio</b>						91

FIRMA DEL EXPERTO

### VALIDACIÓN DE INSTRUMENTO

Apellidos y nombres del experto: *ORPONER PEREZ, POULIO CARRASCO*

Doctor ( <input checked="" type="checkbox"/> )	Magister ( <input type="checkbox"/> )	Ingeniero ( <input type="checkbox"/> )	Otros... especifique
--	---------------------------------------	--	----------------------

Título y/o Grado:

Universidad que labora: Universidad Cesar vallejo

**Fecha:**

Título de Tesis: Data mart para la evaluación de créditos en el área de Campañas Comerciales de la empresa Mibanco

Tabla de evaluación de expertos para el indicador: Índice de cobertura

Ítems	Preguntas	Deficiente 0% - 19%	Regular 20% - 39%	Bueno 40% - 60%	Muy Bueno 61% - 80%	Excelente 81% - 100%
1	¿El instrumento es adecuado al tipo de investigación?					90
2	¿Facilitara el estado de la variable dependiente					90
3	¿el instrumento de recolección de datos facilitara el logro de los objetivos de la investigación?					95
4	¿Es confiable, es decir, debe medir lo mismo en diferentes mediciones?					95
5	¿Esta basado en aspectos teóricos científicos?				80	
6	¿El diseño del instrumento de recolección de datos facilita el análisis y procesamiento de datos?					90
7	¿El instrumento de recolección de datos facilita el logro de los objetivos de la investigación?					95
8	¿El instrumento de recolección de datos tiene relación con el título de la investigación?					90
<b>Promedio</b>						91



FIRMA DEL EXPERTO

## Anexo 7: Entrevista

Anexo N°1: Entrevista al analista encargado de procesar las Campañas Comerciales

Nombre del entrevistado	Jovan Reyes Ubilluz
Cargo	Analista Senior de Campañas Comerciales
Fecha	25-09-2017

1. ¿Cuál es el nombre comercial y la razón social de la empresa?

El nombre comercial es Mibanco y la razón social es MIBANCO - BANCO DE LA MICROEMPRESA S.A.

2. ¿Dónde se encuentra ubicada la empresa?

Mi banco se encuentra en la Calle. Domingo Orué Nro. 165 Res.

3. ¿Cuál es el sector al que pertenece la empresa?

Mibanco pertenece al sector de la micro y pequeña empresa

4. ¿A qué se dedica el área de campañas comerciales?

Se encarga del desarrollo e implementación de las campañas internas para la red y la creación y mantenimiento de la base de datos de clientes, en resumen se encarga de seleccionar a los mejores clientes del banco y ofrecerle facilidades para desembolsar, con productos adecuados a su perfil.

5. ¿Cuál es el proceso principal en el área de campañas comerciales?

El proceso principal es El proceso de Desarrollo e implementación de la base de campañas comerciales internas que nos sirve para la generación del perfil de todos los clientes el cual esta subdividido en 2 subprocesos:

SUBPROCESO ANÁLISIS DE RESULTADO DE CAMPAÑAS HISTÓRICAS

SUBPROCESO GENERACIÓN Y CARGA DE BASE DE CAMPAÑAS (PRE-APROBADOS Y APROBADOS)

6. ¿Considera que existe algún problema en ese proceso o alguno ligado directamente?

Sí, presenta algunos problemas como por ejemplo al momento de consultar la data histórica de Campañas se hace muy pesado, ya que es mucha información y tarda mucho si queremos consultar algo de una tabla, otro inconveniente que tenemos es que en algunas tablas de donde extraemos información de los clientes contiene data duplicada.



Un problema ligado al proceso es el de generar los reportes, se toma mucho tiempo en generarlos

Otro problema ligado al proceso es al momento de actualizar los Excel que contienen un análisis de cosechas, que nos sirve para ver cómo van los clientes en Campañas, el cual nos toma 2 semanas en actualizarlos, ya que esto lo actualizamos a cada inicio de Campaña.

7. ¿Qué personas y áreas intervienen en ese proceso?

Al momento de comenzar a procesar el análisis de resultado de las campañas históricas solo lo ejecuto yo y con ayuda de un practicante que me apoya y si hay algún cambio muy significativo en la base, interviene el equipo de Riesgos.

8. De solicitar documentación, ¿se autoriza la emisión de los mismos?

La emisión de documentos requeridos con fines educativos para la presente investigación



---

FIRMA DEL ENTREVISTADO

**Ficha de Entrevista**

Entrevistador	Bryant Callirgos Cardenas
Fecha	01/04/2018
Lugar	Mibanco
Objetivo	Listar los indicadores utilizados por el área de campañas comerciales de la empresa Mibanco

Entrevistado(a)	Jovan Reyes
Cargo	Analista Senior de campañas

<p>1.¿Podría listar los indicadores que se utilizan para la toma de decisiones en el área de campañas</p> <ul style="list-style-type: none"> <li>✓ Cantidad total de clientes en Mora en un determinado tiempo.</li> <li>✓ Monto total de mora de cada producto a cada cliente en un determinado tiempo.</li> <li>✓ Mejores clientes por zona geográfica, por producto, por mes y año.</li> <li>✓ Cantidad total de dinero perdido por los clientes que cayeron en mora por departamento, provincia y distrito y en un determinado tiempo.</li> <li>✓ Patrón de los clientes que cayeron en mora en un determinado tiempo.</li> <li>✓ Agencias son los que tienen más cantidad de clientes morosos en un determinado tiempo.</li> <li>✓ Ciudades que tienen más clientes morosos en un determinado tiempo.</li> <li>✓ Ganancia anual y mensual de los desembolsos de créditos en un determinado tiempo.</li> <li>✓ Agencia que más desembolsa un producto de campaña y que tipo de clientes son los que más usan esos productos.</li> <li>✓ Índice de morosidad por ciudad en un determinado tiempo.</li> <li>✓ Índice de cobertura por ciudad en un determinado tiempo.</li> <li>✓ Productos más desembolsados y cuanto es el monto total desembolsado.</li> <li>✓ Tipo de asesores de negocios son los más eficientes por meses y años.</li> <li>✓ Asesores son los menos eficientes en un determinado tiempo.</li> <li>✓ Mejores clientes por zona geográfica, por producto, por mes y año.</li> <li>✓ Efectividad de la base de campañas en un determinado tiempo.</li> <li>✓ Cantidad promedio de clientes por tipo de asesores de negocios.</li> <li>✓ Desembolsos por cada tipo de campaña en un determinado tiempo.</li> <li>✓ Ganancia anual y mensual de los desembolsos.</li> <li>✓ Desembolsos realizados por zonas, región y territorio de forma mensual y anual.</li> <li>✓ Cantidad de desembolsos por tipo de clientes y por tipo de productos.</li> <li>✓ Número total de entidades con la que trabaja cada cliente en un determinado tiempo.</li> <li>✓ Deuda total que maneja cada cliente con todas las entidades financieras en un determinado tiempo.</li> </ul>
--

 01/04/2018  
Firma del entrevistado





## Anexo 8: Carta de aprobación de la empresa

CONSTANCIA

JOVAN REYES UBILLUS ANALISTA SENIOR DE LA EMPRESA DE  
MIBANCO BANCO DE LA MICROEMPRESA S.A. QUE SUSCRIBE:


HACE CONSTAR:

Que el estudiante Callirgos Cardenas Jean Bryant de la Escuela de Ingeniería de sistemas de la Universidad Cesar Vallejo, realizo la implementación de un Data mart para la evaluación de créditos en el área de campañas comerciales de la empresa Mibanco teniendo así mejores resultados en la disminución de la morosidad.

Dicha investigación será desarrollada durante el semestre Académico 2018-II, del décimo ciclo académico de la Facultad de Ingeniería.

Se expide la presente constancia para los fines pertinentes.

Lima, 30 de octubre del 2018

  
..... 44890285

Firma



## **Anexo 9 Desarrollo de la metodología**



**UNIVERSIDAD CÉSAR VALLEJO**

**FACULTAD DE INGENIERÍA**

**ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**DATA MART PARA EL PROCESO CREDITICIO EN EL ÁREA DE  
CAMPAÑAS COMERCIALES DE LA EMPRESA MIBANCO**

**TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:  
INGENIERO DE SISTEMAS**

**AUTOR:**

**BRYANT CALLIRGOS CARDENAS**

**ASESOR:**

**DR. ADILIO CHRISTIAN ORDOÑEZ PÉREZ**

**LÍNEA DE INVESTIGACIÓN:**

**SISTEMAS DE INFORMACIÓN Y COMUNICACIONES**

**LIMA - PERÚ**

**(2018)**

## **PRESENTACIÓN**

La presente tesis consiste en la implementación de un Data mart para la evaluación de los créditos en el área de campañas comerciales de la empresa Mibanco.

Mibanco es una entidad privada que se desempeña en el sector financiero a nivel nacional, que tiene como misión transformar las vidas de sus clientes y colaboradores a través de la inclusión financiera, impulsando así el crecimiento del Perú.

El desarrollo de esta tesis se realizará con la metodología Ralph Kimball, ya que proporciona un enfoque de menor a mayor, muy versátil y una serie de herramientas prácticas que ayudan a implementar un Data mart, con pocos recursos.

## ÍNDICE GENERAL

	Página
PRESENTACIÓN	ii
Índice de Figuras	iv
Índice de tablas	vi
1. Planificación del proyecto (Plan)	7
2. Definición de requerimientos del negocio	13
3. Modelado dimensional	17
4. Diseño Físico	28
5. Diseño e implementación del subsistema de ETL	34
6. Especificación de aplicaciones BI	40
7. Desarrollo de Aplicaciones BI	41
8. Implementación	44

## Índice de Figuras

	Página
Figura 1 Planificación del proyecto	10
Figura 2 Diagrama de arquitectura funcional del data mart	15
Figura 3 Diagrama de arquitectura de alto nivel del data mart	16
Figura 4 FACT_DESEMBOLSOS	20
Figura 5 FACT_MORA	21
Figura 6 Granularidad de la dimensión tiempo	22
Figura 7 Campos de la dimensión DM_TIEMPO	22
Figura 8 Nivel de granularidad de la dimensión producto	23
Figura 9 Campos de la dimensión producto	23
Figura 10 Campos de la dimensión Cliente	24
Figura 11 Campos de la dimensión analista	24
Figura 12 Nivel de granularidad de la dimensión agencia	25
Figura 13 Campos de la dimensión Agencia	25
Figura 14 Nivel de granularidad de la dimensión ubigeo	26
Figura 15 Campos de la dimensión ubigeo	26
Figura 16 Campos de la dimensión campañas	27
Figura 17 Diseño Lógico de la base de datos del Data mart	28
Figura 18 Diseño Físico de la base de datos del data mart	29
Figura 19 Creación del procedimiento para actualizar la dimensión producto	35
Figura 20 Creación del procedimiento para actualizar la dimensión Agencia	35
Figura 21 Creación del procedimiento para actualizar la dimensión Campañas	36
Figura 22 Creación del procedimiento para actualizar la dimensión Analista	36
Figura 23 Creación del procedimiento para actualizar la dimensión Clientes	37
Figura 24 Creación del procedimiento para actualizar la dimensión Ubigeo	37
Figura 25 Creación del procedimiento para actualizar la dimensión Fecha	38
Figura 26 Creación del procedimiento para actualizar la tabla hecho Fact_Campañas	39
Figura 27 Creación del procedimiento FACT_MORA	39
Figura 28 Creación del JOB de todos los procedimientos credos	40

Figura 29 Desembolso por año	42
Figura 30 desembolso por territorio	43
Figura 31 desembolso por Campaña	43
Figura 32 Inicio de sesion al servidor Reporting Services	44
Figura 33 Datos del servidor Reporting Services	45
Figura 34 Directorio virtual donde se publicaran los reportes	45
Figura 35 Directorio virtual donde estará publicado los reportes	46
Figura 36 Directorio virtual donde se almacenarán los reportes y donde se crearán las conexiones	46
Figura 37 Configuración del sitio	47
Figura 38 Creación del origen de datos	47
Figura 39 Conexion del data mart	48
Figura 40 Generador de informes	48
Figura 41 Loggeo al sistema de Mibanco	49
Figura 42 Reportes en línea	49
Figura 43 Creación de nuevos reportes	50

## Índice de tablas

	Página
Tabla 1 Recurso humano.....	7
Tabla 2 Materiales de bienes de consumo.....	8
Tabla 3 Materiales de bienes de inversión.....	8
Tabla 4 Presupuesto y financiamiento.....	9
Tabla 5 Beneficios de implementar un Data mart.....	12
Tabla 6 Descripción del diagrama de arquitectura.....	17
Tabla 7 Diccionario de datos.....	30
Tabla 8 Diccionario de datos.....	31
Tabla 9 Diccionario de datos.....	32
Tabla 10 Diccionario de datos.....	33
Tabla 11 Relación entre la información y el sistema transaccional.....	34
Tabla 12 Herramientas de apoyo para explotar la información del Data mart.....	41



## 1. Planificación del proyecto (Plan)

### Objetivo

El objetivo principal para la implementación del data mart es mejorar la evaluación de créditos en el área de campañas comerciales en la empresa Mibanco, a través de análisis de información históricos para tomar decisiones y así poder reducir la morosidad que es uno de los problemas más críticos que posee Mibanco.

### Definición del equipo de trabajo

Para elaborar el desarrollo de esta investigación se define el equipo de trabajo, el cual consta únicamente por el responsable de realizar la investigación, y las funciones serán realizadas por el mismo responsable.

### Recursos humanos

Los recursos humanos para el proyecto se encuentran en la tabla N°1, el recurso humano es el investigador, se clasifico en 2 categorías, la primera en Analista de Business Intelligence y la segunda en Analista programador, ya que son 2 funciones distintas que se realizaron con diferente presupuesto.

**Tabla 1 Recurso humano**

Personal	Cantidad	Duración del proyecto	Pago Mensual (S/.)	Pago por tiempo de duración del proyecto (S/.)
Analista de BI	1	2 meses	2,000	4,000
Analista programador	1	2 meses	1,000	2,000
<b>Total</b>				6,000

Fuente: Mibanco

## Materiales

Los bienes materiales para el proyecto se encuentran en la siguiente tabla N 2

**Tabla 2 Materiales de bienes de consumo**

Ítems	Descripción	Cantidad	Precio	total
1	Hojas bond A4 (1 ciento)	2	10	20
2	Lapiceros	5	1	5
3	Impresiones	300	0.25	75
4	Folder	10	0.80	8
5	CD	15	1	15
<b>Total</b>				123

Fuente: Mibanco

Los bienes de inversión para el proyecto se encuentran en la siguiente tabla N 3, en el desarrollo del proyecto de investigación se utilizó una computadora por 8 meses, el tiempo que duro la investigación y la implementación, por lo cual se pondrá como gasto la depreciación del equipo, la depreciación de equipos de procesamiento de datos es de 25% anual.

**Tabla 3 Materiales de bienes de inversión**

Ítems	Equipos	Cantidad	Precio	Precio actual	Total
<b>Hardware</b>					
1	Computador	1	2,500	2,083.33	416.67
<b>Software</b>					
1	SQL Server Data tools for visual studio 2013	1	7,175	7,175	7,175
<b>Total</b>					7,591.67

Fuente: Mibanco

## **Presupuesto**

El presupuesto y financiamiento para el proyecto se muestran en la tabla N 4, donde el financiamiento del proyecto será en un 90% por la empresa y un 10% por el tesista cumpliendo así el 100% del gasto que equivale a los 13,71467 soles de inversión.

**Tabla 4 Presupuesto y financiamiento**

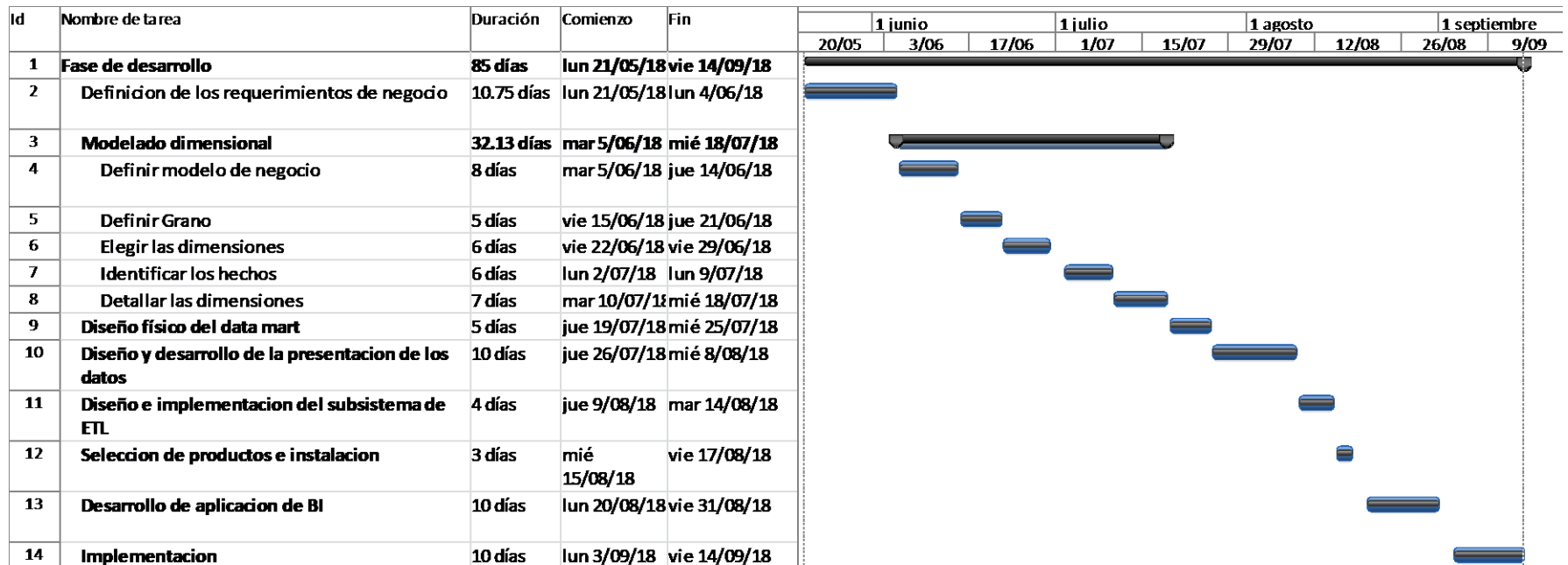
<b>Descripción</b>	<b>Monto</b>
<b>Recursos humanos</b>	6,000
<b>Recursos materiales</b>	
<b>Bienes de consumo</b>	123
<b>Bienes de inversión</b>	7,591.67
<b>Total</b>	13,714.67

Fuente: Mibanco

## Definición del equipo de trabajo

Como en todo proyecto de software, es necesario realizar una planificación de este para tratar de garantizar que el desarrollo se realiza según las fechas previstas. A continuación, se presenta la planificación prevista para el proyecto de Data mart que se está realizando:

**Figura 1 Planificación del proyecto**



Fuente: elaboración propia

## **Alcances del Proyecto**

El proyecto busca ayudar al área de Campañas Comerciales de la empresa Mibanco a través de reportes analíticos con información histórica, veraz y consolidada en tiempo real, permitiendo sacar mejores perfilamientos de clientes, excluir el perfilamiento de los clientes que caen en mora, fidelizar a los mejores clientes y realizar reportes de manera automatizada, además de permitir obtener y analizar los indicadores en tiempo real, oportuno, certero, veraz y confiable.

## **Tareas identificadas**

Para el desarrollo del Data mart, se decidió ejecutar las siguientes tareas:

- Identificar las fuentes de datos para la extracción de la información necesaria.
- Identificar las dimensiones y la tabla hechos para la creación del modelo lógico y físico del Data mart.
- Creación del modelo dimensional y el ETL mediante querys.
- Creación del cubo multidimensional en la herramienta Data tools for Visual Studio.
- Elaboración de reportes y análisis de información en la herramienta de Reporting Services.

## Beneficios

Beneficios de la solución de un Data mart

**Tabla 5 Beneficios de implementar un Data mart**

	Beneficios
Humano	<ul style="list-style-type: none"><li>• Equipo de trabajo efectivo, capaz de tomar las mejores decisiones.</li><li>• Enriquecer el proceso de análisis de información por parte de los analistas Senior.</li></ul>
Tecnológico	<ul style="list-style-type: none"><li>• Herramientas de información acorde con las necesidades de la empresa.</li><li>• Tecnología informática amigable con los usuarios</li></ul>
Material	<ul style="list-style-type: none"><li>• Reducción de tiempos en la emisión de reportes consolidados.</li><li>• Reducción de costos en el proceso de análisis de información.</li></ul>
Información	<ul style="list-style-type: none"><li>• Información veraz y oportuna.</li><li>• Análisis de información adecuadas basadas en reportes dinámicos.</li><li>• Proceso analítico de la data de la institución</li></ul>

Fuente: Elaboración propia

## 2. Definición de requerimientos del negocio

### Requerimientos del negocio

Los requerimientos de negocio son un conjunto de necesidades que ha identificado en el área de campañas comerciales, realizadas al analista senior de campañas de la empresa Mibanco (anexo 7). Cada requerimiento fue resultado con uno o varios indicadores mostrados en los reportes.

- ✚ Se desea saber la cantidad total de clientes en Mora en un determinado tiempo.
- ✚ Se desea conocer cuál fue el monto total de mora de cada producto a cada cliente en un determinado tiempo.
- ✚ Se desea conocer a los mejores clientes por zona geográfica, por producto, por mes y año.
- ✚ Se desea saber la cantidad total de dinero perdido por los clientes que cayeron en mora por departamento, provincia y distrito y en un determinado tiempo.
- ✚ Se quiere saber el patrón de los clientes que cayeron en mora en un determinado tiempo.
- ✚ Se quiere saber qué agencias son los que tienen más cantidad de clientes morosos en un determinado tiempo.
- ✚ Se desea saber que ciudades son los que tienen más clientes morosos en un determinado tiempo.
- ✚ Se desea conocer la ganancia anual y mensual de los desembolsos de créditos en un determinado tiempo.
- ✚ Se desea saber qué agencia es la que más desembolsa un producto de campaña y que tipo de clientes son los que más usan esos productos.
- ✚ Se desea saber el índice de morosidad por ciudad en un determinado tiempo.
- ✚ Se desea saber el índice de cobertura por ciudad en un determinado tiempo.
- ✚ Se desea conocer cuáles son los productos más desembolsados y cuanto es el monto total desembolsado.
- ✚ Se desea conocer que por tipo de asesores de negocios son los más

eficientes por meses y años.

- ✚ Se desea conocer que asesores son los menos eficientes en un determinado tiempo.
- ✚ Se desea conocer los mejores clientes por zona geográfica, por producto, por mes y año.
- ✚ Se quiere saber la efectividad de la base de campañas en un determinado tiempo.
- ✚ Se desea saber la cantidad promedio de clientes por tipo de asesores de negocios.
- ✚ Se desea saber los desembolsos por cada tipo de campaña en un determinado tiempo.
- ✚ Se desea saber la ganancia anual y mensual de los desembolsos.
- ✚ Se desea conocer los desembolsos realizados por zonas, región y territorio de forma mensual y anual.
- ✚ Se desea conocer la cantidad de desembolsos por tipo de clientes y por tipo de productos.
- ✚ Se desea saber el número total de entidades con la que trabaja cada cliente en un determinado tiempo.
- ✚ Se desea saber la deuda total que maneja cada cliente con todas las entidades financieras en un determinado tiempo.

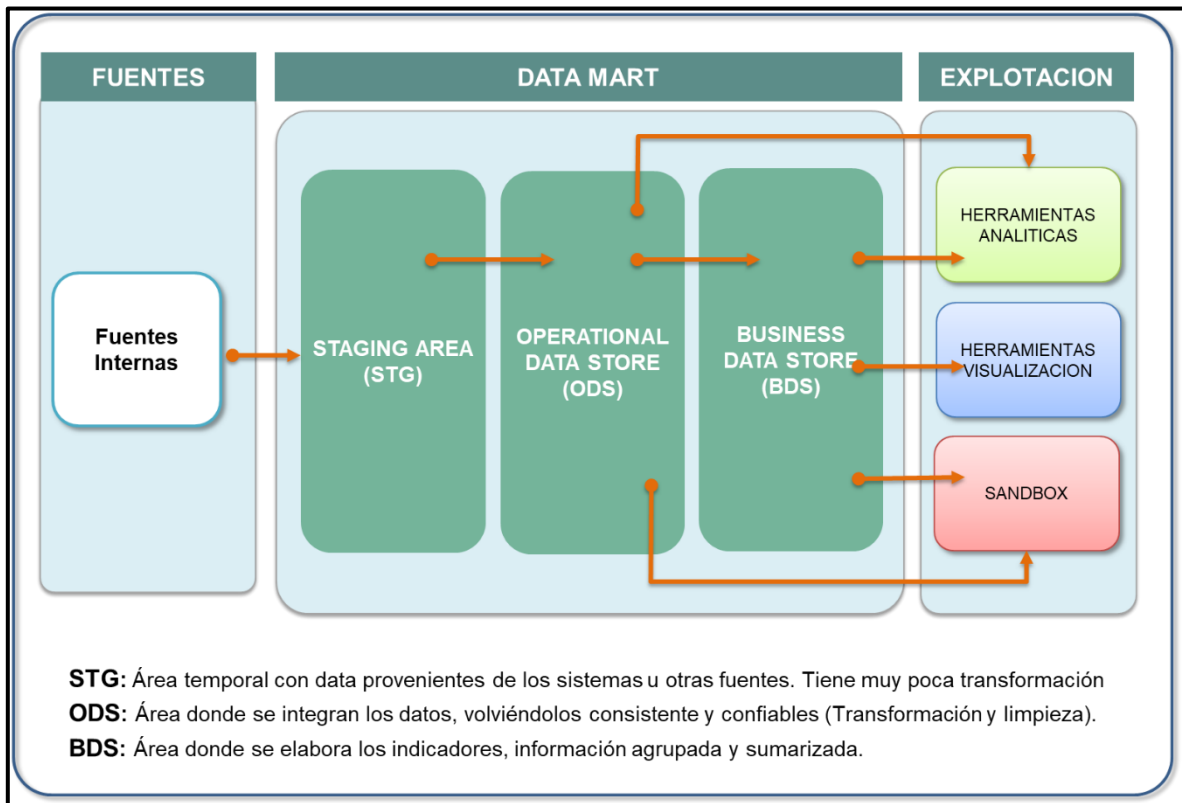
### **Diseño de la arquitectura**

Diagrama de arquitectura para el desarrollo de la investigación:

Para darle una mayor comprensión al desarrollo del Data mart, en la figura 2 se muestra las fases del desarrollo y como estas se relacionan con las herramientas de apoyo, desde las fuentes de datos hasta la elaboración de los indicadores y explotarlas con las herramientas analíticas.



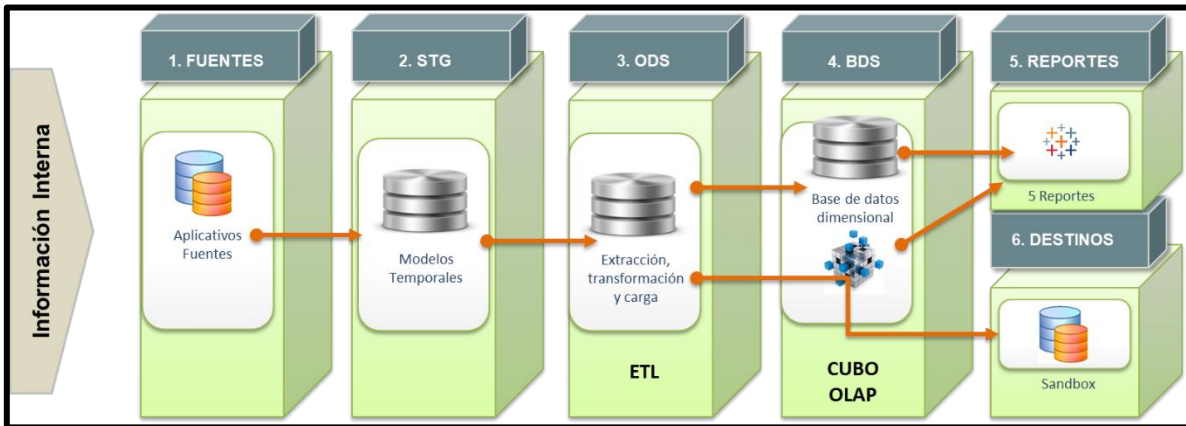
**Figura 2 Diagrama de arquitectura funcional del data mart**



Fuente: Elaboración propia

En la figura 2 se visualiza el diseño de la arquitectura de alto nivel del data mart, desde las fuentes de datos, pasando a través de vistas, así mismo por la parte de la extracción transformación y carga, cargándose toda la información en los cubos OLAP y mostrándose a través de reportes.

**Figura 3 Diagrama de arquitectura de alto nivel del data mart**



Fuente: Elaboración propia

Se detallará a continuación cada punto del diagrama de arquitectura para el desarrollo de la investigación para que sea más comprensible.

**Tabla N°6 Descripción del diagrama de arquitectura**

Nro	Título	Descripción
1	Información interna	La fuente de datos en donde se guardan los datos del Topaz es en el Oracle.
2	Modelos temporales	Limpieza y transformación y consolidación de la información de las fuentes de datos
3	ETL	Extracción, transformación y carga de los datos ya ordenados para la consolidación en el cubo OLAP
4	Cubo OLAP	Modelo dimensional para la explotación de la información de acuerdo con las necesidades de las áreas.

Fuente: Elaboración propia

### **3. Modelado dimensional**

En el proceso de diseño dimensional por Kimball se distingue 5 etapas:

- Definición del proceso de negocio
- Definición del grano
- Elección de las dimensiones
- Identificación de los hechos
- Detalle de las tablas de dimensión

#### **Definición del proceso de negocio**

En la presente investigación se eligió implementar el Data mart en la evaluación de créditos en el área de Campañas Comerciales de la empresa Mibanco.

Para una mayor visualización del proceso revisar el anexo 3.

#### **Definición de la granularidad**

Una vez que se ha definido el proceso de negocio, la siguiente tarea será la definición de la granularidad, o lo que es lo mismo, hasta que nivel de detalle se quiere alcanzar en el modelo del data mart y más concretamente en la tabla de hechos.

El grano más apropiado para el propósito de este proyecto se muestra en el punto del detalle de las dimensiones, donde se mostrará hasta qué punto de detalle se alcanzará por cada dimensión del proyecto del Data mart.

#### **Elección de las dimensiones**

Una vez obtenido los requerimientos por parte del área de Campañas Comerciales se procedió a la elección de las dimensiones, estas dimensiones fueron elegidas ya que cumplirán la labor de satisfacer las necesidades del área y así mismo dar la solución a la problemática planteada ya que servirán para consolidar toda la información de manera ordena y optima, así mismo también para la realización de análisis a detalle de toda la información. Las dimensiones definidas para este proyecto son:

**Dimensión Tiempo:** Para la selección de la dimensión tiempo se analizó los requerimientos solicitados.

Requerimientos:

Se desea saber la cantidad total de clientes en Mora en un determinado **tiempo**.

Se desea conocer la ganancia **anual y mensual** de los desembolsos de créditos.

**Dimensión Producto:** Para la selección de la dimensión producto se analizó los requerimientos solicitados.

Requerimientos:

Se desea conocer a los mejores clientes por zona geográfica, por **producto**, por mes y año.

Se desea conocer la cantidad de desembolsos por tipo de clientes y por **tipo de productos**.

**Dimensión Cliente:** Para la selección de la dimensión Cliente se analizó los requerimientos solicitados.

Requerimientos:

Se desea conocer cuál fue el monto total de mora de cada producto a cada **cliente** en un determinado tiempo.

Se desea saber qué agencia es la que más desembolsa un producto de campaña y que **tipo de clientes** son los que más usan esos productos.

**Dimensión analista:** Para la selección de la dimensión analista se analizó los requerimientos solicitados.

Requerimientos:

Se desea saber la cantidad promedio de clientes por tipo de **asesores** de negocios.

Se desea conocer que por **tipo de asesores** de negocios son los más eficientes por meses y años.

**Dimensión ubigeo:** Para la selección de la dimensión ubigeo se analizo los requerimientos solicitados.

Requerimientos:

Se desea conocer a los mejores clientes por [zona geográfica](#), por producto, por mes y año.

Se desea saber la cantidad total de dinero perdido por los clientes que cayeron en mora por [departamento, provincia y distrito](#) y en un determinado tiempo.

**Dimensión agencias:** Para la selección de la dimensión agencias se analizó los requerimientos solicitados.

Requerimientos:

Se quiere saber qué [agencias](#) son los que tienen más cantidad de clientes morosos en un determinado tiempo.

Se desea saber qué [agencia](#) es la que más desembolsa un producto de campaña y que tipo de clientes son los que más usan esos productos.

**Dimensión campaña:** Para la selección de la dimensión campaña se analizó los requerimientos solicitados.

Requerimientos:

Se desea saber los desembolsos por cada [tipo de campaña](#) en un determinado tiempo

En el punto **detalle de las dimensiones** se explicará a detalle la funcionalidad de cada dimensión.

### **Identificación de los hechos**

Una vez establecida las dimensiones, se comienza a definir la tabla hechos que nos permitirá conseguir la mayor precisión posible en la evaluación de créditos que es el interés de esta investigación. Teniendo en cuenta toda la información que se puede acumular y calcular según los requerimientos del área de campañas comerciales se procedió a crear la tabla Hechos llamada FACT\_DESEMBOLSOS.

En la figura 4 se visualiza la FACT\_DESEMBOLSOS, que servirá como asociación para las dimensiones que posteriormente servirán para la generación de reportes. Para obtener los indicadores se analizaron los requerimientos.

Requerimientos:

Se desea saber el **número total de entidades** con la que trabaja cada cliente en un determinado tiempo.

Se desea saber la **deuda total** que maneja cada cliente con todas las entidades financieras en un determinado tiempo

Se desea conocer cuáles son los productos más desembolsados y cuanto es el **monto total desembolsado**.

Se desea conocer cuál fue el **monto total de mora** de cada producto a cada cliente en un determinado tiempo.

Se desea conocer la **cantidad de desembolsos** por tipo de clientes y por tipo de productos.

**Figura 4 FACT\_DESEMBOLSOS**

FACT_DESEMBOLSOS
COD_CAM
COD_CLI
COD_PROD
CO_ANALISTA
COD_AGE
FECHA
COD_UBIGEO
NUMERO_ENTIDADES
TOT_DEU_DIR_RCC
MONTO_TOTAL_DESEMBOLSADO
MONTO_TOTAL_MORA
CANTIDAD_DESEMBOLSOS

Fuente: Elaboración propia

En la figura 5 se muestra la FACT\_MORA, la cual tendrá como medida a los indicadores de morosidad y que servirá como asociación para las dimensiones que posteriormente servirán para realizar los reportes.

Para obtener los indicadores de la tabla FACT\_MORA se tomó como referencia a los requerimientos.

Se desea saber el **índice de morosidad** por ciudad en un determinado tiempo.

Se desea saber el **índice de cobertura** por ciudad en un determinado tiempo.

Se desea conocer cuál fue el **monto total de mora** de cada producto a cada cliente en un determinado tiempo.

Se desea conocer la **cantidad de desembolsos** por tipo de clientes y por tipo de productos.

Se desea saber la **cantidad total de clientes** en Mora en un determinado tiempo.

**Figura 5 FACT\_MORA**

FACT_MORA
COD_CLI
CO_ANALISTA
COD_AGE
COD_UBIGEO
COD_PROD
CANTIDAD
MONTO_TOTAL_DESEMBOLSADO
MONTO_TOTAL_DETERIORADO
INDICE DE MOROSIDAD
INDICE DE COBERTURA

Fuente: Elaboración propia

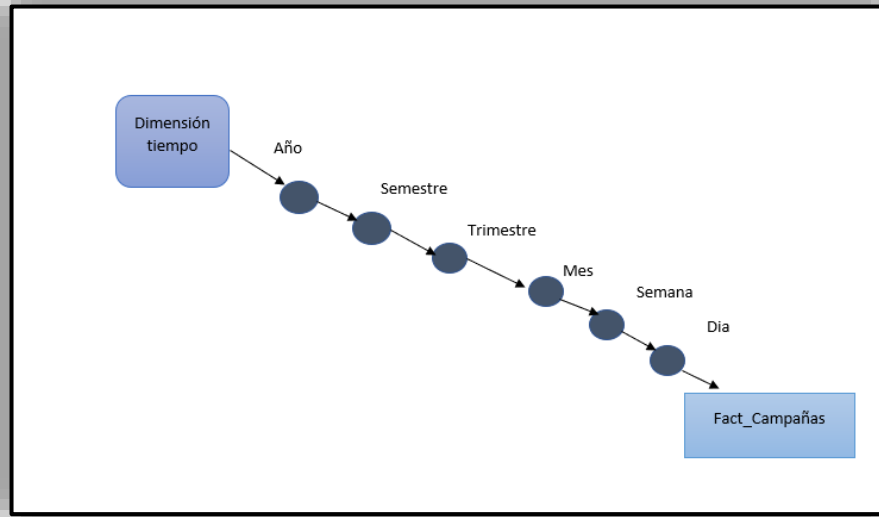
## **Detalle de las tablas de dimensión**

### **Dimensión Tiempo**

Esta dimensión abarca las fechas de los desembolsos de los clientes.

Se podrá observar en esta dimensión los periodos de desembolsos que realizo un cliente en Mibanco. En la Figura 6 se muestra la granularidad del tiempo, llegando hasta el nivel más bajo que exista.

**Figura 6 Granularidad de la dimensión tiempo**



Fuente: Elaboración propia

En la figura 7 se muestra los campos que contendrá la dimensión tiempo, desde su llave primaria que es el campo FECHA hasta el más mínimo detalle que viene hacer el campo día.

**Figura 7 Campos de la dimensión DM\_TIEMPO**

DM_TIEMPO
FECHA
FECHASK
AÑO
TRIMESTRE
MES
SEMANA
DIA
DIASEMANA
NTRIMESTRE
NMES
NMES3L
NSEMANA
NDIA
NDIASEMANA

Fuente: Elaboración propia

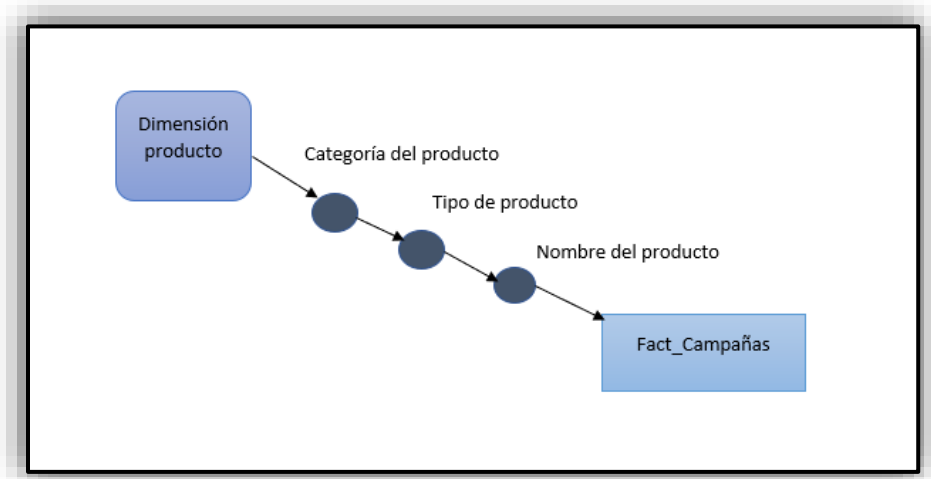


## Dimensión producto

Esta dimensión contiene todos los productos de Mibanco que los clientes desembolsan para terminar de realizar su préstamo. En esta dimensión se encuentran los productos de Campañas y también los productos de catálogo.

En la figura 8 se muestra el nivel de granularidad de la dimensión producto, desagregando la categoría del producto hasta llegar al mínimo detalle que es el nombre del producto.

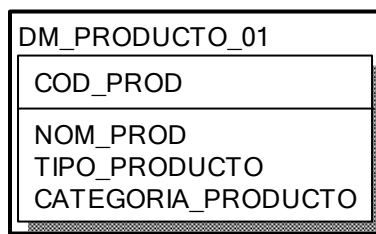
**Figura 8 Nivel de granularidad de la dimensión producto**



Fuente: Elaboración propia

En la figura 9 se muestra los campos que contendrá la dimensión producto, teniendo como llave primaria al código de producto.

**Figura 9 Campos de la dimensión producto**



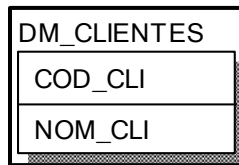
Fuente: Elaboración propia

### **Dimensión Cliente**

Esta dimensión contiene todos los clientes de Mibanco, en la tabla se describirá las características que poseen todos los clientes, los cuales servirán para realizar los análisis correspondientes.

En la figura 10 se muestra los campos que contendrá la dimensión cliente, teniendo como llave primaria al código de cliente.

**Figura 10 Campos de la dimensión Cliente**



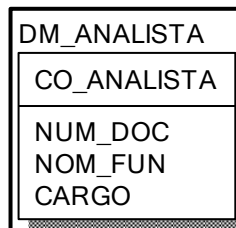
Fuente: Elaboración propia

### **Dimensión analista**

Esta dimensión contiene la información de los funcionarios que trabajan en Mibanco, observaremos que cargo tiene cada analista, si está activo o inactivo.

En la figura 11 se muestra los campos que tendrá la dimensión analista, teniendo como llave primaria al código de analista.

**Figura 11 Campos de la dimensión analista**

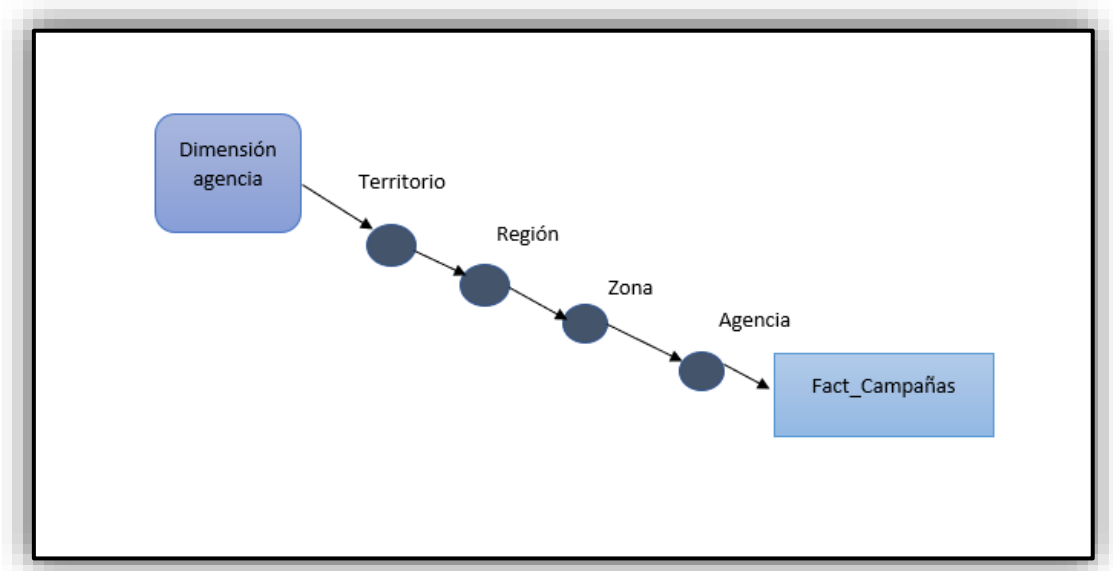


Fuente: Elaboración propia

## Dimensión agencia

Esta dimensión contiene la información de todas las agencias de Mibanco que están distribuidas de forma estratégicas en el Perú. En la figura 12 se muestra el nivel de granularidad de la dimensión agencia, desagregando desde el territorio hasta llegar al mínimo detalle que es la agencia.

**Figura 12 Nivel de granularidad de la dimensión agencia**



Fuente: Elaboración propia

En la figura 13 se muestra los campos que tendrá la dimensión agencia, teniendo como llave primaria al código de agencia.

**Figura 13 Campos de la dimensión Agencia**

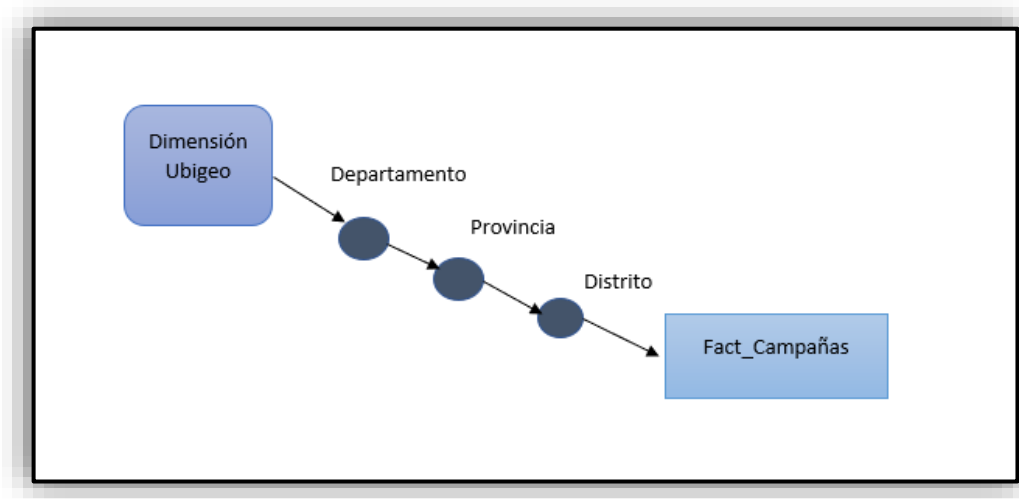
DM_AGENCIA
COD_AGE
NOM_AGE
ZONA
REGION
TERRITORIO

Fuente: Elaboración propia

## Dimensión Ubigeo

Esta dimensión tiene información de los distritos, provincias, departamentos a nivel nacional, el cual nos servirá para analizar información a ese nivel. En la figura 14 se muestra el nivel de granularidad de la dimensión ubigeo, desagregando desde el nivel de departamento, hasta llegar al mínimo detalle con el distrito.

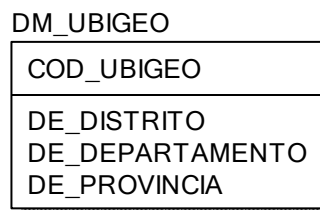
**Figura 14 Nivel de granularidad de la dimensión ubigeo**



Fuente: Elaboración propia

En la figura 15 se muestra todos los campos que tendrá la dimensión ubigeo, teniendo como llave primaria al campo código de ubigeo

**Figura 15 Campos de la dimensión ubigeo**



Fuente: Elaboración propia

### **Dimensión Campañas**

Esta dimensión contendrá información de todas las campañas ya realizadas en Mibanco. En la figura 16 se muestra los campos que tendrá la dimensión campañas, que tendrá como llave primaria al código de campaña.

**Figura 16 Campos de la dimensión campañas**

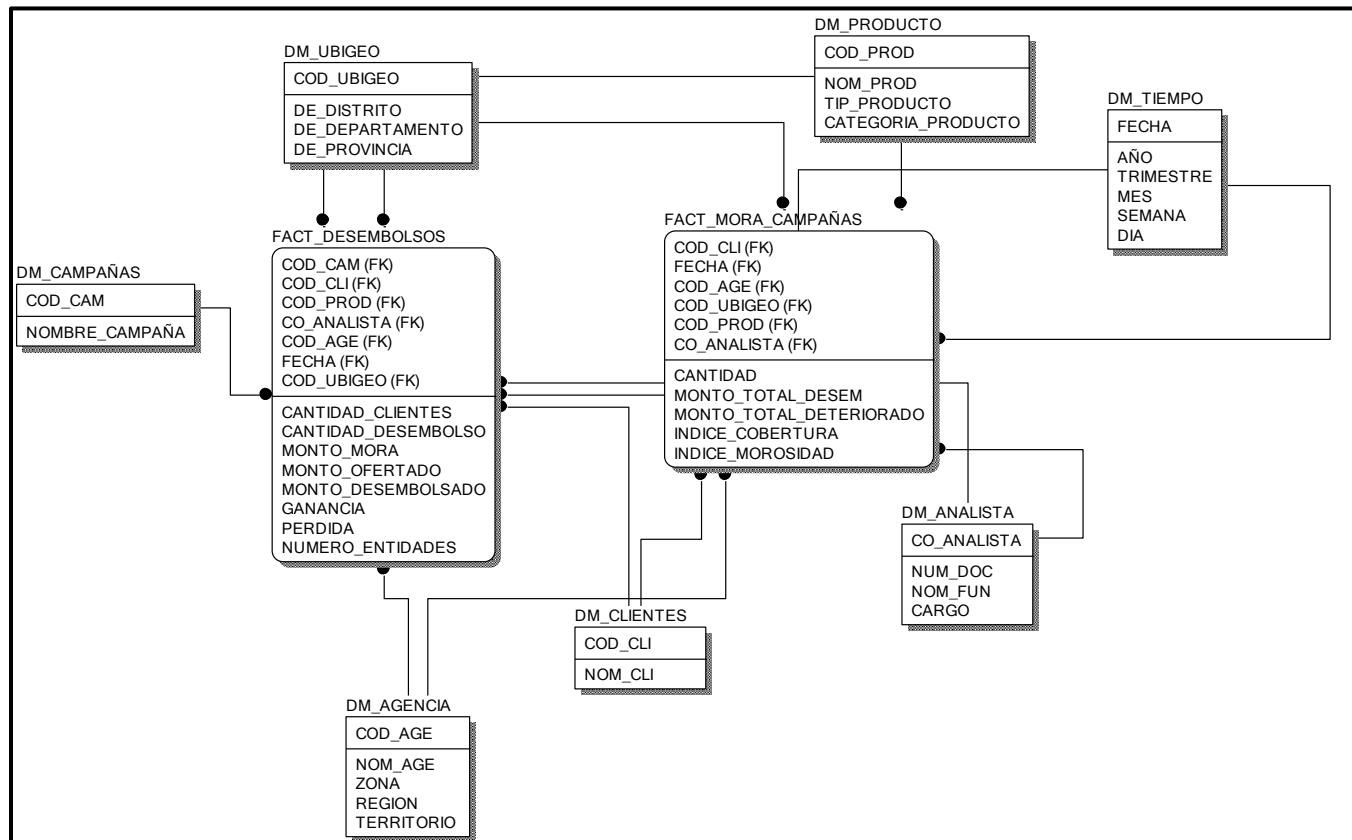


Fuente: Elaboración propia

#### 4. Diseño Físico

**Diseño Lógico:** Para la creación del modelo lógico se tomó en cuenta las dimensiones y tablas hechos analizados a través de los indicadores, tomando forma del modelo constelación (Contiene 2 o más tablas hechos).

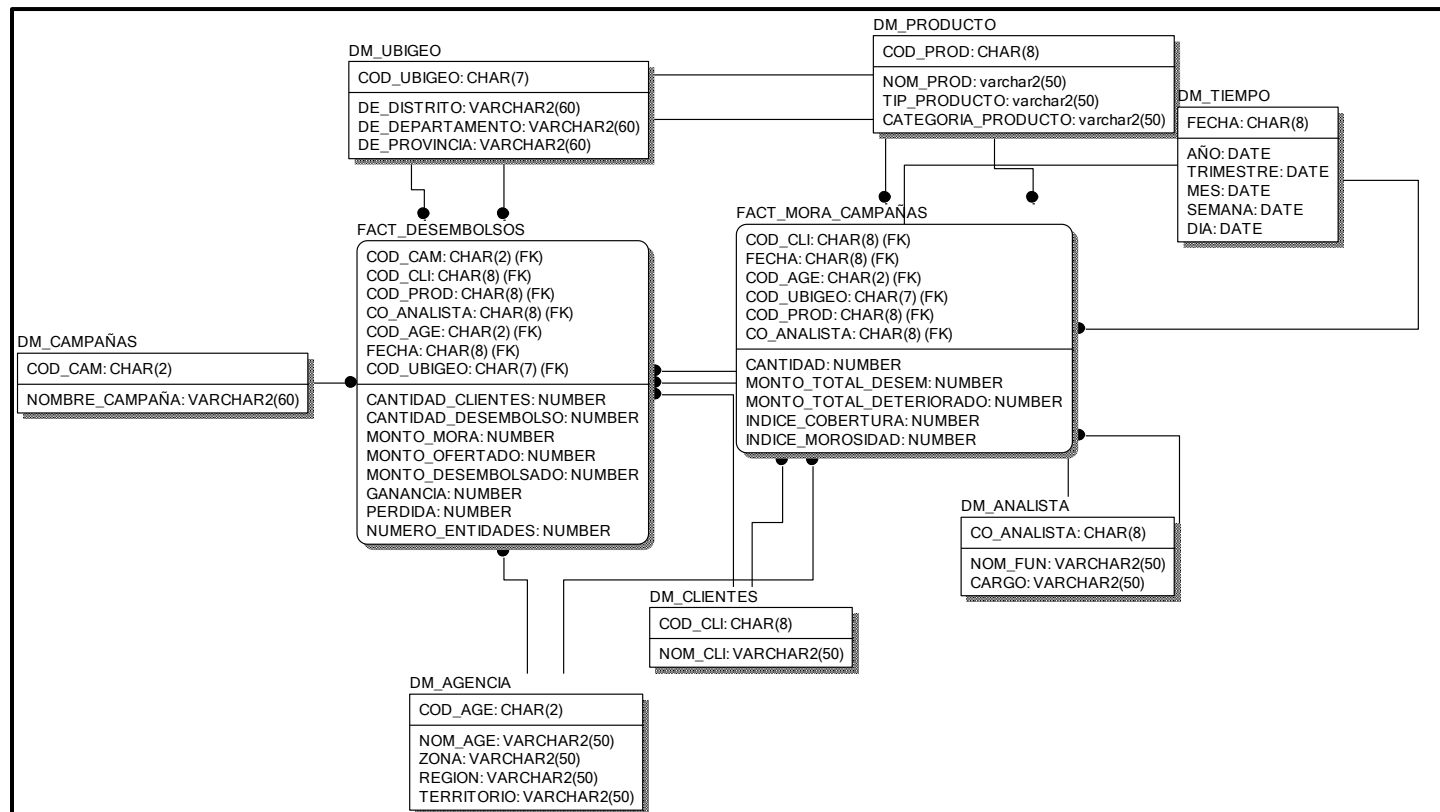
**Figura 17 Diseño Lógico de la base de datos del Data mart**



Fuente: Elaboración propia

**Diseño Físico:** El modelo físico presentado muestra todas las estructuras de la tabla, incluidos el nombre de la columna, el tipo de datos, las llaves primarias y secundarias, las relaciones que hay entre las dimensiones.

**Figura 18 Diseño Físico de la base de datos del data mart**



Fuente: Elaboración propia

## Diccionario de datos

Desde la tabla 7 hasta la tabla 10 se visualiza el diccionario de datos del data mart

**Tabla 7 Diccionario de datos**

<b>Schema</b>	<b>Table</b>	<b>Column</b>	<b>Datatype</b>	<b>Size</b>	<b>Nullable</b>	<b>InPrimaryKey</b>	<b>IsForeignKey</b>	<b>Description</b>
<b>Dbo</b>	DM_Campañas	Cod_cam	Char(2)	2	N	Y	N	Código principal de campaña
<b>Dbo</b>	DM_Campaña	Nombre_campaña	Varchar2(60)	60	N	N	N	Nombre de la campaña
<b>Dbo</b>	DM_Ubigeo	Cod_Ubigeo	Char(7)	7	N	Y	N	Código de ubigeo
<b>Dbo</b>	DM_Ubigeo	DE_Distrito	Varchar(60)	60	N	N	N	Descripción del distrito
<b>Dbo</b>	DM_Ubigeo	DE_Provincia	Varchar(60)	60	N	N	N	Descripción del distrito
<b>Dbo</b>	DM_Ubigeo	DE_Departamento	Varchar(60)	60	N	N	N	Descripción del distrito
<b>Dbo</b>	DM_Agencia	Cod_age	Char(2)	2	N	Y	N	Código de agencia
<b>Dbo</b>	DM_Agencia	Nom_age	Varchar2(50)	50	N	N	N	Nombre de la agencia
<b>Dbo</b>	DM_Agencia	Zona	Varchar2(50)	50	N	N	N	Nombre de la zona
<b>Dbo</b>	DM_Agencia	Región	Varchar2(50)	50	N	N	N	Nombre de la Región
<b>Dbo</b>	DM_Agencia	Territorio	Varchar2(50)	50	N	N	N	Nombre del Territorio

Fuente: Elaboración propia



**Tabla 8 Diccionario de datos**

<b>Schema</b>	<b>Table</b>	<b>Column</b>	<b>Datatype</b>	<b>Size</b>	<b>Nullable</b>	<b>InpPrimaryKey</b>	<b>IsForeignKey</b>	<b>Description</b>
<b>Dbo</b>	DM_Clientes	Cod_cli	Char(8)	8	N	Y	N	Código de cliente
<b>Dbo</b>	DM_Clientes	Nom_cli	Varchar2(50)	50	N	N	N	Nombre del cliente
<b>Dbo</b>	DM_Producto	Cod_prod	Char(8)	8	N	Y	N	Código de producto
<b>Dbo</b>	DM_Producto	Nom_prod	Varchar2(50)	50	N	N	N	Nombre del producto
<b>Dbo</b>	DM_Producto	Tip_producto	Varchar2(50)	50	N	N	N	Tipo de producto
<b>Dbo</b>	DM_Producto	Categoria_producto	Varchar2(50)	50	N	N	N	Categoría del producto
<b>Dbo</b>	DM_Analista	Co_analista	Char(8)	8	N	Y	N	Código de analista
<b>Dbo</b>	DM_Analista	Nom_func	Varchar2(50)	50	N	N	N	Nombre del funcionario
<b>Dbo</b>	DM_Analista	Cargo	Varchar2(50)	50	N	N	N	Cargo del funcionario
<b>Dbo</b>	DM_Tiempo	Fecha	Varchar(8)	8	N	Y	N	Fecha, incluye día mes y año
<b>Dbo</b>	DM_Tiempo	Trimestre	Date		N	N	N	Trimestre de cada año
<b>Dbo</b>	DM_Tiempo	Mes	Date		N	N	N	Mes de cada año
<b>Dbo</b>	DM_Tiempo	Semana	Date		N	N	N	Semana de cada mes
<b>Dbo</b>	DM_Tiempo	Día	Date		N	N	N	Días de tomas las semanas

Fuente: Elaboración propia

**Tabla 9 Diccionario de datos**

<b>Schema</b>	<b>Table</b>	<b>Column</b>	<b>Datatype</b>	<b>Size</b>	<b>Nullable</b>	<b>InpPrimaryKe</b>	<b>IsForeignKey</b>	<b>Description</b>
<b>Dbo</b>	FACT_DESEMBOLSOS	Cod_cam	Char(2)	2	N	N	Y	Código de campaña
<b>Dbo</b>	FACT_DESEMBOLSOS	Cod_cli	Char(8)	8	N	N	Y	Código de cliente
<b>Dbo</b>	FACT_DESEMBOLSOS	Cod_prod	Char(8)	8	N	N	Y	Código de producto
<b>Dbo</b>	FACT_DESEMBOLSOS	Co_analista	Char(8)	8	N	N	Y	Código de analista
<b>Dbo</b>	FACT_DESEMBOLSOS	Cod_age	Char(2)	2	N	N	Y	Código de agencia
<b>Dbo</b>	FACT_DESEMBOLSOS	Fecha	Char(8)	8	N	N	Y	Código de Fecha
<b>Dbo</b>	FACT_DESEMBOLSOS	Cod_Ubigeo	Char(7)	7	N	N	Y	Código de ubigeo
<b>Dbo</b>	FACT_DESEMBOLSOS	Cantidad_clientes	Number		N	N	N	Cantidad de clientes
<b>Dbo</b>	FACT_DESEMBOLSOS	Cantidad_desembolso	Number		N	N	N	Cantidad de desembolsos
<b>Dbo</b>	FACT_DESEMBOLSOS	Monto_mora	Number		N	N	N	Monto que cayó en mora
<b>Dbo</b>	FACT_DESEMBOLSOS	Monto_ofertado	Number		N	N	N	Monto ofertado
<b>Dbo</b>	FACT_DESEMBOLSOS	Monto_desembolsado	Number		N	N	N	Monto desembolsado
<b>Dbo</b>	FACT_DESEMBOLSOS	Ganancia	Number		N	N	N	Ganancia
<b>Dbo</b>	FACT_DESEMBOLSOS	Perdida	Number		N	N	N	perdida
<b>Dbo</b>	FACT_DESEMBOLSOS	Cantidad_entidades	Number		N	N	N	Cantidad de entidades

Fuente: Elaboración propia

**Tabla 10 Diccionario de datos**

<b>Schema</b>	<b>Table</b>	<b>Column</b>	<b>Datatype</b>	<b>Size</b>	<b>Nullable</b>	<b>InpPrimaryKey</b>	<b>IsForeignKey</b>	<b>Description</b>
<b>Dbo</b>	FACT_MORA	Cod_cli	Char(8)	8	N	N	Y	Código de cliente
<b>Dbo</b>	FACT_MORA	Cod_prod	Char(8)	8	N	N	Y	Código de producto
<b>Dbo</b>	FACT_MORA	Co_analista	Char(8)	8	N	N	Y	Código de analista
<b>Dbo</b>	FACT_MORA	Cod_age	Char(2)	2	N	N	Y	Código de agencia
<b>Dbo</b>	FACT_MORA	Fecha	Char(8)	8	N	N	Y	Código de Fecha
<b>Dbo</b>	FACT_MORA	Cod_Ubigeo	Char(7)	7	N	N	Y	Código de ubigeo
<b>Dbo</b>	FACT_MORA	Cantidad	Number		N	N	N	Cantidad de clientes
<b>Dbo</b>	FACT_MORA	Monto_total_desem	Number		N	N	N	Monto total desembolsado
<b>Dbo</b>	FACT_MORA	Monto_total_deteriorado	Number		N	N	N	Monto total mora
<b>Dbo</b>	FACT_MORA	Índice_cobertura	Number		N	N	N	Índice de cobertura
<b>Dbo</b>	FACT_MORA	Índice_morosidad	Number		N	N	N	Índice de morosidad

Fuente: Elaboración propia

## 5. Diseño e implementación del subsistema de ETL

Extracción de la información: Para la presente investigación, se extraerá la información directamente de la base de datos Oracle, ya que esta base de datos se alimenta del sistema transaccional Topaz.

**Tabla 11 Relación entre la información y el sistema transaccional**

Información	Sistema	Descripción
Desembolsos, morosidad, agencias, clientes, ubigeo, campañas, producto, analista	Topaz	Mediante toda la información almacenada del software Topaz, se extraerá todo el detalle almacenado en la base de datos Oracle, para la creación de las dimensiones y así poder poblar el cubo OLAP.

Fuente: Elaboración propia

Una vez teniendo definido de donde se extraerá la información, se procederá a poblar las dimensiones de manera automática mediante procedimientos almacenados que se ejecutaran mediante un JOB de manera automática mensualmente en la misma base de datos Oracle.

En la figura 19 se muestra la creación del procedimiento almacenado para la extracción, transformación y carga de la dimensión producto.

Para el poblado de la dimensión producto se extrajo la información de la tabla transaccional productos, agrupando por tipo de producto y por categoría de producto.

**Figura 19 Creación del procedimiento para actualizar la dimensión producto**

```
CREATE OR REPLACE PROCEDURE ACTUALIZA_PRODUCTO
AS
V_QUERY CLOB;
BEGIN
EXECUTE IMMEDIATE 'TRUNCATE TABLE FACT_DESEMBOLSOS';
DELETE FROM DM_PRODUCTO_01;
V_QUERY := 'insert into DM_PRODUCTO_01
select COD_PROD, descripcion, CASE WHEN cod_prod in (1560 ,
1526, 1547, 1559, 1536, 1713, 1714, 1716,
1076, 1718, 1715, 1717, 1048, 1051,
1077, 1040, 1041, 1042, 1043, 1049, 1050,
1096, 1097 ) THEN ''PRODUCTO-CAMPAÑA''
ELSE ''PRODUCTO-CATALOGO'' END AS TIPO_PRODUCTO,
CASE WHEN cod_prod in (1412,1413,1416,1417,1422,1424,1425,1428,1414,
1415,1406,1407,1426,1427,1643,1631,1408,1409,1410,1411,1018,1019,
1020,1022,1026,1232,1251,1400,1401,1402,1403,1404,1405,1668,1670,1430) THEN ''PRODUCTO-PASIVO''
ELSE ''PRODUCTO-ACTIVO'' END AS CATEGORIA_PRODUCTO
from system.PRODUCTOS';--##TABLA DE DONDE EXTRAERA LA INFORMACION
EXECUTE IMMEDIATE V_QUERY;
COMMIT;
END;
```

Fuente: Elaboración propia

En la figura 20 se crea el procedimiento para la extracción y transformación y carga, para el llenado de la dimensión agencia.

Para el poblado de la dimensión Agencia se extrajo la información de un archivo plano, cargando a la tabla PR\_AGENCIAS y marcando quienes mes a mes se encuentran activas e inactivas.

Procederá primero limpiando la dimensión, luego procederá con el poblado agregando las nuevas agencias creadas en Perú.

**Figura 20 Creación del procedimiento para actualizar la dimensión Agencia**

```
CREATE OR REPLACE PROCEDURE ACTUALIZA_AGENCIAS
AS
V_QUERY CLOB;
BEGIN
EXECUTE IMMEDIATE 'TRUNCATE TABLE FACT_DESEMBOLSOS';
DELETE FROM DM_AGENCIA;
V_QUERY := 'INSERT INTO DM_AGENCIA
select distinct COD_AGE, AGE, ZON_RC, REGION ,TERRITORIO
FROM system.PR_AGENCIAS';--##TABLA DE DONDE EXTRAERA LA INFORMACION
EXECUTE IMMEDIATE V_QUERY;
COMMIT;
END;
```

Fuente: Elaboración propia

En la figura 21 se muestra la creación del procedimiento almacenado para la población de la dimensión DM\_CAMPAÑAS.

Para el poblado de la dimensión DM\_CAMPAÑAS se extrajo la información de la tabla transaccional campañas\_mibanco, agrupándolas por tipo de campañas.

Primero se procederá limpiando la dimensión y luego se volverá agregar los datos actualizados de la tabla transaccional.

**Figura 21 Creación del procedimiento para actualizar la dimensión Campañas**

```
CREATE OR REPLACE PROCEDURE ACTUALIZA_CAMPANAS
AS
V_QUERY CLOB;
BEGIN
EXECUTE IMMEDIATE 'TRUNCATE TABLE FACT_DESEMBOLSOS';
DELETE FROM DM_CAMPANAS;
V_QUERY := 'insert into DM_CAMPANAS
select nom_cam,cod_cam from
system.CAMPANAS_Mibanco
where tz_lock=0 and nom_com_cam is not null
order by cod_cam asc';--##TABLA DE DONDE EXTRAERA LA INFORMACION
EXECUTE IMMEDIATE V_QUERY;
COMMIT;
END;
```

Fuente: Elaboración propia

En la figura 22 se observa la creación del procedimiento almacenado para el poblado de la dimensión DM\_ANALISTA.

Para el poblado de la dimensión DM\_ANALISTA se extrajo la información de la tabla RH\_FUNCIONARIO, agrupándolo por Cargo.

Primero se procederá con la eliminación de los datos de la dimensión, luego se procederá a cargar la nueva información.

**Figura 22 Creación del procedimiento para actualizar la dimensión Analista**

```
CREATE OR REPLACE PROCEDURE ACTUALIZA_ANALISTA
AS
V_QUERY CLOB;
BEGIN
EXECUTE IMMEDIATE 'TRUNCATE TABLE FACT_DESEMBOLSOS';
DELETE FROM DM_ANALISTA;
V_QUERY := '
insert into DM_ANALISTA
select trim(codfuncionario) as co_analista,
nombre as nom_func,TRIM(CATEGORIA) AS cargo
from system.RH_FUNCIONARIO';--##TABLA DE DONDE EXTRAERA LA INFORMACION
EXECUTE IMMEDIATE V_QUERY;
COMMIT;
END;
```

Fuente: Elaboración propia

En la figura 23 se observa la creación del procedimiento almacenado para poder actualizar la dimensión DM\_CLIENTES, haciendo la extracción, transformación y carga.

Para el poblado de la dimensión se extrajo la información de la tabla transaccional cam\_det\_cam que contiene toda la información de todos los clientes en campañas, se agrupo por tipo de cliente.

Se procede con la limpieza de los datos actuales de la dimensión, luego se procederá con el poblado de la dimensión con los nuevos registros.

**Figura 23 Creación del procedimiento para actualizar la dimensión Clientes**

```
CREATE OR REPLACE PROCEDURE ACTUALIZA_CLIENTES
AS
V_QUERY CLOB;

BEGIN
EXECUTE IMMEDIATE 'TRUNCATE TABLE FACT_DESEMBOLSOS';
DELETE FROM DM_CLIENTES;
V_QUERY := 'INSERT INTO DM_CLIENTES
SELECT distinct trim(cod_Cli) as cod_Cli,

max(nom_cli) as nom_cli
FROM SYSTEM.CAM_DET_CAM
where cod_Cli<>0
group by cod_Cli';--#TABLA DE DONDE EXTRAERA LA INFORMACION
EXECUTE IMMEDIATE V_QUERY;

COMMIT;
END;
```

Fuente: Elaboración propia

En la figura 24 se crea el procedimiento almacenado para poblar la dimensión DM\_UBIGEO. Para el poblado de la dimensión, se extrajo la información de la la tabla transaccional ubigeo, completando con los 0 faltantes en esa tabla.

Primero lo que realiza el procedimiento es eliminar los datos de la dimensión, luego volverá a cargar la nueva información más actualizada.

**Figura 24 Creación del procedimiento para actualizar la dimensión Ubigeo**

```
CREATE OR REPLACE PROCEDURE ACTUALIZA_UBIGEO
AS
V_QUERY CLOB;
BEGIN
EXECUTE IMMEDIATE 'TRUNCATE TABLE FACT_DESEMBOLSOS';
DELETE FROM DM_UBIGEO;
V_QUERY := 'INSERT INTO DM_UBIGEO
SELECT LPAD(TRIM(COD_UBIGEO),6,'0') as COD_UBIGEO, DE_DISTRITO, DE_DEPARTAMENTO, DE_PROVINCIA
FROM SYSTEM_UBIGEO';--#TABLA DE DONDE EXTRAERA LA INFORMACION
EXECUTE IMMEDIATE V_QUERY;
COMMIT;
END;
```

Fuente: Elaboración propia

En la figura 25 se muestra la forma de cómo se creó el procedimiento almacenado para poder poblar la dimensión DM\_TIEMPO.

Para el poblado se creó el procedimiento para elegir el rango de fechas y poder realizar los cálculos, por día, semana, mes y año.

Primero se limpia la información de la la dimensión, luego se procede en la actualización con las nuevas fechas ingresadas.

**Figura 25 Creación del procedimiento para actualizar la dimensión Fecha**

```
CREATE OR REPLACE PROCEDURE PROCEDURE_FECHA
AS
FechaDesde date;
FechaHasta date;
BEGIN
DELETE FROM DM_TIEMPO;
FechaDesde := TO_DATE('20060101','YYYYMMDD');
FechaHasta := TO_DATE((TO_CHAR(sysdate,'YYYY')+2 || '1231'),'YYYYMMDD');
WHILE FechaDesde <= FechaHasta LOOP
INSERT INTO DM_TIEMPO
(FechaSK, Fecha, Año, Trimestre, Mes, Semana, Dia, DiaSemana, NTrimestre, NMes, NMes3L, NSemana, NDia, NDiaSemana)
VALUES
(to_char(FechaDesde, 'YYYYMMDD'), FechaDesde, to_char(FechaDesde, 'YYYY'),
to_char(FechaDesde, 'Q'), to_char(FechaDesde, 'MM'), to_char(FechaDesde, 'WW'),
to_char(FechaDesde, 'DD'), to_char(FechaDesde, 'D'),
'T' || to_char(FechaDesde, 'Q') || '/' || to_char(FechaDesde, 'YY'),
to_char(FechaDesde, 'MONTH'), to_char(FechaDesde, 'MON'),
'Sem ' || to_char(FechaDesde, 'WW') || '/' || to_char(FechaDesde, 'YY'),
to_char(FechaDesde, 'DD MON'), to_char(FechaDesde, 'DAY'));
--Incremento del bucle
FechaDesde := FechaDesde + 1;
END LOOP;
DBMS_OUTPUT.PUT_LINE('SE ACTUALIZO LA DIMENSION FECHA: HASTA LA FECHA DE ' || FechaHasta);
END;
```

Fuente: Elaboración propia

En la figura 26 se observa cómo se extra, transforma y carga la información en la FACT\_TABLE\_DESEMBOLSOS.

Para el poblado de la tabla hecho FACT\_CAMPAÑAS se extrajo la información de las dimensiones para relacionar a través de sus llaves primarias y además se realizó los cálculos de los indicadores seleccionados de los requerimientos.

En la tabla FACT\_CAMPAÑAS solo se procederá con insertar nuevos registros ya que es una tabla consolidada histórica.



**Figura 26 Creación del procedimiento para actualizar la tabla hecho Fact\_Campañas**

```
CREATE OR REPLACE PROCEDURE ACTUALIZA_FACT_DESEMBOLOSOS AS
V_QUERY CLOB;
BEGIN
EXECUTE IMMEDIATE 'TRUNCATE TABLE FACT_DESEMBOLOSOS';
V_QUERY :=
'insert into FACT_DESEMBOLOSOS SELECT DISTINCT C.COD_CAM, A.COD_CLI, D.COD_PROD, E.CO_ANALISTA,
M.COD_AGE, F.FECHA, H.COD_UBIGEO, (B.NUM_ENT*1) AS NUMERO_ENTIDADES, REPLACE((B.TOT_DEU_DIR_RCC),',',',',',',') AS TOT_DEU_
(B.PER_INA*1) AS PER_INA, SUM(B.MON_DES*1) AS MON_DES, SUM(I.MON_CAM_01*1) AS MONTO_TOTAL_OFERTADO,
SUM(J.MON_DET) AS MONTO_TOTAL_MORA, SUM(K.CANTIDAD_CLIENTES) AS CANTIDAD_CLIENTES_CAMPANA,
SUM(L.CANTIDAD_CLIENTES) AS CANTIDAD_CLIENTES_DESEM FROM DM_CLIENTES A INNER JOIN SYSTEM.CAM_DET_CAM_RIE B ON A.COD_C
LEFT JOIN DM_CAMPANAS C ON TRIM(B.COD_CAM)=TRIM(C.COD_CAM)
LEFT JOIN DM_PRODUCTO_01 D ON D.COD_PROD=B.COD_PRO
LEFT JOIN DM_ANALISTA E ON E.CO_ANALISTA=B.COD_FUN LEFT JOIN DM_TIEMPO F ON F.FECHA=B.FEC_DES
LEFT JOIN system.temp_cli_ubigeo G ON G.cod_cli=a.cod_cli--#PARA EXTRAER LOS UBIGEOS
LEFT JOIN DM_UBIGEO H ON LPAD(TRIM(G.COD_UBIGEO),6,'0')=LPAD(TRIM(H.COD_UBIGEO),6,'0')
LEFT JOIN SYSTEM.CAM_DET_CAM I ON TRIM(I.COD_CLI)=TRIM(B.COD_CLI) AND I.COD_CAM=B.COD_CAM
LEFT JOIN SYSTEM.clientes_mora_mibanco_2 J on TRIM(B.NUM_PRE)=TRIM(J.COD_PRE)
LEFT JOIN (SELECT COUNT(DISTINCT NUM_DOC) AS CANTIDAD_CLIENTES, COD_CAM FROM SYSTEM.CAM_DET_CAM WHERE FLG_ELI=0
GROUP BY COD_CAM) K ON K.COD_CAM=B.COD_CAM
LEFT JOIN (SELECT COUNT(DISTINCT NUM_DOC) AS CANTIDAD_CLIENTES, COD_CAM FROM system.CAM_DET_CAM_RIE
WHERE TZ_LOCK=0 GROUP BY COD_CAM) I ON I.COD_CAM=B.COD_CAM LEFT JOIN DM_AGENCIA M ON TRIM(M.COD_AGE)=TRIM(B.COD_OFI)
where m.cod_age is not NULL AND C.COD_CAM IS NOT NULL AND D.COD_PROD IS NOT NULL
AND E.CO_ANALISTA IS NOT NULL AND h.cod_ubigeo IS NOT NULL
GROUP BY C.COD_CAM, A.COD_CLI, D.COD_PROD, E.CO_ANALISTA, M.COD_AGE, F.FECHA, H.COD_UBIGEO,
B.NUM_ENT, B.TOT_DEU_DIR_RCC, B.PER_INA';--#TABLA DE DONDE EXTRAERA LA INFORMACION
EXECUTE IMMEDIATE V_QUERY;
```

Fuente: Elaboración propia

En la figura 27 se creó el procedimiento almacenado para poder poblar la tabla FACT\_MORA.

Para el poblado de la tabla hecho FACT\_MORA se realizó las relaciones con las dimensiones y el cálculo de los indicadores obtenidos por los requerimientos.

Se procede con la inserción de los nuevos registros de manera histórica.

**Figura 27 Creación del procedimiento FACT\_MORA**

```
CREATE OR REPLACE PROCEDURE ACTUALIZA_FACT_MORA AS
V_QUERY CLOB;
BEGIN
EXECUTE IMMEDIATE 'TRUNCATE TABLE FACT_MORA_MIBANCO';
V_QUERY :=
'insert into FACT_MORA_MIBANCO
SELECT A.COD_CLI, ANAL.CO_ANALISTA, G.FECHA, F.COD_AGE, E.COD_UBIGEO, H.COD_PROD,
COUNT(DISTINCT A.COD_CLI) AS CANT_OPER,
SUM(B.MON_DES) AS MON_DES,
sum(C.MON_DET) AS MON_DET,
0 AS INDICE_COBERTURA,
round((sum(C.MON_DET)/sum(B.MON_DES)).2) AS RATIO_MOROSIDAD
FROM DM_CLIENTES A
inner JOIN SYSTEM.CAM_DET_CAM_RIE B ON trim(A.COD_CLI)=trim(B.COD_CLI)
INNER JOIN SYSTEM.clientes_mora_mibanco_2 C on TRIM(B.NUM_PRE)=TRIM(C.COD_PRE)
LEFT JOIN system.temp_cli_ubigeo D ON trim(D.COD_CLI)=trim(A.COD_CLI)
LEFT JOIN DM_UBIGEO E ON LPAD(TRIM(D.COD_UBIGEO),6,'0')=LPAD(TRIM(E.COD_UBIGEO),6,'0')
LEFT JOIN DM_AGENCIA F ON TRIM(F.COD_AGE)=TRIM(B.COD_OFI)
LEFT JOIN DM_TIEMPO G ON G.FECHA=B.FEC_DES
LEFT JOIN DM_PRODUCTO_01 H ON H.COD_PROD=B.COD_PRO
LEFT JOIN DM_ANALISTA ANAL ON ANAL.CO_ANALISTA=B.COD_FUN
WHERE F.COD_AGE IS NOT NULL AND H.COD_PROD IS NOT NULL AND E.COD_UBIGEO IS NOT NULL AND ANAL.CO_ANALISTA IS NOT NULL
GROUP BY A.COD_CLI, ANAL.CO_ANALISTA, G.FECHA, F.COD_AGE, H.COD_PROD, E.COD_UBIGEO
';--#TABLA DE DONDE EXTRAERA LA INFORMACION
EXECUTE IMMEDIATE V_QUERY;
```

Fuente: Elaboración propia

En la figura 28 se crea el JOB para poder ejecutar de manera automática todos los procedimientos creados por cada dimensión, se elegirá cada cuanto tiempo se debe ejecutar.

**Figura 28 Creación del JOB de todos los procedimientos credos**

```
DECLARE
X NUMBER;
BEGIN
SYS.DBMS_JOB.SUBMIT
( job => X ,what =>
'BEGIN
ACTUALIZA_AGENCIAS;
ACTUALIZA_PRODUCTO;
ACTUALIZA_ANALISTA;
ACTUALIZA_CAMPANAS;
ACTUALIZA_CLIENTES;
ACTUALIZA_UBIGEO;
PROCEDURE_FECHA;
ACTUALIZA_FACT_DESEMBOLSOS;
END;
' ,next_date => to_date('23/06/2018 01:23:00','dd/mm/yyyy hh24:mi:ss') ,interval => 'SYSDATE + 30'
,no_parse => FALSE
);
SYS.DBMS_OUTPUT.PUT_LINE('Job Number is: ' || to_char(x));
COMMIT;
END;
```

Fuente: Elaboración propia

Los procedimientos creados mostrados en las figuras anteriores servirán ejecutar el proceso ETL. La función del procedimiento es primero limpiar por completo las dimensiones ya que algunos registros pueden sufrir modificaciones, una vez ejecutado el proceso de limpieza se procederá a poblar las dimensiones del Data mart. Finalmente, en la FACT\_DESEMBOLSOS Y FACT\_MORA se procederá solo a insertar la información y no a eliminar, ya que esa información nos servirá de manera histórica. Una vez ejecutado todos los procedimientos se creó un JOB para que de manera mensual realice la limpieza de las dimensiones.

Ya terminado el proceso de ETL se procederá a Actualizar el cubo OLAP en la herramienta Data tools for Visual Studio.

## 6. Especificación de aplicaciones BI

### Informes estándar:

Para que la información sea explotada de la manera óptima se definió que para mostrar los reportes y análisis de información para la presente investigación.

### Herramientas de apoyo para explotar la información del Data mart:

Para poder explotar la información, se presentó 3 opciones de acuerdo con el tipo de rol de los usuarios interesados para que puedan visualizar la información de manera óptima y sencilla lo cual podemos observar en la siguiente tabla N°12.

**Tabla 12 Herramientas de apoyo para explotar la información del Data mart**

Herramienta	Rol	Descripción
Sistema web	Gerentes / Analistas	Para los informes gráficos o reportes a generar de manera automática y de fácil manejo.
Informes personalizados en Excel	Especialistas y Analistas	Los especialistas o Analistas que requieran el acceso al Data mart tendrán que pedir un usuario y contraseña para poder visualizar y explotar toda la información de acuerdo a la necesidad.
Informes personalizados en Reporting Services	Analistas de BI	Para los analistas de BI que deseen explotar la información de acuerdo a la necesidad.

Fuente: Elaboración propia

## 7. Desarrollo de Aplicaciones BI

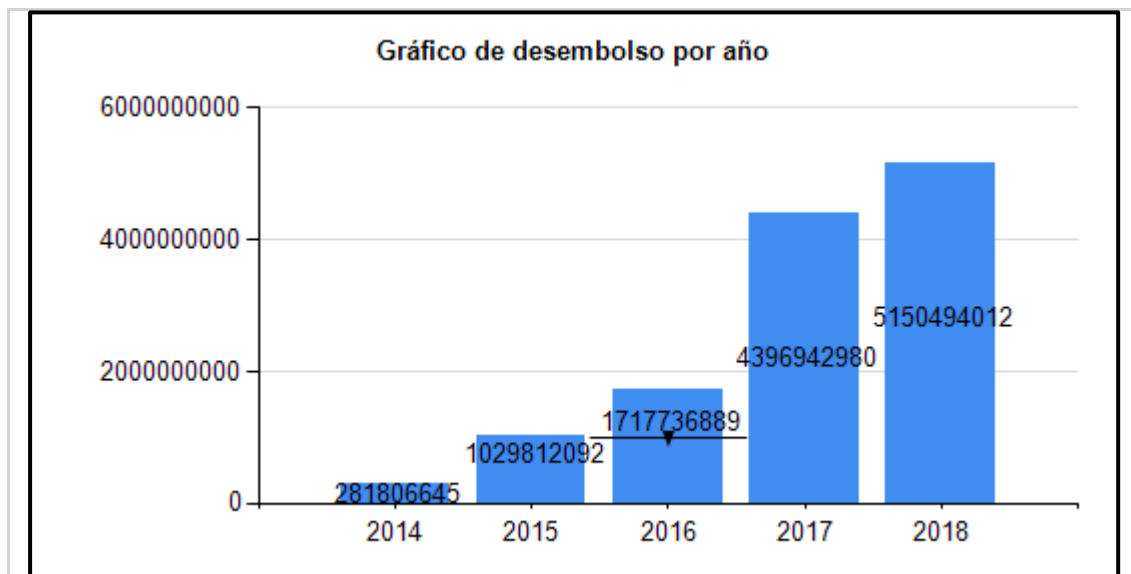
Con la finalidad que los usuarios puedan explotar la información del Data mart se decidió elegir la herramienta de apoyo Reporting Services, Sistema web y Excel, los cuales son de fácil manejo y sencillos para la generación de gráficos dinámicos como reportes.

En la figura 29 se muestra los desembolsos realizados desde el 2014 a la fecha

Figura 29 Desembolso por año

## Reporte de desembolso por año

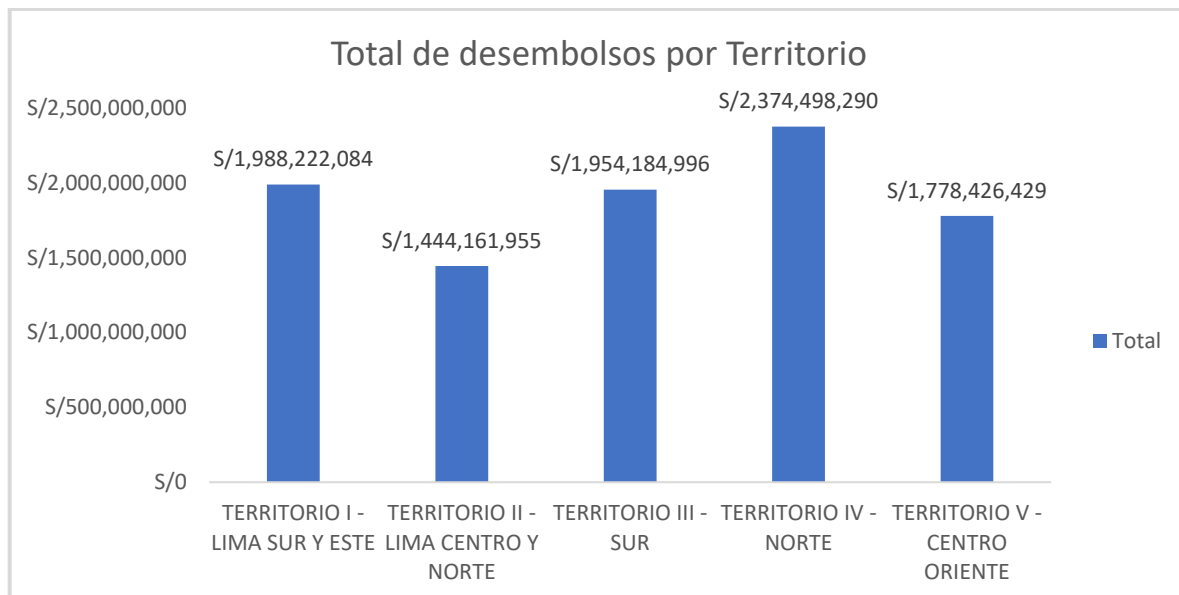
AÑO	MONTO TOTAL DESEMBOLSADO
2014	281806645
2015	1029812092
2016	1717736889
2017	4396942980
2018	5150494012



Fuente: Elaboración propia

En la figura 30 se muestra los montos desembolsados en total por cada territorio

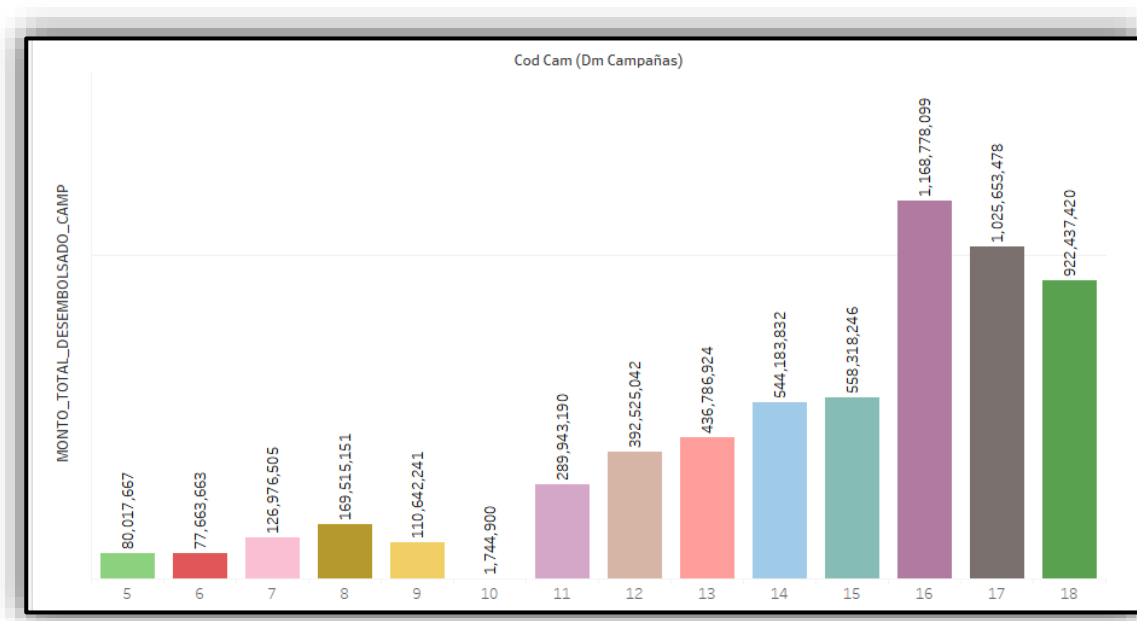
**Figura 30 desembolso por territorio**



Fuente: Elaboración propia

En la figura 31 se observa la cantidad de dinero desembolsado por cada campaña en Mibanco.

**Figura 31 desembolso por Campaña**



Fuente: Elaboración propia

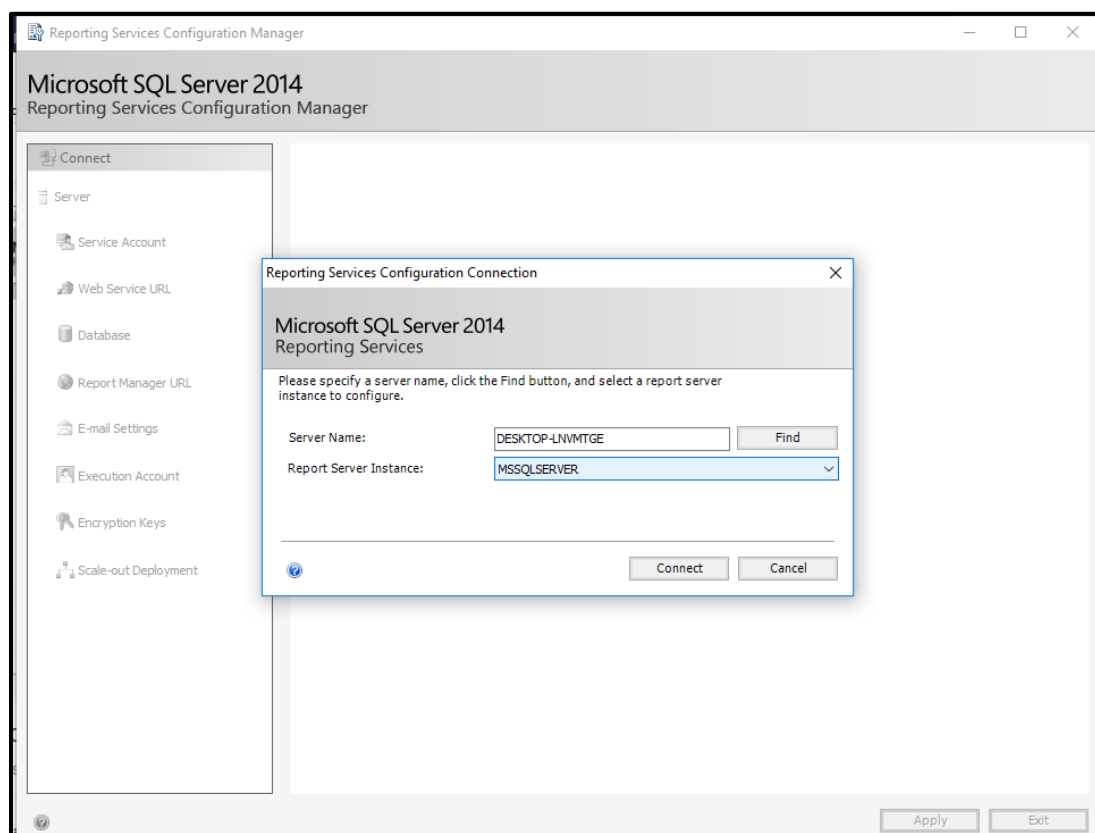
## 8. Implementación

### Despliegue

### Configuración del Reporting Services

En la figura 32 se visualiza el inicio de sesión a la configuración de Reporting Services.

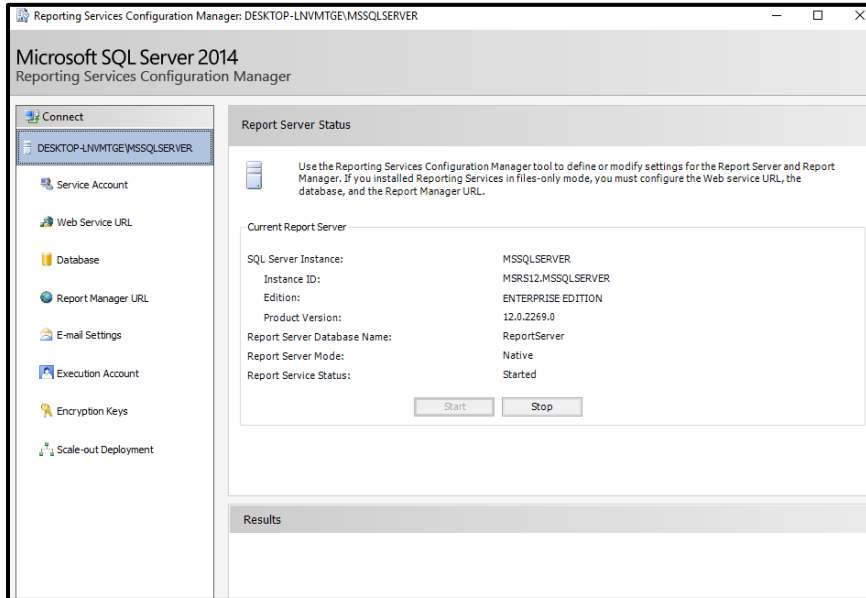
**Figura 32 Inicio de sesión al servidor Reporting Services**



Fuente: Elaboración propia

En la figura 33 se visualiza el estado del servidor de informes, se debe iniciar el servidor.

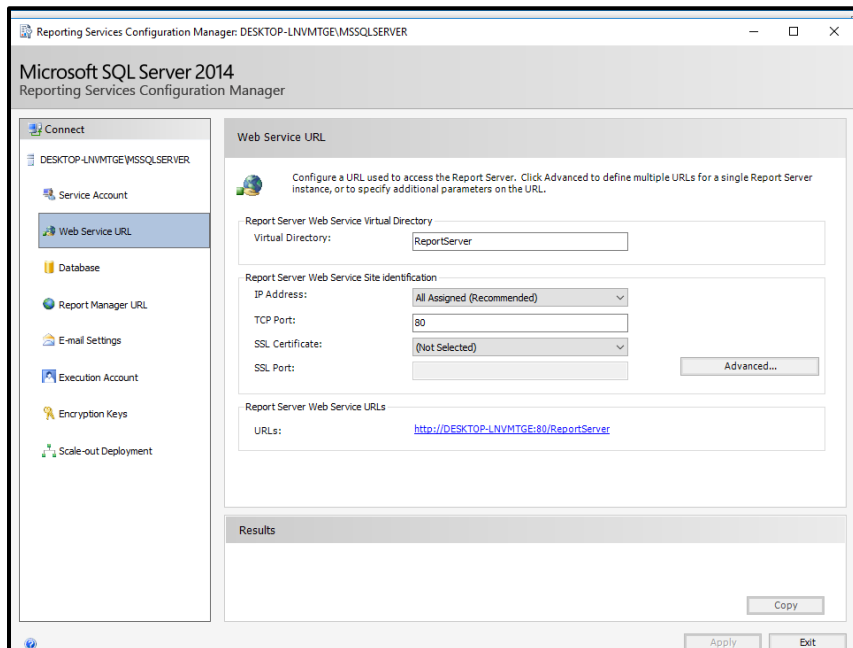
**Figura 33 Datos del servidor Reporting Services**



Fuente: Elaboración propia

En la figura 34 se visualiza la configuración del directorio virtual, o el repositorio de informes donde se aloja.

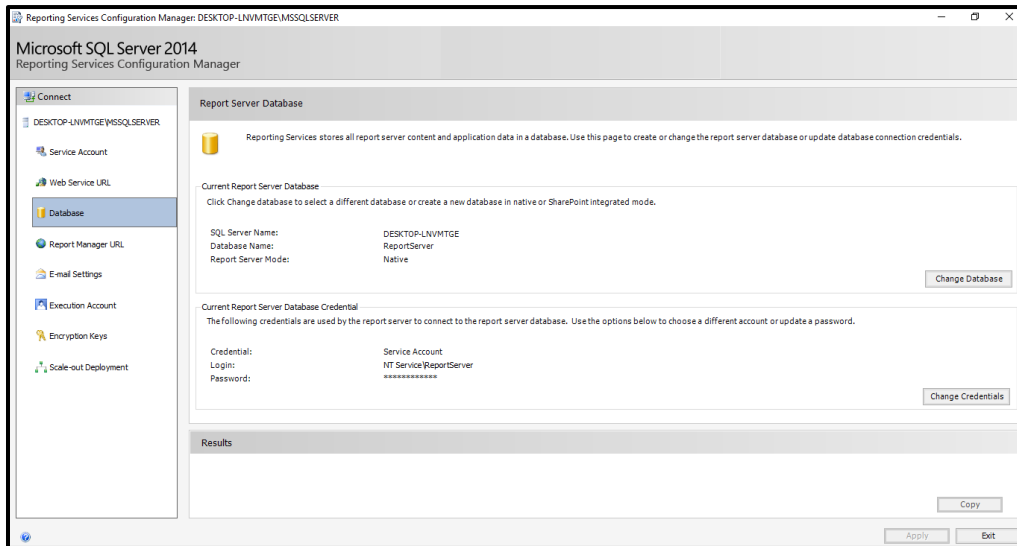
**Figura 34 Directorio virtual donde se publicaran los reportes**



Fuente: Elaboración propia

En la figura 35 se visualiza la configuración del servidor de base de datos que usara el servidor Reporting Services.

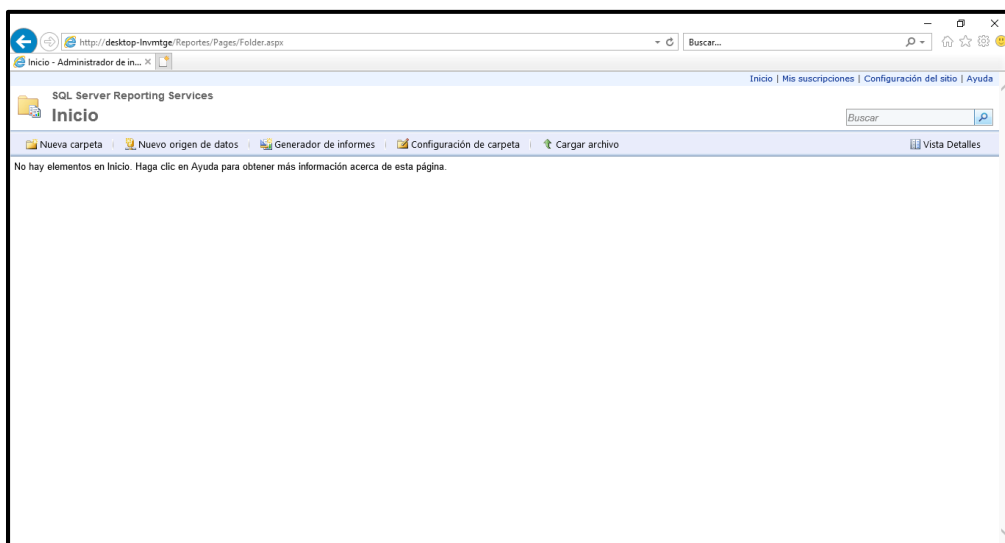
**Figura 35 Directorio virtual donde estará publicado los reportes**



Fuente: Elaboración propia

En la figura 36 se muestra el servidor donde se almacenará y publicaran los reportes, generando la conexión al data mart para la explotación de la información.

**Figura 36 Directorio virtual donde se almacenarán los reportes y donde se crearán las conexiones**

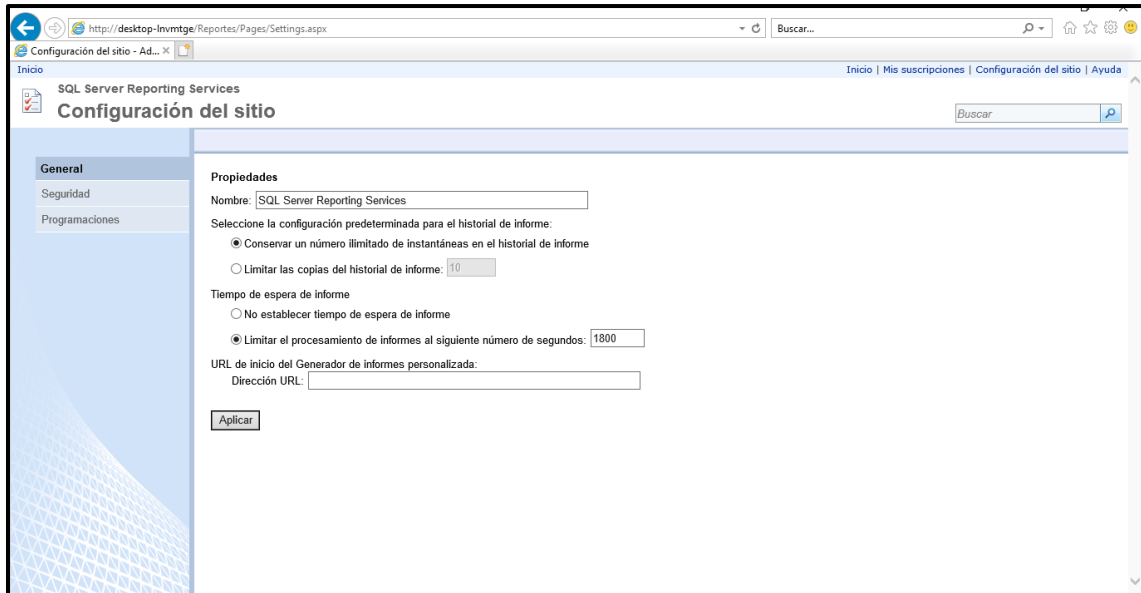


Fuente: Elaboración propia



En la figura 37 se muestra la configuración a realizar para el funcionamiento del servidor Reporting Services

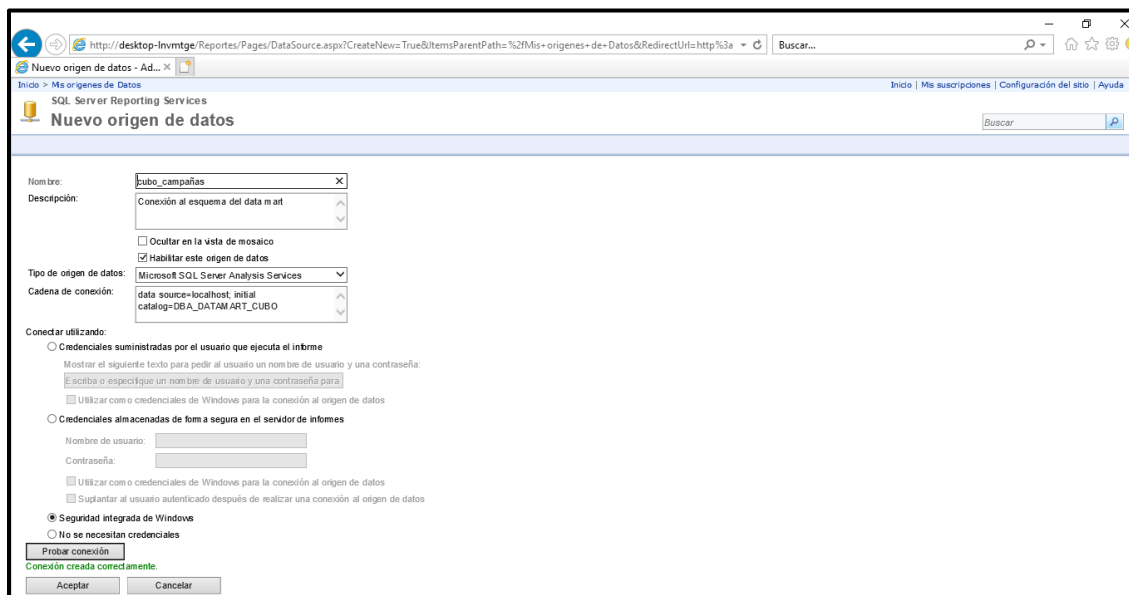
**Figura 37 Configuración del sitio**



**Fuente: Elaboración propia**

En la figura 38 se muestra la configuración para poder acceder al data mart, generando una cadena de conexión y probando la conectividad.

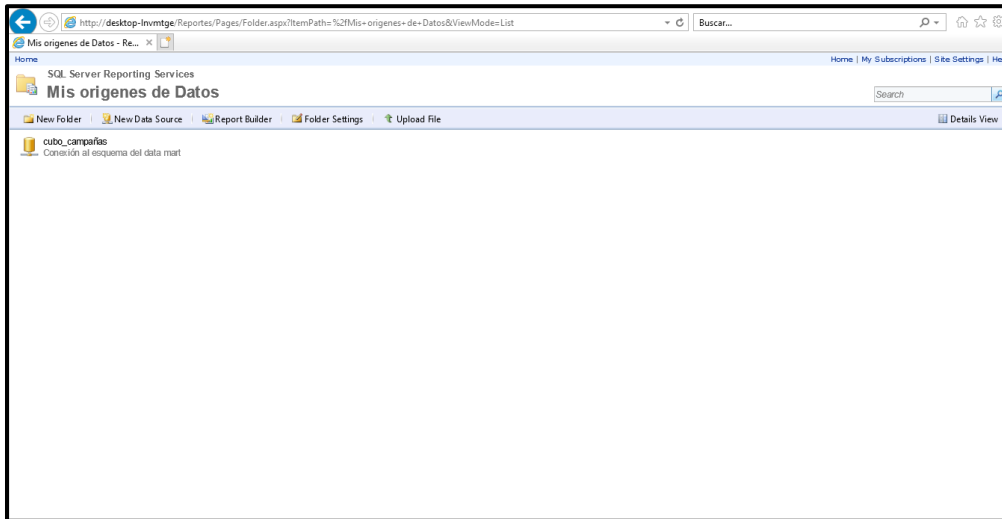
**Figura 38 Creación del origen de datos**



**Fuente: Elaboración propia**

En la figura 39 muestra cómo se llegó a conectar en Reporting Services con el Datamart, listo para la generación de reportes.

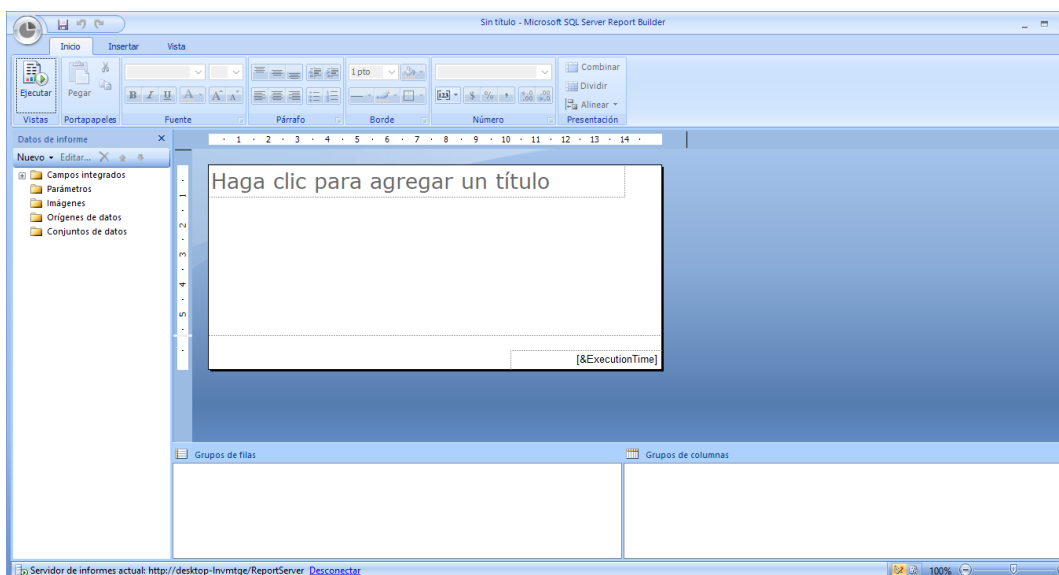
**Figura 39 Conexión del data mart**



Fuente: Elaboración propia

En la figura 40 se muestra la herramienta para la generación de informes, así poder publicarlos en el servidor del Reporting Services.

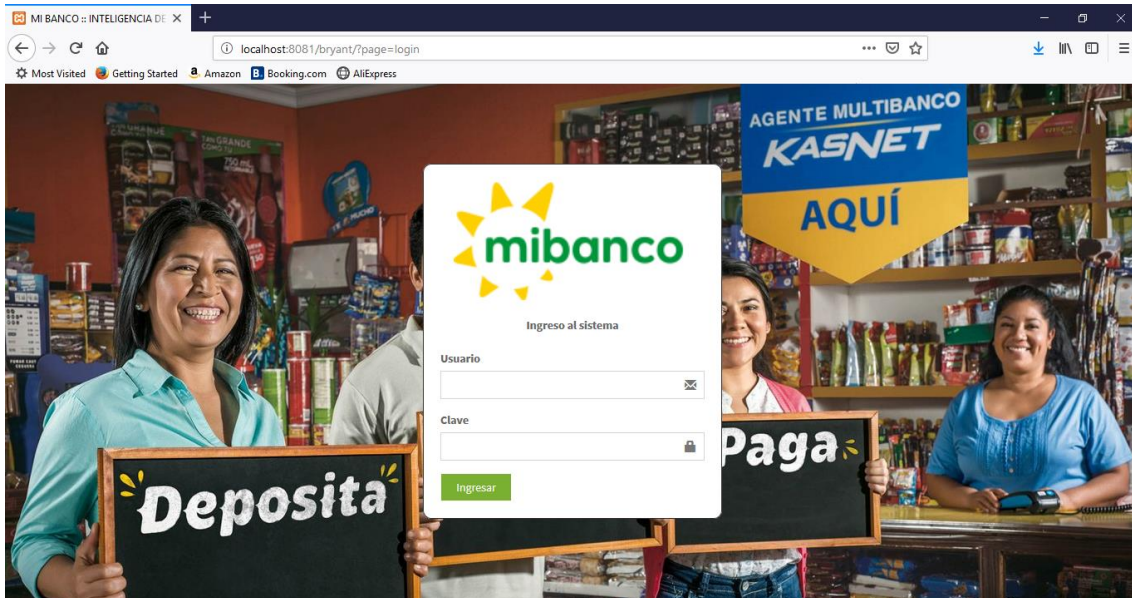
**Figura 40 Generador de informes**



Fuente: Elaboración propia

Una vez configurado el servidor, se procede con el loggeo al sistema implementado para la explotación del Data mart.

**Figura 41 Loggeo al sistema de Mibanco**

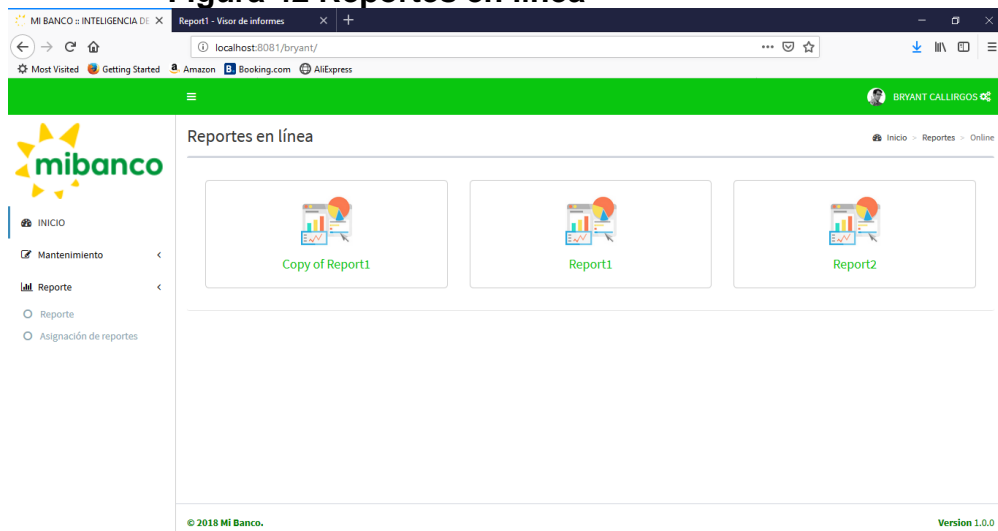


**Fuente: Elaboración propia**

Primero el Usuario se loggeara con el ID brindado, seguidamente ingresara su contraseña.

Una vez dentro del sistema, se procederá a ver los reportes ya creados, se manipularán de manera dinámica y a tiempo real.

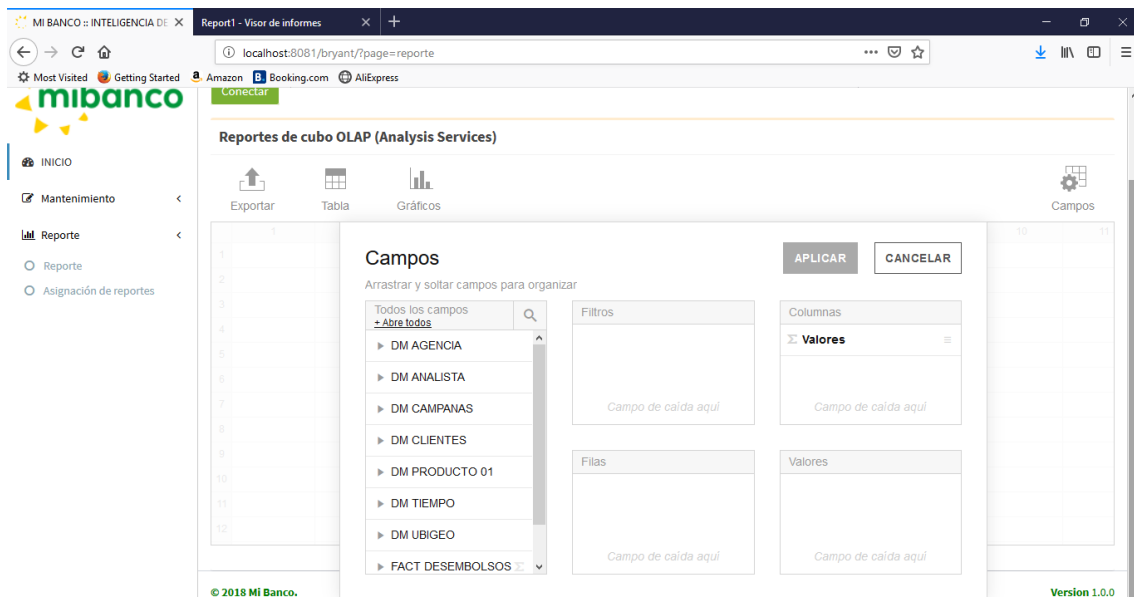
**Figura 42 Reportes en línea**



**Fuente: Elaboración propia**

Así mismo se podrá crear otros reportes dentro del mismo sistema, para generar mayor eficiencia y análisis de información de acuerdo a las necesidades.

**Figura 43 Creación de nuevos reportes**



**Fuente: Elaboración propia**

# Acta de Aprobación de Originalidad de Tesis

	<b>ACTA DE APROBACIÓN DE ORIGINALIDAD DE TESIS</b>	Código: F15A-PP-PR-01.04 Versión: 10 Fecha : 25/02/2019 Página : 1 de 1
---	--	--

Yo, Dr. ORDOÑEZ PEREZ ADILIO CHRISTIAN, docente de la Facultad de Ingeniería y Escuela Profesional de Ingeniería de Sistemas de la Universidad César Vallejo Lima Norte, revisor de la tesis titulada:

**DATAMART PARA EL PROCESO CREDITICIO EN EL ÁREA DE CAMPAÑAS COMERCIALES DE LA EMPRESA MIBANCO**

del estudiante CALLIGOS CARDENAS JEAN BRYANT, constato que la investigación tiene un índice de similitud del 30% verificable en el reporte de originalidad del programa Turnitin.

El suscrito analizó dicho reporte y concluyó que cada una de las coincidencias detectadas no constituyen plagio. A mi leal saber y entender la tesis cumple con todas las normas para el uso de citas y referencias establecidas por la Universidad Cesar Vallejo

Los Olivos, 05 de Noviembre del 2019.



Dr. ORDOÑEZ PEREZ ADILIO CHRISTIAN

Docente Asesor de Tesis

DNI: 40103357

# Pantallazo del Software Turnitin

Feedback Studio - Google Chrome  
ev.turnitin.com/app/cards/es/lang=es&oa=1040937300&u=1058650462&student\_user=183=

feedback studio

Jean Bryant CALLRIGOS CARDENAS DESARROLLO DE TESIS, CALLRIGOS2018

30%

So están viendo fuentes estándar

Ver fuentes en inglés (Beta)

Coincidencias

1	repositorio.ucv.edu.pe	9 %
2	pt.acribd.com	8 %
3	Entregado a Universida...	7 %
4	repositorio.uta.edu.ec	1 %
5	repositorio.upn.edu.pe	1 %
6	Entregado a Universida...	1 %

pt.acribd.com

UNIVERSIDAD CÉSAR VALLEJO

FACULTAD DE INGENIERÍA

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS

DATA MART PARA EL PROCESO CREDITICIO EN EL ÁREA DE  
CAMPAÑAS COMERCIALES DE LA EMPRESA MIBANCO

Página: 1 de 124 Número de palabras: 18177

Text-only Report | High Resolution

# Formulario de Autorización para la publicación de la Tesis

UNIVERSIDAD CÉSAR VALLEJO

Centro de Recursos para el Aprendizaje y la Investigación (CRAI)  
"César Acuña Peralta"

## FORMULARIO DE AUTORIZACIÓN PARA LA PUBLICACIÓN ELECTRÓNICA DE LAS TESIS

### 1. DATOS PERSONALES

Apellidos y Nombres: Callirgos Cardenas Jean Bryant  
D.N.I. : 72481689  
Domicilio : Mz Jv lote 1 urb los libertadores S.M.P  
Teléfono : Fijo : 015853413 Móvil : 991001743  
E-mail : jbcallirgos@gmail.com

### 2. IDENTIFICACIÓN DE LA TESIS

Modalidad:

Tesis de Pregrado

Facultad : Ingeniería  
Escuela : Ingeniería de Sistemas  
Carrera : Ingeniería de Sistemas  
Título : Ingeniero de Sistemas

Tesis de Post Grado

Maestría

Doctorado

Grado : .....  
Mención : .....

### 3. DATOS DE LA TESIS

Autor (es) Apellidos y Nombres:  
Callirgos Cardenas Jean Bryant

Título de la tesis:

Datamart para el proceso crediticio en el área de campañas comerciales de la empresa Mibanco

Año de publicación : 2019

### 4. AUTORIZACIÓN DE PUBLICACIÓN DE LA TESIS EN VERSIÓN ELECTRÓNICA:

A través del presente documento,

Si autorizo a publicar en texto completo mi tesis.



No autorizo a publicar en texto completo mi tesis.



Firma :  .....

Fecha: 17/12/2019



# Autorización de la Versión final del trabajo de Investigación



**UNIVERSIDAD CÉSAR VALLEJO**

## AUTORIZACIÓN DE LA VERSIÓN FINAL DEL TRABAJO DE INVESTIGACIÓN

CONSTE POR EL PRESENTE EL VISTO BUENO QUE OTORGA EL ENCARGADO DE INVESTIGACIÓN DE

La Escuela de Ingeniería de Sistemas

---

A LA VERSIÓN FINAL DEL TRABAJO DE INVESTIGACIÓN QUE PRESENTA:

Callirgos Cardenas Jean bryant

INFORME TITULADO:

Data mart para el proceso crediticio en el área de campañas comerciales de la empresa Mibanco

---

PARA OBTENER EL TÍTULO O GRADO DE:

---

Ingeniero de Sistemas

SUSTENTADO EN FECHA: 12/12/2018

NOTA O MENCIÓN: 15



FIRMA DEL ENCARGADO DE INVESTIGACIÓN