

Universidad CENFOTEC

Maestría en Tecnología de Base de Datos

Documento final de Proyecto de Investigación aplicada

Desarrollo de un prototipo de inteligencia de negocios para el departamento de crédito de la Junta-UCR sede Rodrigo Facio.

Estudiantes

Quirós Cedeño María

Coulon Ollivier Francisco

Enero, 2019

Declaratoria de derecho de autor

El presente proyecto de investigación aplicada fue realizado por el equipo de trabajo conformado por María Quirós Cedeño y Francisco Javier Coulon. El mismo contiene los resultados de indagación en Internet, literatura relacionada a la temática y los conocimientos adquiridos de cursos llevados durante el proceso de la Maestría, de los cuales, Inteligencia de negocios y Administración de bases de datos son la base.

Con respecto a la utilización del presente material, se solicita no hacer reproducción de cualquier índole (total o parcial) de este trabajo, por cualquier medio o procedimiento, debido al tipo de información que se estuvo manejando en el proyecto. Esto se especifica para los primeros cinco años, posterior a esto, el material es libre de ser utilizado para fines académicos para aquellos que lo requieran.

Agradecimientos

El presente trabajo investigativo lo dedicamos principalmente a Dios, por ser el inspirador y darnos fuerza para continuar en este proceso de obtener uno de los anhelos más deseados. Al mismo tiempo, queremos agradecer al profesor José Cabezas, por todo el tiempo dedicado al proyecto, la guía con respecto a las distintas fases del proyecto, especialmente en la segunda parte. De corazón, agradecemos toda la paciencia para con nosotros.

Tabla de contenido

Resumen ejecutivo 1

Capítulo primero. Introducción 2

Generalidades 2

Antecedentes del problema 2

Definición y descripción del problema 3

Justificación 4

Viabilidad 5

Punto de vista técnico. 5

Punto de vista operativo. 5

Punto de vista Económico. 5

Objetivos 5

Objetivo general. 6

Objetivos específicos 6

Alcances y limitaciones 6

Alcances. 6

Limitaciones. 7

Marco de Referencia Organizacional y Socioeconómico 8

Historia. 8

Tipo de negocio y mercado meta. 8

Misión 9

Visión. 9

Políticas Institucionales. 9

Estado de la cuestión 9

Planteamiento previo a la revisión sistemática. 9

Problema. 10
Pregunta de investigación. 10
Formulario de la pregunta. 11
Proceso de revisión/búsqueda. 11
Preparación de los términos utilizados en las búsquedas. 11
Palabras claves de acuerdo con la estructura PICOC. 11
Sinónimos de palabras clave. 12
Elaboración de la consulta base con cláusulas AND y OR. 12
Criterios de inclusión / exclusión. 13
Extracción de los artículos. 13
Lista de fuentes utilizadas. 13
Proceso de selección de artículos. 14
Detalle de los artículos seleccionados. 14
Artículo 1. 15
Artículo 2. 15
Artículo 3. 16
Artículo 4. 16
Artículo 5. 17
Artículo 6. 17
Capítulo segundo. Marco Teórico 18
Inteligencia de Negocios 18
Ventajas. 19
Desventajas. 19
Proceso para Inteligencia de Negocios. 21
Proceso según Beatriz Canales. 21

Dirigir y planear. 21
Recolección de información. 21
Procesamiento de datos. 21
Análisis y producción. 21
Difusión. 21
Proceso según Gangadharan. 22
Análisis. 22
Diseño. 22
Desarrollo. 22
Despliegue. 22
Evolución. 22
Inteligencia de negocios y Administración de procesos de negocio 23
Inteligencia de negocios y toma de decisiones 25
Claves para la toma de decisiones. 26
Atender la información relevante. 26
Alternativa creativa. 26
Consecuencia. 26
No tener miedo. 26
Evaluación de la Inteligencia de Negocios 27
Reportes tradicionales vs. Inteligencia de Negocios. 28
Criterios de evaluación de herramientas inteligencia de negocios. 29
Simplicidad y adopción. 29
Capacidades de la herramienta. 29
Adopción de la herramienta por parte de los usuarios. 30
Inteligencia de negocios y mejora de procesos 30

¿Qué es un proceso? 30
Inteligencia de negocios, conocimiento y aprendizaje organizacional 31
Patrocinio del más alto nivel. 32
Restricción de recursos 32
Comprensión de las reglas del negocio. 32
Calidad de datos existentes. 33
Justificación del retorno de la inversión. 33
Silos departamentales. 33
Calidad de datos. 33
Dimensiones de la calidad de datos. 34
Puntualidad. 34
Exactitud. 34
Completitud. 34
Consistencia. 34
Accesibilidad 34
Soluciones a problemas de rendimiento de base de datos. 35
Cursores. 35
Vistas. 36
Vistas materializadas o indexadas. 36
TSQL. 37
Vistas dinámicas de gestión. 37
Procedimientos almacenados. 39
Sp_executesql. 39
SP configure. 40
Lenguaje de definición de datos. 42

Rebuild index. 42
Reorganize Index. 43
Comando de Consola de bases de datos. 43
DBCC FreeProcCac. 44
DBCC FreeSystemCache. 44
DBCC FlushProcInDB. 44
DBCC DropCleanBuffer. 44
DBCC FreeSessionCache. 44
Herramientas de evaluación de bases de datos 44
Asesor de ajuste. 44
Ventajas de esta herramienta. 45
Perfilador SQL Server. 45
SQL Server Management Studio 45
Servicios de integración de SQL Server. 46
Definición. 46
Requisitos. 46
Escenarios de uso. 46
Servicios de análisis de SQL Server 47
Definición. 47
Cubos multidimensionales 47
Procesamiento analítico en línea multidimensional. 48
Procesamiento analítico en línea relacionado. 48
Procesamiento analítico en línea híbrido. 48
Capítulo tercero. Marco Metodológico 49
Tipo de investigación 49

Alcance investigativo 49
Enfoque 50
Diseño 51
Población y muestreo 51
Instrumentos de recolección de datos 52
Técnicas de análisis de información 52
Estrategia de desarrollo de la propuesta 53
Etapas del proyecto. 53
Etapa 1 – Análisis de la base de datos. 53
Etapa 2 –Diseño de cubo. 54
Etapa 3 – Implementación de la base de datos multidimensional. 54
Actividades principales de cada etapa. 54
Etapa Uno. 54
Etapa Dos. 54
Etapa tres. 54
Capítulo cuarto. Análisis del diagnóstico 56
Nivel de satisfacción y uso de los datos 56
Utilización de la tecnología. 57
Necesidad de datos 57
Tamaño de la base de datos 58
Arquitectura de datos. 58
Capítulo quinto. Propuesta de la Solución 60
Tecnologías a utilizar 60
SQL Server. 60
Servicios de integración de SQL Server. 60

Tableau. 60
Power BI. 60
Servicios de Análisis de SQL Server. 60
Diseño de la arquitectura de la solución 60
Diseño del área de Staging. 63
Conexiones. 63
Estructura de la base de datos. 63
Proceso de extracción, transformación y Carga. 65
Detalle del proceso de carga en el área de almacenamiento temporal. 65
Diseño del área de almacenamiento multidimensional. 66
Conexiones. 67
Estructura del almacén de datos. 67
Tabla de hechos. 67
Tabla de dimensión de préstamos. 68
Dimensión de fechas. 69
Tabla de dimensión de clientes. 69
Dimensiones de filtrado. 70
Proceso de extracción, transformación y Carga. 70
Diseño del ambiente analítico 74
Archivos de solución. 74
Configuración en formato tabular. 74
Dimensiones. 75
Tabla de hechos. 77
Traducciones. 78
Diseño de las visualizaciones 79

Visualizaciones con Tableau. 81
Resumen de operaciones. 81
Resumen de transacciones. 82
Resumen de monto de créditos. 82
Visualizaciones con PowerBI. 83
Resumen de operaciones. 83
Resumen de transacciones. 85
Resumen de monto de créditos. 86
Capítulo sexto. Conclusiones y Recomendaciones 87
Conclusiones 87
Recomendaciones 90
Capítulo séptimo. Reflexiones finales. 92
Respecto a los propósitos del proyecto 92
Respecto a la solución propuesta 92
Respecto a los retos del área 93
Capítulo octavo. Trabajos a futuro 94
Afinamiento de la base de datos 94
Minería de datos 94
Rendición de cuentas 95
Visualizaciones móviles 95
Futuras líneas de investigación 96
Referencias 97
Bibliografía 99
Anexos 101
Anexo A. Encuesta diagnóstica del departamento de contabilidad. 101

Intención. 101
Estructura de las preguntas. 101
Descripción de las preguntas. 101
Pregunta #1 ¿Cuál es la frecuencia en la que debes realizar reportes de actividades de créditos? 101
Pregunta #2 ¿Qué tan confortable es el acceder a los datos? 101
Pregunta #3 ¿Tienes dependencias para la extracción de datos? 101
Pregunta #4 ¿Qué tan actual requieren los datos? 102
Pregunta #5 ¿Qué tan necesario visualizan tener una base de datos analítica? 102
Pregunta #6 ¿Conocen acerca de PowerBI? 102
Pregunta #7 ¿Conocen acerca de Tableau? 102
Pregunta #8 ¿Conocen acerca de SSAS? 102
Pregunta #9 ¿Sienten que la base de datos es rápida para consultas? 102
Pregunta #10 ¿Requieren generar gráficos para representar los datos? 102
Resultados de la encuesta. 103
Anexo B. Diagrama de base de dato multidimensional. 106
Anexo C. Diagrama de base datos analítica. 107

Tabla de figuras

Figura 1 Flujo de trabajo para el estado de la cuestión (Fuente: Elaboración propia).-------14
Figura 2 Proceso de Inteligencia de Negocios según Wittschen (Tomada del sitio web gestiopolis)----------20
Figura 3 Ciclo de Vida de un sistema de Inteligencia de negocios (tomado de Gangadharan)----------22
Figura 4 Inteligencia de negocios según Kaptadata----------27
Figura 5 Planes de ejecución guardados (Fuente Elaboración Francisco Coulon) ----------39
Figura 6 Planes de ejecución guardados (Fuente Elaboración Francisco Coulon) ----------40
Figura 7 Salida del procedimiento sp_configure (Fuente elaboración Francisco Coulon) --41
Figura 8 Resultado de cambio de configuración de memoria con sp_configure (Fuente Elaboración Francisco Coulon) ----------42
Figura 9 Gráfico con la frecuencia de realización de reportes de actividades.----------57
Figura 10 Diseño de la arquitectura de la solución (Fuente: Elaboración propia). ----------62
Figura 11 Conexiones para el área de Staging. (Fuente: Elaboración propia).----------63
Figura 12 Objetos de la base de datos de Staging. (Fuente: Elaboración propia).----------64
Figura 13 Proceso de carga de staging. (Fuente: Elaboración propia).----------65
Figura 14 Especificación del proceso de carga de dimensión en el área de almacenamiento temporal. (Fuente: Elaboración propia). ----------66
Figura 15 Conexiones para el proceso de carga al almacén de datos. (Fuente: Elaboración propia).----------67
Figura 16 Carga de dimensión en la base de datos multidimensional. (Fuente: Elaboración propia).----------71
Figura 17 Proceso de Búsqueda para la carga de la base de datos multidimensional. (Fuente: Elaboración propia). ----------72
Figura 18 Proceso de verificación y carga en la base de datos multidimensional. (Fuente: Elaboración propia). ----------73
Figura 19 Archivos de solución para la parte analítica. (Fuente: Elaboración propia).------74
Figura 20 Servidor en formato tabular. (Fuente: Elaboración propia).----------74

Figura 21 Estructura de los registros dentro del proyecto analítico. (Fuente: Elaboración propia).----------75
Figura 22 dimensión de fechas para la solución. (Fuente: Elaboración propia).----------76
Figura 23 Tabla de hechos. (Fuente: Elaboración propia).----------77
Figura 24 Traducciones para la solución analítica con el uso del complemento BIDS. (Fuente: Elaboración propia). ----------78
Figura 25 Tablero resumen de operaciones. (Fuente: Elaboración propia). ----------81
Figura 26 Tablero resumen de transacciones. (Fuente: Elaboración propia). ----------82
Figura 27 Tablero resumen de los montos por tipo de transacción. (Fuente: Elaboración propia).----------83
Figura 28 Resumen de operaciones con PowerBI. (Fuente: Elaboración propia). ----------84
Figura 29 Resumen de transacciones usando PowerBI. (Fuente: Elaboración propia). -----85
Figura 30 Tablero resumen de monto de créditos utilizando PowerBI. (Fuente: Elaboración propia).----------86
Figura 31 Anexo 1. Diagrama de bases de datos multidimensional. ---------- 106
Figura 32 Anexo B. Diagrama de base de dato analítica. ---------- 107
Figura 33 Gráfico de la frecuencia con lo que se debe analizar reportes en el diagrama. - 103
Figura 34 Gráfico con respecto a lo confortable del acceso de datos.---------- 104
Figura 35 Gráfico con respecto a la necesidad de utilizar elementos gráficos para presentar datos.---------- 105

Resumen ejecutivo

Actualmente, las organizaciones requieren adaptarse rápidamente a los cambios negocios que ocurren de manera tan rápida, inesperada y constante, por lo que el buen uso de la inteligencia de negocios se convierte en un factor clave para tomar decisiones que le permitan seguir compitiendo. En el mercado existen herramientas para este tipo de objetivos, algunas de estas con precios elevados y otras son gratuitas. La utilización de una u otra dependerá de la naturaleza de lo que se plantea y de los recursos disponibles.

El proyecto se enfocó en la realización un prototipo de herramienta de inteligencia de negocios para el análisis de los créditos de la Junta-UCR. Este fue realizado en la Junta-UCR, la cual es una organización de índole financiera para un sector cerrado (los empleados de la Universidad de Costa Rica), ubicada en la sede central de la UCR, la Ciudad Universitaria Rodrigo Facio.

Este proyecto nació de la necesidad de análisis de la información que existe en el departamento y que es generada con las transacciones que ocurren todos los días en el mismo. Si bien existen esfuerzos que han buscado el logro de dicho objetivo, el sistema no genera todos los reportes necesarios, traduciéndose en mayor cantidad de tareas para los interesados en los informes de operaciones y reportes a nivel administrativo.

Con un prototipo y un análisis de la base de datos, se pretendió proponer una alternativa de solución a las necesidades de datos, los mismos considerados como cruciales en la eficiencia de procesos dentro del departamento.

Capítulo primero. Introducción

En la actualidad, las organizaciones basan sus tomas de decisión en datos, por lo que estos se convierten en la base para el funcionamiento efectivo de la empresa. Sin embargo, esta tarea se hace más compleja, principalmente debido a los avances de las tecnologías de información, las cuales, aparte de la variedad de opciones que han surgido, permiten a las organizaciones manejar grandes volúmenes de datos (Twun, 2013, p. 1).

Al momento de aplicar inteligencia de negocios en una organización, surgen una serie de beneficios que son tangibles e intangibles. Uno de los aspectos más relevantes es el hecho de poder contar con información de mejor calidad, ya que un proyecto de esta índole requiere que se revise los datos existentes y mejorarlos, si fuese el caso, antes de continuar con el proceso. Además que se logra tomadas de decisión mucho más efectivas y fundamentadas.

Este documento es el resultado del proyecto. En el mismo se incluye un análisis de la situación actual del departamento en cuestión, se brinda una reseña general de la organización. Además de ello, se analizan la solución propuesta y los resultados obtenidos.

Generalidades

En el mercado hay muchas herramientas para la inteligencia de negocios, algunas con precios elevados y otras son gratuitas. Estas se utilizan para la obtención rápida y sencilla de datos provenientes de los sistemas de gestión empresarial, para así poder realizar un análisis de dichos datos de una forma más rápida e eficiente.

Este análisis es aprovechado por las organizaciones para la toma de decisiones de manera rápida y con un nivel de seguridad de que la dirección que se tome sea la correcta.

Antecedentes del problema

Este proyecto nació de la necesidad del departamento de informática de la Junta-UCR sede UCR sede Rodrigo Facio, en el cual se ha observado un crecimiento importante en los últimos años, a nivel de operaciones y de responsabilidades. Todo lo anterior debido

a la expansión de actividades del mismo. Esto implica que aspectos como la forma en que se toman las decisiones sean basados en criterios fundamentados con datos y se deje de lado el estilo "corazonada" para tales efectos.

Cuando se hace mención de "corazonada", según la Real Academia Española (RAE), se hace referencia al "impulso espontáneo con que alguien se mueve a ejecutar algo arriesgado y difícil" (Real Academia Española, 2014, 23º ed.). Esto es precisamente lo que se busca evitar con el planteamiento de este proyecto.

Al haber más responsabilidades y, por ende, más tareas a completar, se requiere que los tiempos de respuestas y los procesos sean optimizados, lo cual no está al día de hoy. Lo anterior basado en la gran cantidad de trabajo manual que se realiza actualmente.

Otra de las situaciones problemáticas que existen es el hecho de que el sistema no fue programado internamente, sino que fue por medio de una subcontratación. De aquí surgió la estructura y la forma actual de esta base de datos. Esto implica que la mantenibilidad sea de alta complejidad y junto con la falta de tiempo el sistema ha caído en un nivel de rendimiento que se encuentra: por debajo de las expectativas de los distintos usuarios, debido a que la Junta ha trabajado dicha base de datos desde el año 2014, pero se han presentado problemas como que el sistema se cae, por lo que hay procesos que duran mucho y el administrador de la base de datos de la institución no ha tenido el tiempo suficiente para revisar a detalle los índices y demás aspectos.

Definición y descripción del problema

La Junta ha trabajado con una base de datos desde el año 2014, la cual fue realizada por un grupo externo e independiente a la organización y no cuenta con los procesos de mantenimiento óptimos debido al tiempo disponible y su estructura.

Además, del aspecto de que la base de datos fuese creada por otra empresa, esta no tiene diccionario de datos, ni cuenta con diagramas o algún tipo de información que facilite el entendimiento de la misma. La Junta siente la necesidad de revisar dicha base de datos, ya que hay procesos que duran demasiado tiempo, como por ejemplo los del cierre de mes.

También, la organización no cuenta con una herramienta que le ayude al análisis de los datos actuales. Por consecuencia, existen ocasiones en las cuales se han tomado decisiones equivocadas, lo cual ha significado pérdida de dinero. Existe falta de algunos reportes importantes en el "corazón" bancario, por lo cual se solicita constantemente información al departamento de tecnologías de información. El mismo debe de generar directamente de la base de datos. Por lo tanto, con la creación de una herramienta se solventaría la creación de reportes y se ayudaría al departamento financiero para llevar a cabo el análisis de la información con la que cuentan.

Actualmente, la Junta trabaja algunos procesos utilizando documentos de Microsoft Excel, los cuales son generados por el departamento de tecnologías de información y son extraídos de la base de datos transaccional. Estos reportes son importantes para las actividades diarias tales como lo son las evaluaciones a la morosidad, cantidad de créditos, cantidad de créditos cancelados entre otros

SIBU es el sistema principal de la Junta. En él se realizan procesos de tamaño importante; además, este contiene muchos registros, debido a las transacciones que se hacen diariamente con la población de casi 10 mil afiliados. Sin embargo, el consumo en tiempo para algunos procesos toma alrededor de dos horas, por lo cual se considera necesario hacer un análisis del diseño de la SIBUERP (nombre de la base de datos) para ver el tipo de índices que existen, como se está utilizando la base de datos y mejorar el proceso de extracción de datos de la misma.

Justificación

En la actualidad, si bien existen esfuerzos que buscan, a partir de los datos, generar elementos que ayuden a la toma de decisión, el sistema no genera todos los reportes. Lo anterior también se debe que existen reportes que son estilo "ad-hoc", que se realizan en una sola ocasión para algún tema en específico o porque alguna de las auditorías lo solicita, de allí que la justificación del proyecto no solo busca que el departamento de crédito genere sus propios informes, sino para que el departamento de tecnologías de información se enfoque de lleno en sus proyectos.

Con este proyecto se permitiría a la organización contar con un sistema que facilite, en gran medida, la toma de decisiones relacionadas con crédito, además de que se tenga más claridad acerca de la situación actual del negocio y al mismo tiempo que acelere los tiempos de respuesta en la generación de reportes requeridos por los distintos usuarios.

Viabilidad

Se consideró un proyecto viable, primeramente, porque es una necesidad existente de la institución, en cuanto a la gestión de los datos y manejo de la información para la toma de decisiones.

Punto de vista técnico. Técnicamente el proyecto se consideró viable debido a que las tecnologías que existen en la actualidad se adaptan a soluciones de inteligencia de negocios. Además, por su naturaleza (prototipo), no se estaría afectando ningún proceso actual de la organización hasta que se pueda concretar su implementación.

Punto de vista operativo. Esta herramienta sería utilizada por los compañeros que conforman el departamento de crédito, los miembros de la Junta Directiva y el Gerente General de la Junta. Además, ellos pueden guiar al grupo durante la planeación, ejecución y evaluación de los distintos artefactos a generar en el proceso.

Punto de vista Económico. Desde el punto de vista económico, no es necesario contar con un servidor adicional, ha como se pudo probar durante la ejecución del prototipo. La Junta para el año 2018 presupuestó recursos, los cuales ayudaría a los compañeros en la utilización de las distintas herramientas, incluyendo la prueba de concepto.

Objetivos

Se seleccionó la taxonomía de Benjamín Bloom, en su versión original (1956), para el planteo de los objetivos, debido a que es la taxonomía cognitiva que representa de mejor manera los distintos elementos que conforman la finalidad del proyecto que se propone en

este documento. Asimismo, se considera esta taxonomía como apropiada para un proyecto de tecnologías en bases de datos en general.

Objetivo general. Desarrollar un prototipo de inteligencia de negocios para el departamento de crédito de la Junta-UCR sede Rodrigo Facio.

Objetivos específicos. A continuación se presentan los objetivos específicos, los cuales se enfocan en dos partes importantes de la investigación: la revisión del estado actual de la base de datos y su compatibilidad con el plan de proyecto y la elaboración de la solución de inteligencia de negocios.

1. Revisar el estado actual de los objetos de la base de datos existente que permitan la detección de oportunidades de mejora en su estructura y los pasos a seguir para su implementación.
2. Documentar los objetos de datos existentes que sean compatibles con la solución de inteligencia de negocios planteada.
3. Elaborar la documentación base para la solución de inteligencia de negocios para que pueda ser utilizado en toma de decisiones, reestructuraciones y solución de problemas de manera eficiente.
4. Desarrollar un cubo multidimensional que permita dar respuestas a múltiples consultas planteadas por los usuarios.
5. Evaluar el prototipo elaborado en un ambiente de desarrollo, con datos reales de la organización, para medir el nivel de efectividad y el rendimiento.

Alcances y limitaciones

Alcances. El proyecto estuvo enfocado en la implementación de un prototipo de una herramienta de inteligencia de negocios para poder ayudar en que las tomas de decisiones sean más analíticas que por intuición.

Este proceso fue realizado en el departamento de tecnologías de información, en coordinación con el departamento de créditos y sus dependencias, de acuerdo con sus

funciones. Específicamente créditos de la Junta-UCR la cual queda ubicada en la sede central de la Universidad de Costa Rica.

Se buscó también que los usuarios del sistema, tuvieran una forma de visualización de los diversos resultados o consultas que el sistema va a soportar. Esto se utilizó tecnologías como Power y Tableau para su elaboración.

También, la base de datos al estar completamente en inglés, sus objetos pueden convertirse en limitaciones de uso para los diversos usuarios, durante las visualizaciones, por ende se ha incluido la parte de traducciones, para poner los componentes extraídos en español de Costa Rica.

Limitaciones. Siguiendo con la sección anterior de alcances del proyecto, se consideró necesario aclarar elementos que no están comprendidos en el proyecto como los que a continuación se explicitan:

- No se hizo la implementación en ambiente real (producción) del departamento. El esfuerzo estuvo orientado en la prueba de concepto del cubo multidimensional.
- Solo la herramienta que presente los datos necesarios para la continuación de funciones.
- No se aplicó ningún tipo de auditoría en procesos de negocio ni en implementaciones de sistemas.
- No se consideraron aspectos de cambios en protocolos de seguridad, cambios en el otorgamiento de permisos del departamento.
- No se rediseñó la base de datos existente.
- No se llenó la base de datos multidimensional ni la analítica, solo se tomó una pequeña muestra de los datos totales para poder hacer la demostración del prototipo.

Marco de Referencia Organizacional y Socioeconómico

Historia. La Junta Administradora del Fondo de Ahorro y Préstamo de la Universidad de Costa Rica (JAFAP), fue creada con la promulgación de la Ley Orgánica de la Universidad de Costa Rica, el 26 de agosto de 1940 , debido a la necesidad de que sus funcionarios tuvieran acceso a un régimen de seguros al momento de jubilarse.

Posteriormente, con la Ley 4273 de la República, se le otorgaron entre otras funciones específicas las siguientes:

- Dirigir y administrar el fondo de ahorro y préstamo conforme con la reglamentación establecida por el Consejo Universitario.
- Constituir depósitos bancarios.
- Realizar inversiones con bonos del Estado y sus instituciones.
- Realizar operaciones de préstamos y descuentos a favor de los empleados de la Universidad de Costa Rica, con la simple garantía de su fondo patrimonial acumulado.
- Acepta las donaciones, herencias o legados que se hagan a su favor.
- Realizar todas aquellas otras funciones y actividades que para el buen cumplimiento de los fines estipulados, le encomiende el Consejo Universitario.

Tipo de negocio y mercado meta. La Junta es una financiera para un sector cerrado, los cuales son los empleados de la Universidad de Costa Rica; dicha financiera se dedica a brindar créditos con tasas más favorables que el mercado de afuera y también tiene una gama de ahorros con tasas de interés competentes con otras financieras.

No cuenta con un mercado meta, ya que es solamente para los funcionarios de la Universidad de Costa Rica, pero constantemente se revisan tasas tanto en los créditos como en los ahorros, para ser la primera opción de los afiliados del departamento.

Misión. "Contribuir con el mejoramiento del bienestar integral y la calidad de vida de las personas afiliadas." (Junta UCR, s.f.)

Visión. “Ser el principal aliado financiero de las personas afiliadas." (Junta UCR, s.f.)

Políticas Institucionales. La junta actualmente está en un proceso de documentar los procesos y realización de políticas, por lo cual se va a nombrar con las que cuenta actualmente:

- Política de Comunicación.
- Política Administrativa e informe de Gestión.
- Política Salarial.
- Política de Seguridad.
- Políticas de Tecnologías de Información.
- Política de Base de Datos.

Estado de la cuestión

Planteamiento previo a la revisión sistemática. El objetivo de esta revisión de literatura fue poder determinar técnicas, prácticas, maneras en los cuales las herramientas de tecnologías de inteligencia de negocios llegan a mejorar los procesos de negocio de una organización.

Para esta primera etapa se han seguido los siguientes pasos:

1. Identificar la necesidad de la revisión.
2. Definir la pregunta de investigación.
3. Desarrollar el protocolo de revisión.
4. Identificación de fuentes.
5. Selección de estudios (artículos, textos)
6. Medición de la calidad de los artículos encontrados.
7. Resumen de los datos encontrados.

Problema. En los últimos tiempos ha crecido el interés de las empresas por la información; se ha empezado a darle valor a toda la información que se obtiene de los clientes para poder darle a cada uno lo que busca, por lo que la manera de tomar decisiones cambio y por ello se necesita estas herramientas de inteligencia de negocios que ayudan a procesar la información con la que se cuenta.

Pregunta de investigación. Lo indicado anteriormente, conlleva a la siguiente pregunta ¿Cómo puede impactar las tecnologías de inteligencia de negocios en los procesos de negocio de una organización pública?

Se refiere "¿Cómo?" a las distintas actividades que estas herramientas pueden llegar a afectar. Por ejemplo

1. Actividades del personal de primera línea.
2. Actividades gerenciales.
3. Actividades para externos.
4. Toma de decisiones.

Formulario de la pregunta. Evidenciar el impacto de la inteligencia de negocios en los procesos actuales de toma de decisiones con los que cuenta las empresas.

Proceso de revisión/búsqueda. En esta subsección se muestra el proceso seguido para la revisión y búsqueda de los documentos que sirven como base de la investigación.

Preparación de los términos utilizados en las búsquedas. Tal y como lo recomienda Kitchenan (2007), las palabras clave de la pregunta se clasificaron de acuerdo con la estructura de PICOC (Población, Intervención, Contexto, Salidas y Comparación por sus términos en inglés). Sumado a lo anterior, la definición de sinónimos se hizo a través de sitios web como "Thesaurus" y el conocimiento de cada uno de los integrantes del equipo del proyecto.

Además, para la elaboración de la consulta, se realizó utilizando los términos clave obtenidos separados por la cláusula OR y cada elemento de la estructura PICOC se ha separado por la cláusula AND. Sin embargo a partir de esta consulta generada, se hicieron adaptaciones para los distintos motores utilizados debido a sus capacidades y formato requerido.

Palabras claves de acuerdo con la estructura PICOC. Estas se han traducido al inglés, debido a que será el modo de actuar de esta revisión. La pregunta quedaría de la siguiente manera:

"How business intelligence tools can impact the business processes in an organization?"

- Population: Business intelligence.
- Intervention: Business processes.
- Comparison: Process performance
- Outcome: Business intelligence tools
- Context: Organization

Sinónimos de palabras clave. Se muestra a continuación las palabras claves y sus sinónimos utilizados para la elaboración de las preguntas del proceso de búsqueda:

- Population: “Business intelligence”, “BI”
- Intervention: “Business processes”, “Process”, “Business case”
- Comparison: “Process performance”, “measure”, “impact”, “measuring”, “effect”
- Outcome: “Business intelligence tools”, “BI Tools”
- Context: “Organization”, “public organization”, “institution”, “system”, “government”

Elaboración de la consulta base con cláusulas AND y OR. La consulta base que se utilizó estuvo basada en las palabras clave y una combinación de las mismas; a estas se le hicieron adaptaciones para los distintos motores de búsqueda.

Para Google Académico, la consulta utilizada fue:

(((Business intelligence) OR BI) AND ((Business Process) OR (Process) OR (Business Case)) AND ((impact OR measure OR measuring OR effect OR effects OR (process performance)) AND (government OR (public organization) OR (institution) OR organization or system) AND (technology))

Usando el motor de búsqueda de la biblioteca digital de ACM, se utilizó la siguiente consulta:

recordAbstract:(+'Business +intelligence') AND content.ftsec:(BI 'Business Process' 'process performance' 'government' 'public organization' institution 'GBIS')

Por su parte, en el sitio Science Direct se utilizó la siguiente consulta:

(({Business intelligence} or {BI}) and ({Business Process} or {Business Case}) and ({impact} or {measure} or {performance}) and ({government} or {public organization} or {institution} or {organization} or {system}) and ({technology}))

Criterios de inclusión / exclusión. Se excluyeron artículos de acuerdo con el contenido del título y el resumen. No se adentra en el texto completo de cada uno. El proceso fue realizado por ambos integrantes del equipo del proyecto y, posterior a su realización, se hizo una revisión. Además de lo indicado anteriormente, se tomaron en cuenta las siguientes condiciones:

1. Deben se artículos en formato PDF y de acceso libre o abierto.
2. Artículos que no tengan el texto completo también se excluyen.
3. La fecha de creación no puede ser menor a 2009 por lo que debe ser documentos elaborados en los últimos años.
4. Se excluyen también lo que son libros y literatura gris.
5. Se excluyen documentos o estudios que son considerados como "lo mismo" de acuerdo a otro artículo (repetidos)
6. El resumen debe tener alguna de las palabras indicadas en las consulta.
7. La capacidad de poder dar respuesta a la pregunta de investigación.

Extracción de los artículos. Para la lectura de los artículos seleccionados y su correspondiente registro, se descargó el archivo .PDF, se registró el enlace de donde se extrajo y también se registró la fecha en la cual se procedió a su descarga; lo anterior especialmente para llevar a cabo el rastreo del proceso.

Lista de fuentes utilizadas. La presente revisión se realiza con base en las métricas presentadas por Scimago JR, la cual es una web especializada que muestra el nivel de confiabilidad que tienen las distintas fuentes que se utilizarán. Las seleccionadas estuvieron en un rango aceptable para el presente año 2018 e incluyendo métricas de años anteriores. En la siguiente lista de motores de búsqueda se hará la revisión sistemática:

- IEEE Xplore Digital Library.
- Google Scholar.
- ACM Digital Library.
- AAA Digital Library.
- Science Direct.

Proceso de selección de artículos. Por cada una de las fuentes de consultas seleccionadas, se siguió con el siguiente proceso:

1. Realizar la búsqueda con las cadenas previstas.
2. Revisar la cantidad de resultados y refinar la cadena (adaptar al motor de búsqueda)
3. ¿Existen resultados que sirven para responder la pregunta, con el título o resumen?
    a. Se guarda la fuente y el documento.
4. Se revisa el resumen con más detalle para ver si puede fue útil o no para el proceso.
5. Se repitió el proceso por los distintos documentos que aparecen.

El proceso seguido se indica de manera de diagrama de flujo de la Figura 1.

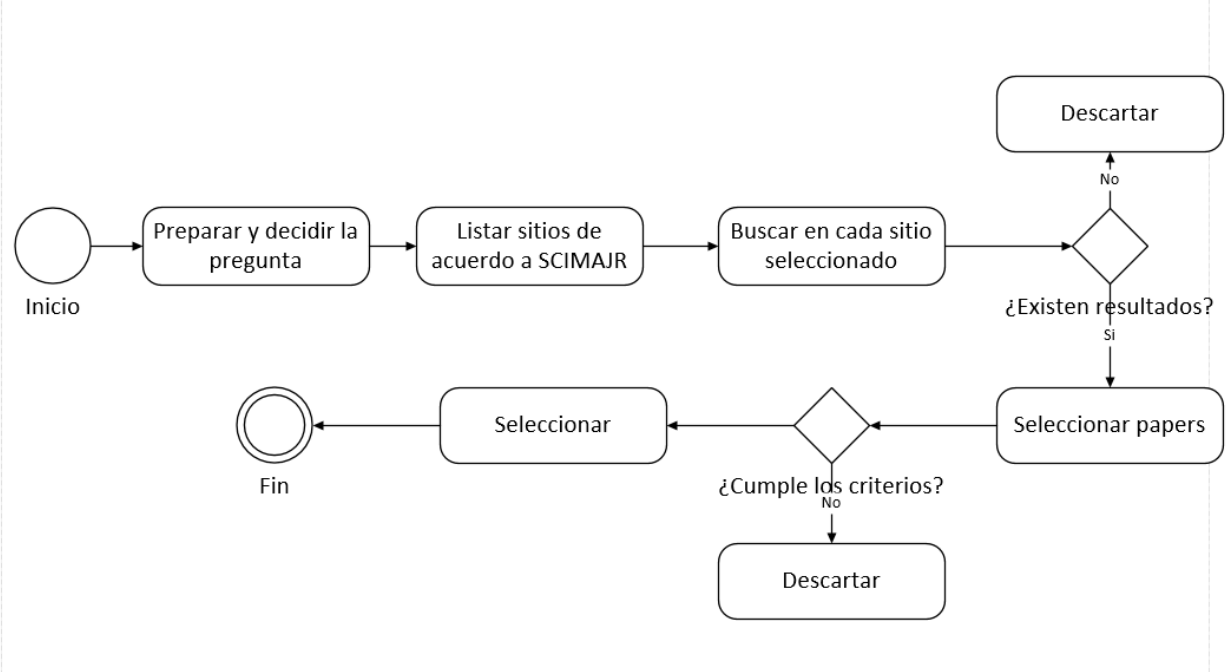


Figura 1 Flujo de trabajo para el estado de la cuestión (Fuente: Elaboración propia).

Detalle de los artículos seleccionados. Por cada uno de los artículos que se han seleccionado, se estructuró su resumen de la siguiente manera:

1. Título y fecha de publicación.
2. Autor o autores.

3. Aspectos relevantes: desde el punto de vista de la investigación y el planteo del proyecto de investigación aplicada.

Artículo 1.

Título y fecha de publicación

Business Value Creation through Business Processes Management and Operational Business Intelligence Integration (2010).

Autor/es: Olivera Marjonovic

Aspectos relevantes

Este artículo habla de estrategias orientadas al cliente; existe un volumen considerable de información que las empresas necesitan almacenar, por lo cual se crea la necesidad de herramientas que van a ser puestas en las manos de los tomadores de decisiones y con ellas los trabajadores van a tener más información de sus clientes.

La inteligencia de negocios cambia la naturaleza del proyecto para los participantes del proceso, dándoles mejores ideas y más oportunidades para hacer sus procesos operacionales conocimiento intensivo.

En algunos casos, las decisiones implican resultados deterministas, el desafío radica en los casos atípicos donde los resultados o incluso los procesos para alcanzarlos, pueden no ser conocidos.

Artículo 2.

Título y fecha de publicación

Using or Not Using Business Intelligence and Big Data for Strategic Management: An Empirical Study Based on Interviews with Executives in Various Sectors (2016)

Autor/es: Gökhan Silahtaroğlu y Nihat Alayoglu

Aspectos relevantes

Este artículo menciona acerca del porqué para las organizaciones resulta difícil el poder adoptar tecnologías de inteligencia de negocios. Hace una serie de entrevistas a

ejecutivos, quienes dan algunos detalles del porqué no se usan dichas tecnologías. Lo que resulta más útil para la investigación son aquellos elementos que deberían tener estas soluciones de inteligencia de negocios y que en ocasiones no la tienen y, por ende, alejan a los gerentes.

Artículo 3.

Título y fecha de publicación

A tool to evaluate the business intelligence of enterprise systems (2011)

Autor/es: M.Ghazanfari M.Jafari y S.Rouhani

Aspectos relevantes

Se muestra la importancia de tener una herramienta, técnica o modelo que permita evaluar y medir el nivel de aporte en inteligencia de negocios que tienen las herramientas existentes para la toma de decisiones de la organización.

El proceso de los seis factores puede resultar en una herramienta importante a la hora de medir la solución que se está proponiendo inicialmente para este proyecto de investigación aplicada. Igualmente puede servir para dirigir las actividades o esfuerzos planteados durante el desarrollo del proyecto.

Artículo 4.

Título y fecha de publicación

Framework Guidelines to measure the impact of business Intelligence and Decision Support Methodologies in the Public sector. (2016)

Autor/es: Roberto Boselli, Mirko Cesarini y Mario Mezzanzanica

Aspectos relevantes

Este artículo indica elementos claves en lo que es la inteligencia de negocios en el ámbito público. En el mismo se indica que la parte pública está atrasada con respecto a lo

privado en estos temas. Igualmente hace mención de posibles beneficios de implementar sistemas de soporte de decisiones tanto a nivel de procesos como de resultados.

Artículo 5.

Título y fecha de publicación

Effect of Business Intelligence and Analytics on Business Performance (2016).

Autor/es: Mohammad Daneshvar Kakhki y Prashant Palvia

Aspectos relevantes

Los autores buscan mostrar la relación entre la inteligencia de negocios y lo que es el performance de una empresa o del negocio. Buscan cómo demostrar una relación positiva entre ambos conceptos. Además utilizan datos de 116 organizaciones en la industria de la información.

Artículo 6.

Título y fecha de publicación

Business intelligence and organizational learning: An empirical investigation of value creation processes (2017)

Autor/es: Lior Fink, Nir Yogev y Adir Even

Aspectos relevantes

Este estudio busca relacionar el proceso de aprendizaje organizacional y la inteligencia de negocios; según los autores este ha sido un tema que no se ha manejado en conjunto o como un todo. El estudio busca maneras de cómo los componentes y capacidades de inteligencia de negocios pueden crear valor al negocio.

Capítulo segundo. Marco Teórico

A continuación se presenta el contenido conceptual de las diversas fuentes seleccionadas en el capítulo uno y que toma en cuenta diversas puntos de vistas publicados en sitios web. La información presentada se tomó como base para el proceso de organización, planeamiento de actividades, diseño e implementación de la solución realizada.

Inteligencia de Negocios

Según el sitio web Sinnexus "Business Intelligence es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios." (2018, s. f., párr. 1).

El término de inteligencia de negocios, es un concepto global en el que los sistemas de este tipo basan su integración y análisis de los recursos de datos organizacionales. Fink, Yogev y Even (2017).

Uno de los activos más importantes con los que cuenta una organización es la información, ya que se va recolectando con el tiempo y dependiendo de su uso podría traer muchos beneficios; lo que sucede con dicha información es que se requiere tiempo para poderla procesar además de conocimiento. La información es la clave para el éxito de cualquier empresa y ayudará a la toma de decisiones y está en las bases de datos que cuenta cada compañía.

En muchas compañías se da el caso que almacenan mucha información y tal vez dicha información está en muchas bases datos y hasta pueden ser de diferente proveedor de motor de base de datos. Por lo que para ganar ventaja o hacer la diferencia se requiere procesar dichos datos de forma inteligente. Por ejemplo, en situaciones donde se quiera saber si un producto o servicio tendrán éxito o cuál será la reacción de los afiliados tras su lanzamiento, si se gestionan de manera adecuada tanto los sistemas como las tecnología y sobre todo de la interpretación correcta de la información recabada. Por lo tanto, la gestión eficiente de la información, así como la inteligencia de negocios permite ampliar la visión estratégica, reducir el riesgo y la incertidumbre en la toma de decisiones empresariales y construir ventajas competitivas de largo plazo.

En el caso de la Junta, se cuenta con mucha información valiosa centralizada en una base de datos, lo cual es una gran ventaja. Sin embargo, el sistema base o corazón de la compañía no cuenta con todos los reportes para dicha información, por lo que estos son generados por el departamento de tecnologías de información, lo que hace que dichos reportes no todos los empleados los conozcan y por otro lado, también se genera mucho más trabajo para el departamento.

Ventajas. Como lo expone Wivo (2017), la inteligencia de negocios trae consigo ventajas como las mencionadas a continuación:

- Aumenta la productividad.
- Mejora las ventas.
- Acerca los resultados a las metas establecidas.
- Mejora el retorno de inversión.
- Conoce mejor el comportamiento de los clientes.
- Mejora la visibilidad y control.
- Convierte los datos en información útil y accionable.

Desventajas. No se pueden dejar de lado o subestimar las posibles desventajas que este concepto trae consigo; por ejemplo:

- Demasiado control al departamento de Tecnologías de información o usuarios.
- Ignorar las necesidades de los usuarios de estas soluciones de inteligencia de negocios.
- Subestimar los costos.
- Ignorar requerimientos futuros de este tipo de proyectos.

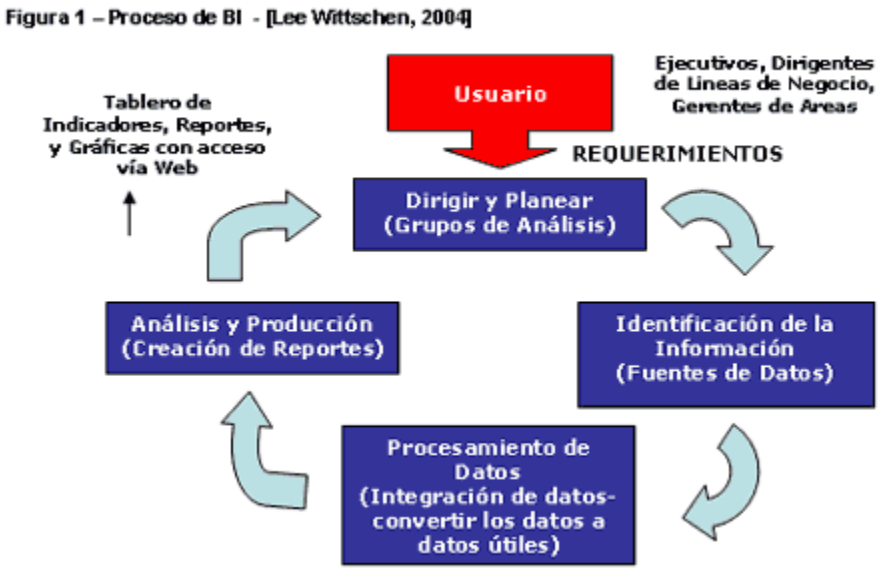


Figura 2 Proceso de Inteligencia de Negocios según Wittschen (Tomada del sitio web gestiopolis)

Proceso para Inteligencia de Negocios. La inteligencia de negocios es lo que se está utilizando para crecer y satisfacer a los clientes, pero cómo se implementa en las organizaciones, es un proceso de una serie de pasos como lo describe Beatriz Canales (2004) en su sitio web y que se resumen en los siguientes párrafos.

Proceso según Beatriz Canales. La autora divide la metodología de inteligencia de negocios en las siguientes etapas:

Dirigir y planear. Esta representa la fase 1 del proceso. Aquí es donde se hace la redacción de los requerimientos, de lo que se pretende alcanzar. Los implicados son los ejecutivos y líderes de negocios. Generalmente los objetivos vienen acompañados de preguntas que buscarán dirigir el proceso.

Recolección de información. Siguiendo los pasos del ciclo de vida de los datos, estos se convierten en información y luego son transformados en conocimiento. Sin embargo, para alcanzar este conocimiento se debe, en la mayoría de los casos iniciar con datos crudos. Estos son incompletos y confusos y también contienen partes no relevantes para lo que se definió en la fase pasada. De allí que, a partir de los diversos sistemas, se deben extraer datos, proceder a limpiarlos y determinar de qué manera llegan a ser útiles para lo que se pretende.

Procesamiento de datos. Es la tercera parte del proceso. Una vez que se tiene definido lo que se pretende hacer, y los conjuntos de datos identificados, se procede a realizar los diferentes proceso de extracción, transformación y carga para los diversos ambientes de inteligencia de negocios.

Análisis y producción. Este es el resultado final del proceso. Ya con la solución implementada entonces se procede a revisar que todo lo pretendido inicialmente se ha alcanzado con la solución. También es probable que se encuentre nuevas preguntas que requieren ser contestadas (nuevos requerimientos).

Difusión.

Esta es la última fase y consiste simplemente en hacer entrega de productos a los clientes que lo soliciten.

Proceso según Gangadharan. Por su parte Gangadharan, G. R., & Swami, S. N. (2004) dividen la metodología de inteligencia de negocios en las siguientes etapas:

Análisis. En la cual se hace la justificación del porqué se debe hacer, sus beneficios e indicadores de calidad. En esta parte se entiende que se debe hacer.

Diseño. Una vez que se tiene lo recopilado en la fase anterior se determina como se va a hacer, que elementos tiene. En esta parte, se busca hacer cortas interacciones (al estilo ágil) tal y como se muestra en la imagen siguiente

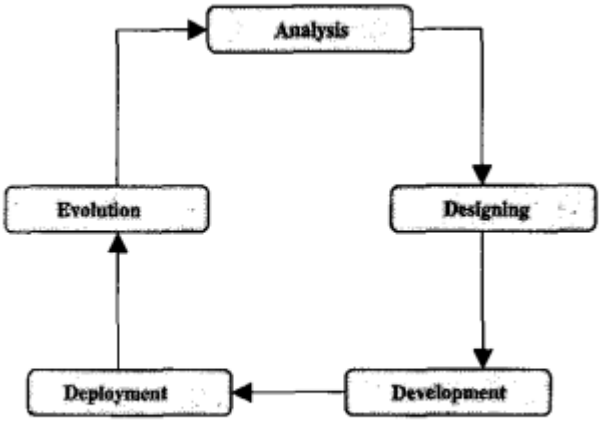


Figure 1. Life Cycle of BI System

Figura 3 Ciclo de Vida de un sistema de Inteligencia de negocios (tomado de Gangadharan)

Desarrollo. Teniendo el ¿Qué? Y el ¿Cómo? En esta etapa se pasa a la parte de desarrollo propiamente dicho. De acuerdo con lo definido y el presupuesto, así será el tiempo y el tipo de herramientas a utilizar.

Despliegue. Esta fase contiene el desarrollo completado y validado, en esta fase se hace el proceso de despliegue hacia los usuarios finales.

Evolución. El autor agrega una fase extra que es la evolución, la cual se convierte en el fin y a la vez inicio del proceso.

En caso de la Junta, lo que se implementó fue una herramienta ayude a la toma de decisiones en los procesos allí realizados, con la ayuda de diversas visualizaciones, por lo cual se trabajó en conjunto con el personal de inteligencia de negocios que indica qué información se requería para la elaboración de dicha herramienta.

Se utilizó la metodología presentada por Gangadharan, G. R., & Swami, S. N. (2004, June) para poder hacer del proyecto algo que tomara en cuenta las necesidades reales del usuario y que este último pueda ir observando el valor conforme se vaya desarrollando.

Inteligencia de negocios y Administración de procesos de negocio

La administración de procesos de negocio (BPM) es un enfoque estructurado que emplea métodos, policías, métricas, prácticas de administración y herramientas de software para manejar y optimizar las actividades de la organización y procesos (Marjanovic, 2010).

Esta definición se complementa con lo indicado en sitios webs como "conexionesan" (2015) el cual indica "El Business Performance Management, también conocido por las siglas BPM, es un conjunto de procesos analíticos y de gestión que permite a las empresas conocer su nivel de desempeño y saber qué medidas deben tomar para acercarse a sus objetivos" (párr. 2).

Se suele considerar al Business Performance Management (BPM) como un sistema que permite a las organizaciones recoger, añadir y analizar información proveniente de distintas fuentes para tomar las mejores medidas para la dirección del negocio

La conjunción de las herramientas de inteligencia de negocios y de gestión de rendimiento ha proporcionado un arma clave para que las organizaciones se adapten a las cambiantes condiciones del mercado y estén capacitadas para tomar decisiones de negocio rápidas, eficientes y eficaces.

Con los BPM se ha observado que este sistema trabaja con una serie de herramientas y con ellas se desarrolla un análisis FODA y este se encuentra

complementado por el análisis de las cinco fuerzas de Porter y esto es utilizado para saber y tener más conciencia de cómo está la industria.

La transformación de la inteligencia de negocios a la administración de procesos de negocio permite utilizar los activos de información de una empresa, de manera que se pueda incluir las funciones y herramientas de inteligencia de negocios. Con un proceso de administración de negocios (BPM) alineado a la inteligencia de negocios se pueden mejorar los tiempos de ejecución de las tareas de procesos y elimina los cuellos de botella en el proceso.

Como lo informa ITnow "El BPM mejora de manera continua el ciclo de vida de los procesos compuesto por la evaluación, el diseño, construcción, transformación, y operación para así identificar los aspectos del proceso que pueden ser mejorados." (ITnow, 2015, párr. 4)

Los datos son un elemento importante dentro este ciclo de mejora continua y para el control de los procesos; sin su medición no se podría controlar de forma eficaz y eficiente las actividades que integran los procesos.

Actualmente parece que las empresas están más maduras en cuanto a la inteligencia de negocios, pero necesita alcanzar el mismo grado de madurez con las soluciones de gestión de procesos de negocio, ya que el trabajo de estas dos herramientas juntas se puede alcanzar muchos logros.

Se puede pensar que no tienen nada en común estas dos herramientas, ya que la inteligencia de negocios y la administración de proyectos de negocio provienen de distintos puntos de vista organizacionales: la primera está centrada en datos, mientras que la segunda está centrada en los procesos; no existe una conexión entre ambas, pero los procesos necesitan distintos tipos de metodologías de mejora, que van mucho más allá del control de flujo, muchas veces utilizada para la mejora de procesos altamente estructurados y repetitivos. Aún más importante, los procesos de negocio ahora son considerados como la clave para determinar el valor de negocio de las aplicaciones de inteligencia de negocios.

Es necesario resolver los requerimientos de las empresas desde el punto de vista de la optimización del negocio, y no para solucionar problemas puntuales de la operación.

Algunas herramientas de administración de procesos de negocio cuentan con esta funcionalidad embebida y se conoce como monitorización de actividades de negocio (BAM por sus siglas en inglés), que mide el tiempo que tardan en ejecutarse las actividades de un proceso automatizado. Actualmente este monitoreo se lleva con herramientas de inteligencia de negocios independientes del BPMs que permiten conocer tanto los indicadores de gestión del negocio, como la actividad de sus procesos en tiempo real, consultando para ello directamente los datos resultantes de la gestión del proceso.

BPM por sí solo no es suficiente porque con él no se logra ver el contexto dinámico de los procesos. Al agregar inteligencia de negocios se puede tener una visión de la trayectoria del rendimiento, en donde los indicadores se comparan con los objetivos del negocio lo cual permite mejorar los procesos, las decisiones y los resultados.

Inteligencia de negocios y toma de decisiones

La inteligencia de negocios permite minimizar los riesgos en la toma de decisiones. La información y el análisis que estas técnicas proporcionan arrojan luz sobre el proceso de toma de decisiones. A través de un análisis interno de los datos existentes, la inteligencia de negocios trata de formar un conjunto de estrategias que otorguen información a las empresas.

Herramientas y concepto relacionados con sistemas de información tienen una importancia primordial cuando se trata de hacer planes estratégicos y decisiones en la vida de la organización. Silahtaroğlu, G., & Alayoglu, N. (2016).

Por otro lado el uso adecuado de la información suministrada por la inteligencia de negocios, brinda a la empresa herramientas para la toma de decisiones en muchas áreas de la misma.

Básicamente lo que hace esta metodología es estructurar datos que estaban desestructurados haciendo un análisis de información que de conocimiento efectivo a la empresa acerca de cómo marcha cada aspecto de esta.

Claves para la toma de decisiones. Como lo dice Ana Juana en su artículo "5 claves para la toma de decisión empresarial" para el sitio web Cultura Crm (2017) para tomar decisiones hay que tener:

Atender la información relevante. Sin una estrategia de análisis de datos no se puede ver ante un volumen de información demasiado amplio. Cuando se tiene datos en exceso, se puede perder la atención y obviar la información relevante y útil.

Alternativa creativa. Los datos pueden ser duros; al final a través de la inteligencia de negocios se descubre el camino correcto a seguir, pero la herramienta no dirá cómo hacerlo.

Consecuencia. El miedo de tomar una decisión empresarial es el riesgo que se asume y la responsabilidad si todo sale mal. Por esta razón, las empresas usan herramientas de inteligencia de negocios para reducir todos los riesgos.

No tener miedo. Cualquier directivo con una media responsabilidad tiene miedo a tomar una decisión y que no sea la correcta. Pero no debes tener miedo. Si trabajas con herramientas de inteligencia de negocios, la toma de decisiones está avalada por los datos y en ellos puedes respaldarte

Aplicando inteligencia de negocios, las Organizaciones pueden profundizar en los datos más importantes para convertirse en una compañía realmente guiada por datos y tomar decisiones informadas alrededor de mejores prácticas, utilizando indicadores de desempeño que son obtenidos de los sistemas transaccionales en operación dentro de la compañía."

Los sectores con una marcada estacionalidad a menudo tienen importantes problemas para lograr optimizar adecuadamente el inventario. Por ejemplo, si las ventas de un determinado producto se disparan en verano o durante las fiestas de navidad, resulta complicado almacenar la cantidad adecuada para maximizar los beneficios.

En este sentido, algunas empresas del sector conservero o de la alimentación en general han logrado aumentar la rentabilidad en cifras cercanas al 10% mediante técnicas de inteligencia de negocios basadas en la implementación de algún sistema de soporte de la decisión (DSS) y un análisis exhaustivo de los datos históricos de ventas e inventarios de productos almacenados.

Muchas empresas logran mejorar o aumentar sus ingresos y la rentabilidad gracias a las herramientas de inteligencia de negocios; no importa la índole de la empresa estas herramientas son útiles para cualquiera, ya que con ella se puede conocer lo que los clientes buscan. Los puntos de venta que podría ir mejor, que local sería mejor cerrar; en fin, tener una solución de estas es para tomar decisiones basándose en información estable y confiable.



Figura 4 Inteligencia de negocios según Kaptadata

Evaluación de la Inteligencia de Negocios

La definición incorrecta de los requerimientos del negocio o la pérdida de vista de los mismos puede causar que se elija un sistema de reportes que te presenta estadísticas no accionables. Esto ocurre cuando quienes toman las decisiones pierden la visión global de

las necesidades de la empresa y se dejan llevar por características o aspectos visuales de la herramienta que pueden no ser necesarias.

Para realizar esta evaluación en el ámbito público, hay que tomar en cuenta que el sector público es un sistema complejo Boselli, R., Cesarini, M., & Mezzanzanica, M. (2011) y que para lograr una correcta evaluación de la misma ha de tomarse en cuenta aspectos como objetivos, información, explotación del conocimiento y procesos de toma de decisión.

Reportes tradicionales vs. Inteligencia de Negocios. El artículo en el sitio web "Mundo Erp" indica que el sistema de reportes es estático. La información puede quedar obsoleta en el momento de disponer de ella, dado que no puede sufrir actualizaciones. Además, si en un momento, a partir de unos datos, se requiere profundizar más en esa información, implica generar nuevos informes que ratifiquen o desmientan ésta, lo que conlleva enormes esfuerzos adicionales. Los sistemas de reportes son herramientas reactivas, pues generan acción una vez se han completado los informes. Algunas de las consecuencias que generan estos tipos de reportes en una compañía son:

- Una pérdida de competitividad.
- Retraso en las actividades generadoras de valor.
- Limitaciones de la información.
- Escasa integración entre los datos.
- Información sin visiones globales.

Todas estas limitaciones afectan la toma de decisiones e implican una pérdida de competitividad y rentabilidad. Los sistemas de inteligencia de negocios vienen a solventar este tipo de limitantes, permitiendo una automatización en los procesos de obtención de la información y transformándola en conocimiento. Algunas de sus ventajas son:

- Reducción de costos.
- Gran interacción entre los datos.
- Alcance a todos los usuarios.
- Unificación en un sistema de toda la información de interés.
- Reducción de errores.

- Múltiples dimensiones y métricas.

Un sistema de inteligencia de negocios es altamente dinámico y ágil, con información en tiempo real (o casi), y con una gran integración entre los datos. En definitiva, la inteligencia de negocios es una herramienta "proactiva" y permite anticipar decisiones en el entorno de la organización.

Criterios de evaluación de herramientas inteligencia de negocios. ¿Qué debería de tomar en cuenta o evaluar una compañía, para ver si es funcional tener una herramienta de inteligencia de negocios? Según el sitio web "Storecheck", hay tres puntos clave a ser tomados en cuenta para la implementación de una solución de inteligencia de negocios, los cuales se resumen en:

Simplicidad y adopción. Se debe evaluar a través de períodos de prueba lo que las principales herramientas y tecnologías ofrecen, la facilidad de uso de la plataforma, para cualquier otro tipo de usuarios.

Algo común es el hecho de que la implementación del sistema puede requerir gran cantidad de recursos enfocados en la carga de los datos que la herramienta de inteligencia de negocios hará uso. Por lo que este tipo de herramientas deben de buscar la simplificación de este tipo de actividades.

Capacidades de la herramienta. Este punto se refiere a la capacidad de potenciar la información y análisis a todos los niveles de la compañía, es decir, no sólo funcionar como una simple base de datos. Al hablar de una herramienta de esta índole, no solo se habla de un sistema de información sino de un sistema integral para la toma de decisiones.

Cabe recordar que el objetivo principal es que la organización llegue a mejorar sus procesos en general. En la actualidad, es normal ver cómo las empresas usan herramientas de inteligencia de negocios para extraer datos, sin embargo al llegar al análisis, este se hace de manera manual lo que son reportes y/o gráficos.

Adopción de la herramienta por parte de los usuarios. Sucede también, como en los proyectos de software tradicionales, que los usuarios no adoptan la herramienta y esta, se convierte en un recurso subutilizado que puede ser percibido como una mala inversión.

Lo anterior puede ser generado por varias razones, entre ellas por la falta de capacitación o resistencia al cambio por parte de los usuarios. Resulta también tradicional que el nuevo sistema sea promocionado por figuras de autoridad dentro de la organización.

Inteligencia de negocios y mejora de procesos

¿Qué es un proceso? Proceso está definido como la sucesión de actos o acciones realizados con cierto orden, que se dirigen a un punto o finalidad, así como también al conjunto de fenómenos activos y organizados en el tiempo.

Todas las actividades de las empresas, desde la planificación del de compras, hasta la forma de gestionar un proceso de reclamo, pueden considerarse como procesos. Por lo tanto, para que una empresa pueda operar de manera eficaz es necesario que identifique y gestione numerosos procesos interrelacionados entre sí, de manera que dichos procesos se ayuden mutuamente.

El conocimiento se considera como una parte fundamental de los procesos de negocio. Se encuentra en documentos, modelos o repositorios formales, aunque también se encuentra en rutinas organizacionales, procesos y prácticas. El conocimiento organizacional incluye conocimiento explícito que puede ser externalizado, documentado, codificado, compartido dentro del mismo contexto y manejado por la tecnología. Esto también significa que el conocimiento, y especialmente su aspecto tácito, no es algo que pueda ser envasado y distribuido por la tecnología de manera de ser entregado a las personas indicadas en el momento indicado.

Se considera como una buena práctica dentro de los proyectos de inteligencia de negocios, darles un adecuado enfoque de procesos. Esta metodología permite conocer el negocio de tal manera que el producto final, los reportes e indicadores, cumplan con satisfacer las principales necesidades de información de la empresa y brindan adecuado soporte a la toma de decisiones.

La inteligencia de negocios nos va ayudar en el proceso:

- Convirtiendo los datos de los procesos en información estructurada, consolidada y de calidad.
- Obteniendo información en tiempo real.
- Reduciendo costes y alcanzando la mayor eficiencia.

Inteligencia de negocios, conocimiento y aprendizaje organizacional

Los directores de sistemas de información han de definir qué significa el éxito en términos de la inteligencia de negocios. Esto puede variar de manera subjetiva, estos encargados han de poder unir y combinar los gustos y necesidades de diversas fuerzas. Por ejemplo, hay individuos que el éxito de la inteligencia de negocios son los reportes, otros son las capacidades analíticas o la versión única de la verdad.

Esto de la inteligencia de negocios no es solamente tener un almacén de datos, sino también crear beneficios a partir de la implementación hecha. Si no se considera de esta manera, la inteligencia de negocios va a llegar a ser una simple plataforma y el valor pretendido por la inversión será difícil de poder ser demostrada.

Como no lo dice el sitio web INCAE: “Realizar cambios de cualquier índole genera ciertas dudas y pone a prueba la capacidad de los involucrados.” Implementar inteligencia de negocios como parte de la estrategia de la empresa implica considerar y superar una serie de retos que conllevan amenazas, pero también una puerta hacia grandes oportunidades de mejora.

Patrocinio del más alto nivel. Es importante, al igual que cualquier otro tipo de proyecto, tener apoyo de los altos mandos. Esto implica poder convencer acerca de la utilidad, involucrar a los que se consideran como fuerzas que amenazan la iniciativa.

Restricción de recursos. Se empieza con el proyecto de inteligencia de negocios y se da cuenta de que hay que capacitar gente, limpiar los datos y de pronto el proyecto es caro, implicaba cambio organizacional y no estaba planificado.

Comprensión de las reglas del negocio. Cuando se llega a implementar una solución de inteligencia de negocios, no se tiene claro las reglas del negocio. Esto ocasiona que a la hora de la implementación se tiene que trabajar con los procesos, y estos que no estaban bien documentados en conjunto con las reglas que no estaban claras hacen que el proyecto comience a crecer en complejidad y costos.

Calidad de datos existentes. La calidad de los datos es una amenaza. Alguna de ellas se debe a problemas de proceso de negocio. Otras es por la falta de control en el crecimiento de las aplicaciones empresariales. Además a nivel de auditoría, no todas las empresas se sienten cómodas registrando todas sus operaciones.

Justificación del retorno de la inversión. Hay inversiones que son fáciles de medir, sin embargo también existen las que no son fácilmente medibles. Se ha hecho proyectos donde un grupo de ejecutivos que duraban dos días montando la información y luego de la habilitación de la inteligencia de negocios, esos dos días están liberados. Esto con una serie de fórmulas se puede determinar el retorno inversión.

Silos departamentales. La mayoría de empresas está dividida en departamentos. Esto implica que existan fracturas o silos que dificultan el poder ver los procesos de principio a fin y los datos que se necesitan a lo largo de todo el proceso. Al usar herramientas de inteligencia de negocios no se va a terminar afectando departamentos, sino afectando procesos.

Existen dos elementos dentro de la Organización que hacen que varíe radicalmente el impacto de un proyecto de inteligencia de negocios, estos son Las personas que conforman la Organización y la propia Organización.

Calidad de datos.

La calidad de datos se convierte en un requisito para alcanzar el siguiente nivel en inteligencia de negocios. El resultado de la falta de calidad en los datos se experimenta normalmente en el día a día de las operaciones Batini, C., & Scannapieca, M. (2006). En esta sección se hace una revisión generalizada de lo relacionado a calidad de datos y sus dimensiones.

La calidad de datos es más que simplemente la exactitud de los datos, existen otras dimensiones como al completitud, consistencia y actualidad las cuales son necesarias para poder caracterizar de manera completa la calidad de los datos Batini, C., & Scannapieca, M. (2006).

Dimensiones de la calidad de datos. En esta sección se hace referencia a las distintas dimensiones de calidad de datos, es decir los criterios en los que se puede basar para indicar si los datos tienen o carecen de calidad.

Puntualidad. Se refiere a que los datos pueden verse en un momento específico de tiempo, es decir los cambios a lo largo del tiempo se van almacenando y se puede ver que paso en un momento determinado.

Exactitud. Hace referencia de la relación entre un valor existente y lo que realmente es considerado como la representación correcta del fenómeno en la vida real que busca representar Batini, (2006). Hay dos tipos de exactitud, la sintáctica que se relaciona con el valor del elemento y la semántica la cual se trata de cuan cerca está el valor con el valor real.

Completitud. Este concepto trata de que en los datos no esté faltando nada. Por ejemplo a nivel de columna se puede determinar su completitud de acuerdo a la cantidad de valores que están faltando para esa propiedad específica en la tabla. Esto puede suceder por reglas de negocio o por un error en el sistema. De ahí que según Batini, se puede definir como la extensión en la cual los datos tienen la amplitud, profundidad y alcance para lo que se está realizando (2006).

Consistencia. Esta dimensión considera aspectos de integridad. Esta captura violaciones a las reglas semánticas definidas sobre un set de datos. Un ejemplo de ello son las limitaciones de integridad, cuando estas son definidas pero no respetadas, se está fallando en el aspecto de consistencia. Igualmente tiene que ver con los registros, los cuales deberían seguir un estándar o formato, es decir ser consistentes.

Accesibilidad Según Batini (2010) la accesibilidad mide la capacidad del usuario para acceder a los datos desde su propia cultura, estado físico y funciones además de la tecnología disponible. Esto se relaciona más con aspectos de usabilidad, sin embargo la manera como se hace disponible los datos se convierte en factor determinante a la hora de medir esta dimensión de calidad.

Soluciones a problemas de rendimiento de base de datos.

Cursores. Para iniciar esta sección, se presenta lo que comúnmente se dice de los cursores son lentos, causan problemas de rendimiento y aumentan la complejidad del código.

Lo anterior no es sacado de un libro, es en sí lo que se escucha y se ve cuando se busca información de cursores, sea para SQL Server, o MYSQL o algún otro motor.

Por su parte, uno de los autores que se conocen más en temas de rendimiento de bases de datos, Aaron Bertrand, en su artículo "What impact can different cursor options have?" confirma este asunto. El afirma que "en la mayoría de los casos es mejor reescribir cursores por lógica basada en sets de TSQL. Sin embargo el no cierra la posibilidad de que estos sean necesarios en ciertas situaciones. Igualmente existen otros autores que indican que los cursores se crearon por "alguna razón" y que se mantienen por lo mismo.

Existen algunas opciones que ofrece el motor de SQL Server a través del TSQL para mejorar el rendimiento de las consultas realizadas con cursores.

LOCAL FAST_FORWARD: estas son dos opciones que en los experimentos realizados y los resultados de diferentes autores, terminan siendo las opciones más útiles. Por defecto SQL Server al crear cursores, estos se hacen de manera global (depende de una configuración de la base de datos). Si le quita esas opciones por defecto especificando el tipo de cursor, en este caso particular con estas dos opciones se crearía a nivel local y "con opciones de optimización habilitadas"

LOCAL STATIC: Esta opción, en los experimentos que se hicieron, causa el efecto contrario. En sí la opción "static" lo que hace es crear una copia de los datos para el cursor y trabajarlos desde la base de datos TEMPDB. Entonces con unos 390 mil registros el tiempo de duración iba por 7 minutos antes de que se cancelara la operación.

Vistas. En esta sección se realiza una reseña acerca de las vistas y sus prácticas recomendadas para poder mejorar el rendimiento de la base de datos.

Vistas materializadas o indexadas. Cuando se está trabajando con muchas tablas, existen ciertas consultas que requieren de múltiples objetos tipo tabla para poder generar un resultado, sea por legibilidad al usuario, sea por el requerimiento de una aplicación.

Esto es lo que hace las vistas, toman una consulta compleja y la "esconden" para que el usuario le sea más fácil el poder obtener los datos. Sin embargo por debajo lo que hace el motor es sustituir toda esa vista por la definición que en ella se encuentra. Entonces en resumen, con las vistas se puede

- Hacer compatible con un idioma específico.
- Limitar al usuario escondiendo ciertas columnas.

Las vistas tradicionales presentan algunos problemas. Empezando por la parte de rendimiento, no por la vista en sí, sino por la complejidad que se busca abstraer dentro de la misma.

Ahora bien, SQL Server ofrece una solución a ciertas consultas que los resultados son casi siempre los mismos y se hacen muchas veces durante el día. Esto se puede lograr a través del uso de las vistas materializadas.

Según el autor Ahmad Yaseen en su artículo "SQL Server indexed views", los beneficios vienen de datos que son actualizados frecuentemente. El autor indica en ese artículo además que todos los beneficios de rendimiento que se pueden obtener, se quitan con el hecho de estar manteniendo todas esas estructuras y que por ende son más útiles en ambientes de almacenes de datos y OLAP.

En un artículo interesante de vistas indexadas en la página de C Sharp Corner, se indica que se muestran generalidades de este tema. De esa guía se puede destacar lo siguiente

- En la versión Empresarial o para desarrollador no se requiere usar ningún HINT para lo que es el uso de los índices de vistas. En las otras ediciones se requiere el uso de "NOEXPAND".

- Se requiere que las tablas base están creadas con ANSI_NULLS correctamente establecidos (ON).
- Requiere la directiva with SCHEMABINDING.
- Si no se establece eso entonces va a dar error.

Algunos problemas que podrían suceder con las vistas.

- Luego de alterar una vista indexada los índices desaparecen.
- El modificar columnas usadas en la vista indexada conlleva mucho trabajo en el motor de base de datos.
- El rendimiento de las consultas si se mejora en gran cantidad con respecto a las vistas “normales”.

Las vistas indexadas además requieren lo siguiente, según la documentación de Microsoft

- Permisos de creación y alteración de la vista.
- No se puede usar distinct, count, min, max, top, unión, order by y otras expresiones o funcionalidades TSQL
- No usar otras vistas.

TSQL. SQL Server a través de TSQL ofrece opciones tales como las expuestas a continuación.

Vistas dinámicas de gestión. Esta es otra de las facilidades que trae el motor SQL Server. Estas fueron introducidas desde SQL Server 2005. Estas vistas ofrecen información que anteriormente era bien complicado el poder obtenerlas. Hoy en día existe un “sin número” de este tipo de vistas en el motor, sin embargo siempre hay ciertas que son con un nivel superior de utilidad que otras.

Según la documentación de Microsoft existen dos tipos de vistas dinámicas de gestión:

- A nivel de servidor
- A nivel de bases de datos.

Por otro lado debido a la gran cantidad de vistas disponibles, se han clasificado en distintas categorías, que son bastantes. Se van a cubrir como ejemplo algunas las vistas más referenciadas en la web.

Por ejemplo. Se mencionan algunas vistas útiles. Tales como

- Sys.dm_exec_cached_plans: que trae la información de planes cacheados de consultas a SQL Server.
- Sys.dm_exec_sessions: relacionadas a sesiones en SQL Server.
- Sys.dm_exec_connections: que trae la información de las conexiones a SQL Server.
- Sys.dm_db_index_usage_stats: esta trae información útil acerca de seeks, escaneos, lookups para con los índices. Es decir, el uso de índices.
- sys.dm_io_virtual_file_stats – estadísticas de IO para las bases de datos y de logs.
- sys.dm_tran_active_transactions – estados de las transacciones activas actualmente
- sys.dm_exec_sql_text – código TSQL ejecutado
- sys.dm_exec_query_plan – Devuelve información concerniente al plan de ejecución de consultas.
- sys.dm_os_wait_stats – Información acerca de recursos que están causando waits o esperas.
- sys.dm_os_performance_counters – Estadísticas de rendimiento de SQL Server

Procedimientos almacenados. En esta sección se menciona una serie de procedimientos almacenados útiles para analizar el proceso de ejecución de una consulta y elementos que son considerados como útiles a la hora de optimización.

Sp_executesql. A nivel de procedimientos almacenados, SQL Server ofrece una serie de utilidades que ayudan en el proceso de mejora de las consultas y de las operaciones.

Una de las técnicas que ayudan a mejorar el rendimiento es la capacidad de reutilizar planes de ejecución. Al mismo tiempo existen prácticas que impiden que tal objetivo se cumpla, un ejemplo sería.

```
declare @pelicula_id int=15133;
exec('SELECT [nombre] FROM [pelicula] where pelicula_id='+@pelicula_id);
```

El ejemplo anterior, aunque no represente un caso que suceda en el día a día directamente, si representa una técnica que es la creación de consultas con los parámetros del usuario. Esta técnica lo que hace es, si ejecutas la misma operación para distintos identificadores de películas, va crear un plan de ejecución por cada valor distinto que se encuentre, en este caso para la película 15133 habrá un nuevo plan, si se pasara 15134 habrá otro nuevo y así sucesivamente.

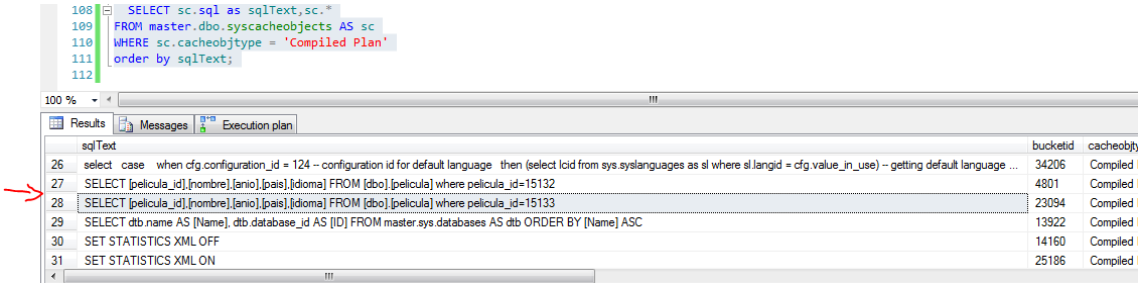


Figura 5 Planes de ejecución guardados (Fuente Elaboración Francisco Coulon)

Nota: aparecen campos extras ya que en el documento se ha dejado solo el nombre de la película por efectos de espacio y claridad de la consulta.

Existe un procedimiento almacenado que permite evitar este tipo de situaciones con las consultas parametrizadas, de la siguiente manera.

```
declare @pelicula_id int=15133;
exec sp_executesql N'SELECT[nombre] [pelicula] where pelicula_id=@pelicula_id',
N'@pelicula_id int',
@pelicula_id;
```

Si se ejecuta la consulta anterior, se obtendrá el mismo resultado, pero a nivel interno ocurrirá lo siguiente

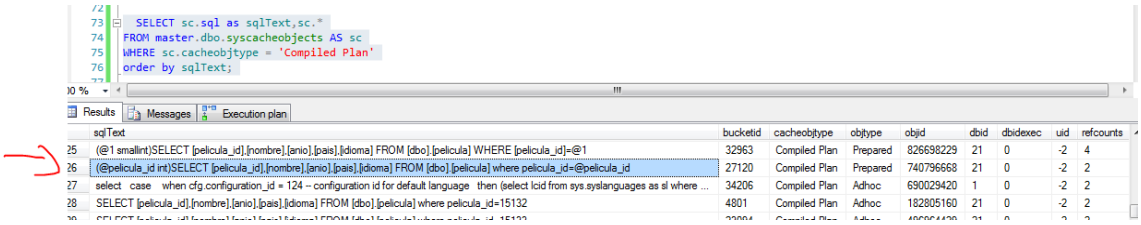


Figura 6 Planes de ejecución guardados (Fuente Elaboración Francisco Coulon)

Si se vuelve a ejecutar con otros parámetros diferentes (por ejemplo usando 15138 el plan de ejecución no varía porque está "parametrizado" y puede ser reutilizado.

SP configure. Esto de SP configure es una herramienta para mostrar y cambiar las configuraciones del servidor de manera programática. Tiene muchas opciones, en esta subsección se pretende mostrar que "existe" esta funcionalidad por el hecho de ser un tópico más avanzado y que no todo aplica para rendimiento.

Si se ejecuta el siguiente procedimiento

```
EXEC sp_configure
```

Se obtiene la siguiente salida

```
6  EXEC sp_configure
```

.00 %

Results | Messages

	name	minimum	maximum	config_value	run_value
34	locks	5000	2147483647	0	0
35	max degree of parallelism	0	32767	0	0
36	max full-text crawl range	0	256	4	4
37	max server memory (MB)	128	2147483647	2147483647	214748...
38	max text repl size (B)	-1	2147483647	65536	65536
39	max worker threads	128	65535	0	0
40	media retention	0	365	0	0
41	min memory per query (KB)	512	2147483647	1024	1024
42	min server memory (MB)	0	2147483647	10	16
43	nested triggers	0	1	1	1
44	network packet size (B)	512	32767	4096	4096
45	Ole Automation Procedures	0	1	0	0
46	open objects	0	2147483647	0	0
47	optimize for ad hoc workloads	0	1	0	0
48	PH timeout (s)	1	3600	60	60
49	precompute rank	0	1	0	0
50	priority boost	0	1	0	0
51	query governor cost limit	0	2147483647	0	0
52	query wait (s)	-1	2147483647	-1	-1
53	recovery interval (min)	0	32767	0	0
54	remote access	0	1	1	1
55	remote admin connections	0	1	0	0
56	remote login timeout (s)	0	2147483647	10	10
57	remote proc trans	0	1	0	0
58	remote query timeout (s)	0	2147483647	600	600
59	Replication XPs	0	1	0	0

Figura 7 Salida del procedimiento sp_configure (Fuente elaboración Francisco Coulon)

Las columnas que muestra son las siguientes

1. name – el nombre de la configuración.
2. minimum – el valor mínimo permitido para la configuración en cuestión.
3. maximum – el valor máximo permitido para la configuración en cuestión.
4. config_value – indica el valor actual configurado.
5. run_value – cuál es el valor que está corriendo actualmente.

Por ejemplo, si se va a cambiar algunas configuraciones que pueden ayudar a mejorar el rendimiento de las consultas como lo es la memoria mínima por consulta y el máximo de memoria permitida para el servidor. También lo está configuraciones como el factor de llenado por defecto de los índices (para los que les gustan las configuraciones por defecto)

```
EXEC sp_configure 'max server memory (MB)', 2048
```

Y luego si se ve las configuraciones se verá lo siguiente

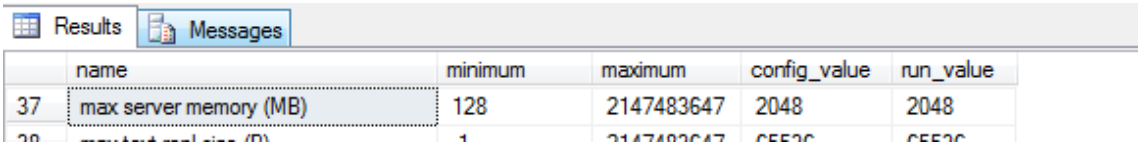


Figura 8 Resultado de cambio de configuración de memoria con sp_configure (Fuente Elaboración Francisco Coulon)

Lenguaje de definición de datos. Este conocido también como DDL por sus siglas en inglés, provee ciertas funcionalidades que mejoran el rendimiento de las consultas a través de elementos como los mencionados a continuación.

Rebuild index. Antes de entrar de lleno con la sección, se va a recordar el concepto de fragmentación de índices. Según la documentación de Microsoft lo que sucede es que a lo largo del tiempo las bases de datos van sufriendo cambios, desde "ingresos", "actualizaciones", "eliminaciones" en sus datos. Esto ocasiona que después de cierto tiempo los índices se fragmenten y después de cierto nivel de fragmentación se afecta el rendimiento.

Esta fragmentación quiere decir que las páginas en las cuales se tiene ordenado los datos no es la misma que el orden físico dentro del archivo de datos. Microsoft confirma que se puede solucionar este problema mediante la reconstrucción o reorganización de los índices.

La gran diferencia entre reconstruir y reorganizar es que la reconstrucción rehace todo "nuevo" y utiliza más recursos. Reconstruir significa que se "eliminan" y se "crean".

Reorganize Index. Esta operación busca lo mismo que lo de la sección anterior, pero requiere menos capacidad de procesamiento. Según la recomendación encontrada en la documentación oficial de Microsoft, fragmentación menor o igual a 5% no debería hacerse "nada" por el hecho de lo que implica el reorganizar o reconstruir. Todo lo que se "arreglaría" se pierde por lo que se ocupa para tal objetivo y los beneficios no van a mejorar.

Comando de Consola de bases de datos. Esta es otra de las categorías que se van a mencionar en este documento. Este término hace referencia a los comandos de consola de base de datos para SQL Server. Según la documentación de Microsoft, existen varias categorías de comandos

- comandos de mantenimiento.
- Misceláneos.
- De información.
- De validación.

¿Pero no habla ninguno de optimización? Quizás directamente no, sin embargo para optimizar hay que probar, y en ocasiones cuando se prueba no se está haciendo bien ese proceso. Un ejemplo es el siguiente caso, basado en lo que dice Glenn en todas sus presentaciones de optimización.

Alguien que está optimizando una consulta, se da cuenta que hay que cambiar unas relaciones, cambiar la forma en que se ejecuta la consulta, luego de ciertos minutos ejecuta la nueva versión, luego la vuelve a ejecutar y se da cuenta que ha creado una consulta "perfecta". Luego al rato después de haber sido aplicada en producción los problemas regresan. Se analiza y tiempo después se da cuenta que todas esas validaciones hechas no eran válidas en realidad porque después de la primera ejecución todos los datos se quedaban en caché, no había lecturas físicas. Por lo cual no se podía medir exactamente qué tanta mejora había.

De ahí que surge la importancia de conocer algunos comandos de consola de bases de datos que van a ayudar a probar realmente una consulta.

Algunos de estos comandos importantes a la hora de probar alguna solución. La parte común es que no deberían ser usados en producción.

DBCC FreeProcCac. Este comando elimina todos los elementos del caché de plan. También puede eliminar algunos elementos específicos indicando el manejador del plan o elementos asociados a un pool de recursos. Esto lo que hace entonces es que un procedimiento, sea recopilado en vez de usarse algún plan en la caché. Esto hace que el rendimiento baje en cierta medida por cierto tiempo.

DBCC FreeSystemCache. Este elimina todas las entradas de caché que no se usen de todas las caches del sistema. Este comando va limpiar el caché de planes que se tengan en el servidor, haciendo que las consultas estén más lentas mientras se vuelven a generar.

DBCC FlushProcInDB. Este comando permite especificar una base de datos específica. Este elimina todos los planes por esa base de datos particular.

DBCC DropCleanBuffer. Este comando sirve para poder tener el servidor como si se hubiese reiniciado, pero sin tener que reiniciar. Limpia toda la caché.

DBCC FreeSessionCache. Elimina o limpia la caché de consultas hechas por diversas fuentes hacia la instancia de SQL Server.

Herramientas de evaluación de bases de datos

Asesor de ajuste. Una de las herramientas que vienen con la instalación del motor de base de datos es la herramienta de asesoramiento de ajuste. Esta herramienta se encarga de analizar consultas, sea una o sean varias y va retornar algunas recomendaciones de mejora en las mismas consultas.

Además de ello, basado en lo que se indica en la página de MSSQLTIPS y lo experimentado con la herramienta esta puede también

- Crear índices.
- Crear particiones.

- Crear estadísticas.
- Auto generar el TSQL necesario para aplicar recomendaciones.

Ventajas de esta herramienta. Con esta herramienta se puede gráfica y de manera fácil se pueden tener resultados interesantes. También los resultados son entendibles y se puede aplicar las recomendaciones con la misma herramienta o exportar los resultados. Inclusive se puede ver código TSQL para ejecutarlo si se requiere.

Perfilador SQL Server. Esta es otra herramienta que es dada por el paquete de SQL Server. En sí lo que permite es ver qué consultas se están haciendo de cualquier índole al servidor. Además se pueden ver aspectos como

- Duraciones.
- Lecturas.
- Escrituras.
- Datos del equipo.

Una de las partes interesantes de la herramienta es que es gráfica y también tiene buena integración con el asesor de ajuste. Esta herramienta en sí no ayuda a mejorar las consultas, pero se convierte en una herramienta donde se puede detectar cargas de trabajo específicas para poder usarlas en el asesor de ajuste o utilizando la experiencia.

SQL Server Management Studio. Esta es otra de las herramientas que trae el paquete de instalación de SQL Server (ya para la versión 2014 en adelante ya no se incluye).

En SQL Management Studio tiene una funcionalidad que le dice el plan de ejecución de manera gráfica, que este podría ser un primer paso para descubrir que falta o sobra un índice. Esto es útil en casos donde quizás no se sepa porque está el problema, entonces con este sistema podría tenerse una idea inicial de que está pasando con respecto al rendimiento de la consulta.

Igualmente, la mayoría de funcionalidades es gráfica, por lo que permite darse cuenta y hacer optimizaciones de manera sencilla. Para usuarios no tan avanzados se convierte en una herramienta administrativa útil.

Servicios de integración de SQL Server.

En esta sección se hace referencia al paquete de servicios ofrecidos por Microsoft para la integración de datos de la organización. Este se convierte en una primera opción de herramienta para poder ser aplicado en el proyecto debido a su naturaleza y también por su importancia a lo largo de los cursos de la carrera.

Definición. Los servicios de integración de Microsoft, es una plataforma lista para la ser usada en desarrollo y provee la capacidad de extraer y cargar de fuentes de datos diversas e igualmente con el destino Tok, W. H., Parida, R., Masson, M., & Ding, X. (2012).

Requisitos. Los requerimientos mínimos para poder correr los servicios de integración en el equipo se resumen a continuación. Estos son basados en la lista oficial de Microsoft y lo indicado en el libro de SQL Server Integration Services de Tok, W. H., Parida, R., Masson, M., & Ding, X. (2012).

1. Framework .NET 3.5 SP1.
2. SQL Server Edición estándar o superior.
3. SQL Management Studio.
4. Procesador de al menos 1.4GHz o superior.
5. 6 gigas de espacio.
6. 1 giga de memoria.
7. Requisitos propios al motor de bases de datos de SQL Server.

Escenarios de uso. En esta subsección se muestra una lista de posibles escenarios donde pueden hacerse uso de los servicios de integración de SQL Server.

Estos van desde tareas de migración de datos, consolidación de datos de fuentes diversas, carga de almacenes de datos, limpieza y formateo de datos, identificación y

captura de cambios de datos, análisis de datos, creación de vistas de fuentes de datos federadas además de procesamiento de eventos en tiempo real.

Servicios de análisis de SQL Server

En esta sección se hace referencia al paquete de servicios ofrecidos por Microsoft para el análisis de datos de la organización. Este se convierte en una primera opción de herramienta de análisis y procesamiento para poder ser aplicado en el proyecto debido a su naturaleza y también por su importancia a lo largo de los cursos de la carrera.

Definición. Los servicios de análisis de Microsoft, es parte del ecosistema de inteligencia de negocios de Microsoft, el cual no consiste en un simple producto Ferrari, A., Russo, M., & Webb, C. (2012). Es una base de datos tipo OLAP la cual por su naturaleza es optimizada para análisis y cálculos los cuales son comunes en un ambiente de inteligencia de negocios Ferrari, A., Russo, M., & Webb, C. (2012). SSAS se puede considerar como una capa extra la cual busca que las consultas no se hagan directamente a la base de datos sino que a los modelos diseñados previamente, esto permite centralizar la información y las consultas que el usuario debe escribir son más fáciles de manejar.

Cubos multidimensionales

Es una estructura que contiene información para que sea analizado y evaluado. Esta estructura contiene múltiples componentes, siendo las dimensiones y medidas las principales de ellas (Microsoft, 2018).

A nivel de implementación, en un modelo multidimensional fuera del tipo tabular, el cubo se convierte en la base, de hecho los datos se modelan en una serie de cubos y dimensiones y no tablas. Cada cubo está hecho por un grupo de medidas el cual se encuentran en una tabla de hechos. Ferrari, A., Russo, M., & Webb, C. (2012). La mayor ventaja de usar cubos es el hecho de que los datos se ponen todos en un solo lugar para su análisis.

En términos de almacenamiento, el modelo multidimensional se puede almacenar en tres diferentes formas que se indican a en las secciones subsecuentes.

Procesamiento analítico en línea multidimensional. Más conocido como OLAP y es cuando los datos se almacenan siguiendo el modo o estilo de almacenamiento de los servicios de análisis de SQL propiamente. Este es utilizado en su vasta mayoría de implementaciones. .Ferrari, A., Russo, M., & Webb, C. (2012).

Procesamiento analítico en línea relacionado. Más conocido como ROLAP, en este caso los datos no se almacenan en los servicios de integración de SQL Server y este se convierte en una capa de metadatos haciendo que las consultas se deban hacer directamente a las fuentes de datos

Procesamiento analítico en línea híbrido. Hace referencia al modelo Híbrido de OLAP en el cual se sigue el estilo expuesto en la parte 2.18.2 pero donde los valores calculados se almacenan en MOLAP.

Capítulo tercero. Marco Metodológico

En este capítulo se presenta de manera detallada la metodología utilizada en la investigación y ejecución del proyecto. En las secciones subsecuentes se analizaran aspectos como el diseño y el enfoque que fueron tomados para la investigación, el tipo de investigación realizada como preámbulo al planeamiento de la ejecución y otros elementos destacables para el éxito del presente proyecto.

Tipo de investigación

En esta sección se indican diversas clasificaciones en las que la investigación/proyecto se encuadra. Si bien es una investigación aplicada debido a su naturaleza de poder aplicar lo que se aprende para resolver algún tipo de problema o situación en este caso fue en el departamento de cobros, existen otros criterios en los cuales calza.

De acuerdo al manejo de las variables se clasificó este proyecto investigativo como no experimental y de campo. Es experimental puesto que se realiza sin manipular deliberadamente variables (Sampieri, Collado, & Baptista, 2014, pág. 152). No se pretendió hacer ningún tipo de experimento sino que se buscó solo hacer análisis de lo que actualmente está sin ninguna alteración y solo se pretende estar en el departamento beneficiado con la solución.

Por último y tomando en cuenta el estilo de estudio a adoptar en esta investigación, se tiene el método inductivo, que buscó a partir de conocimientos adquiridos a lo largo de la carrera, documentación presentada en la primera parte de este documento y las situaciones presentadas en la organización emitir un criterio. Un proceso inductivo explora y describe para luego generar perspectivas (Sampieri et al. 2014, pág. 8) Sin embargo este criterio va a ofrecer solamente un conocimiento probable y no necesariamente aplicable en todo tipo de situaciones debido a la naturaleza de las necesidades departamentales.

Alcance investigativo

Como se indicó en las secciones de objetivos y alance, el objetivo principal de este proyecto es lograr presentar un prototipo o una prueba de concepto a los encargados del

departamento, para que pueda ser analizado y evaluado. Según el alcance, este será de manera predominante el tipo exploratorio, debido a que a como se presentó en la literatura, a nivel de ámbito público, la soluciones de inteligencia de negocios no tienen mucho auge o aplicación hoy en día por lo que no existe una metodología clara.

Tal y como lo indica (Sampieri et al. 2014) las investigaciones exploratorias por lo general, antecede a las demás investigaciones (p.90) y esta se basa principalmente al estado del conocimiento mostrado en las secciones anteriores. El objetivo de las mismas es poder examinar un tema y problema de investigación poco estudiado, del cual se tienen muchas dudas o no se ha abordado antes (Sampieri et al. 2014, p.91)

Igualmente adoptó un modo de investigación explicativa, debido a que se pretende determinar causas del porque las soluciones de inteligencia de negocios, si bien existen en el ámbito público, estas no están con la importancia que en otros ámbito de la sociedad existen. Tal y como lo indica Sampieri, estas se centran en explicar porque ocurre un fenómeno y en qué condiciones se manifiesta o porque se relacionan dos o más variables (Sampieri et al. 2014, p.95)

Enfoque

Según su enfoque clasificó como una investigación mixta, debido a que habrá elementos cuantitativos como también caracterizaciones que no entran en el lado de los números o contables. Además se fundamenta por lo indicado por Creswell (2013a), Niglas (2010) y Unrau, Grinnell y Williams (2005) citados en Sampieri et al. (2014), los cuales indican que la elección del enfoque mixto se elige cuando el investigador considere que se adapte al planteamiento del problema y/o que el investigador posea más conocimientos y entrenamiento.

Tiene elementos cuantitativos porque será “secuencial y probatorio” (Sampieri et al., 2014). Por ejemplo a nivel de rendimiento de bases de datos, comparó numéricamente el rendimiento antes y posterior a la implementación de la solución.

Según Sampieri (2014) el enfoque cualitativo se selecciona cuando el propósito es examinar la forma en que los individuos perciben y experimentan los fenómenos que los rodean, profundizando en sus puntos de vistas, interpretaciones y significados (p.358). En este caso se aplicó al momento de evaluar lo que es el nivel de satisfacción del proceso existente con respecto a los resultados luego de la implementación.

Diseño

Para el presente proyecto el diseño adoptado fue el no experimental para poder medir la propuesta o prueba de concepto a desarrollar con respecto a lo que actualmente existe sin alterar las variables.

En un primer momento hizo fue una recopilación de datos alrededor de las métricas del motor de bases de datos que midan la cantidad de los datos existentes para soportar los distintos procesos necesarios para las operaciones de la organización.

Cuando ya se contó con dichos datos se pasó a una segunda etapa, la elaboración de un diseño inicial para el almacén de datos. Con ello nos dimos cuenta la compatibilidad de la estructura existente con la idea del prototipo.

En una tercera etapa se implementó la solución de la prueba de concepto del cubo multidimensional. Este tiene datos pre-calculados y listos para ser utilizados y se hizo una evaluación de dichos desarrollos a través de la creación de visualizaciones. Así se pudo determinar que la estructura planteada si responde a lo que se pretende.

Población y muestreo

El proyecto estuvo orientado al departamento de cobros de la junta administrativa de la universidad de Costa Rica. En él participaron los encargados de los procesos afectados, siendo los mismos clientes de la solución, igualmente con el proceso de evaluación.

Por otro lado se tuvieron los motores de bases de datos a utilizar, los cuales están ubicados dentro de las instalaciones, de allí se sacaron métricas de bases de datos para las distintas evaluaciones realizadas.

Por la naturaleza de los datos a recolectar, estos se clasificaron en un conjunto de datos discrecional.

Instrumentos de recolección de datos

En una primera instancia se elaboró una serie de encuestas que fueron respondidas por los encargados del departamento, estos sirvieron como métrica para ver el estado de satisfacción con lo que se hace al día de hoy. Este proceso se dividió en tres partes principales

1. Elaboración de las preguntas.
2. Aplicación de las encuestas a los usuarios.
3. Tabulación de dichas encuestas.
4. Preparación de observaciones y mediciones obtenidas.

Igualmente las instancias de bases de datos que existen fueron otra fuente importante de datos, en ellas se extrajeron una serie de métricas que son utilizadas para la elaboración de mediciones de salud y rendimiento de un motor de bases de datos.

Además de ambas fuentes anteriormente mencionadas, la implementación del cubo trajo consigo otro elemento que es el cubo mismo, para poder extraer datos del mismo y en combinación a lo que el usuario final considera de la solución se podrá determinar la efectividad de la misma.

Técnicas de análisis de información

A partir de las encuestas extraídas, se analizaron los datos con la herramienta de Microsoft Excel para sacar gráficos acerca de las opiniones de los diversos usuarios del departamento en cuestión.

Se pretendió utilizar diversas herramientas de medición o evaluación de la situación actual. Una de ellas fue el diagrama de espina de pescado o conocido también como el Diagrama de Ishikawa, el cual representa elementos (que son causas) de algún contexto

(que se convierte en el efecto). Con esto se pretendía describir de manera simple y concreta lo que está sucediendo tanto a nivel de procesos en el departamento como a nivel de bases de datos, sin embargo por utilidad o relevancia en el proyecto, se consideró que no era necesario aplicar este tipo de técnica.

La otra herramienta que se pretendió utilizar en el presente proyecto fueron los diagramas 80-20 o más conocido como diagrama de Pareto, el cual busca igualmente representar de manera gráfica datos que se recolectan sobre una situación en particular, identificando los elementos prioritarios o que deberían ser tratados en primera instancia. Con este se pretendió tener un panorama más claro acerca de cuáles factores son los principales que afectan a nivel de bases de datos y de usuarios.

Estrategia de desarrollo de la propuesta

En esta sección se pretende dar con mayor detalle el proceso que se siguió para la ejecución del proyecto. En una primera parte se expone las distintas etapas del proyecto. En una segunda parte, se detallan las actividades principales de cada una de esas fases primeramente indicadas. Al final se muestra un resumen del cronograma (indicando fechas) con respecto a la implementación de la solución y distintas herramientas de apoyo utilizadas y que se pretenden utilizar durante la duración del proyecto.

Etapas del proyecto. El proyecto se dividió en tres etapas principales. Estas se detallan a continuación.

Etapa 1 – Análisis de la base de datos. En esta etapa se hizo una revisión de la base de datos con que se cuenta. Se determinó aspectos como la distribución de datos. En esta etapa además se hizo una determinación del estado de "salud" de la base de datos, desde la dimensión de los datos en sí y también a nivel de rendimiento de la base de datos.

En esta primera se tuvo ya una revisión de los objetos de bases de datos y una documentación de las distintas características de la base de datos en donde se hizo las operaciones iniciales del proyecto.

Etapa 2 –Diseño de cubo. En esta segunda etapa, ya con la información de la primera etapa, se inició el diseño del cubo multidimensional, que se convierte en la parte principal del proyecto y se concentró en la etapa tres. Esto implica diseño de procesos de extracción y carga, elaboración del almacén de datos, entre otras.

Etapa 3 – Implementación de la base de datos multidimensional. En esta etapa se procedió a realizar la implementación del cubo utilizando las distintas herramientas de las cuales se optó. Las salidas de esta etapa se convierten en la parte principal del proyecto. Además se incluye la elaboración de las visualizaciones en dos herramientas distintas.

Actividades principales de cada etapa. El desglose de las actividades principales realizadas se presenta a continuación y se complementan con el cronograma de la siguiente subsección.

Etapa Uno.

- Planeamiento y revisión de los recursos técnicos disponibles.
- Análisis de la base de datos.
- Hacer la consulta a los diversos integrantes del departamento para comprender los distintos problemas que se presentan.
- Realización de consultas.
- Levantado de mediciones en distribución de datos y estado actual de la base de datos, en conjunto con una documentación de la misma.

Etapa Dos.

- Análisis de datos para el diseño del cubo.
- Diseño inicial del cubo multidimensional.
- Determinación de las consultas a la cuales el cubo deberá responder.
- Documentación de la primera fase.
- Revisión del estado del proyecto.

Etapa tres.

- Implementación del cubo multidimensional.
- Pruebas de funcionalidad del cubo.
- Carga del cubo.

- Elaboración de las visualizaciones.
- Documentación del cubo.
- Documentación del proyecto.

Capítulo cuarto. Análisis del diagnóstico

A continuación se realiza un análisis de los datos obtenidos, de lo planteado en la sección referente a las técnicas de recolección de datos. En un primer momento se realizó una serie de encuestas para los encargados del departamento, para conocer acerca de la situación actual y el uso de los datos existentes.

Igualmente se realizó un análisis de los objetos de datos para medir el nivel de complejidad de la misma y las implicaciones de la selección o no de elementos de datos particulares.

Nivel de satisfacción y uso de los datos

Se logra comprobar, a través de las encuestas que en términos generales, la satisfacción y la utilización de los datos esta:

- No es el mejor debido a las limitaciones de la base de datos y a las complejidades y dependencias que se requieren para acceder a los datos.
- Solo la mitad de los entrevistados indican que sus reportes se han de hacer de manera mensual.
- Todos los entrevistados indicaron que el modo como se accede a los datos es “Poco” o “Muy poco” confortable para el usuario y existen dependencias para poder sacarlos. Tal como se indica en apartados anteriores, esta dependencia es el departamento de tecnologías de información.
- Al mismo tiempo se logra observar que más del 66 por ciento de los entrevistados requieren generar gráficos para presentar los detalles de los reportes.

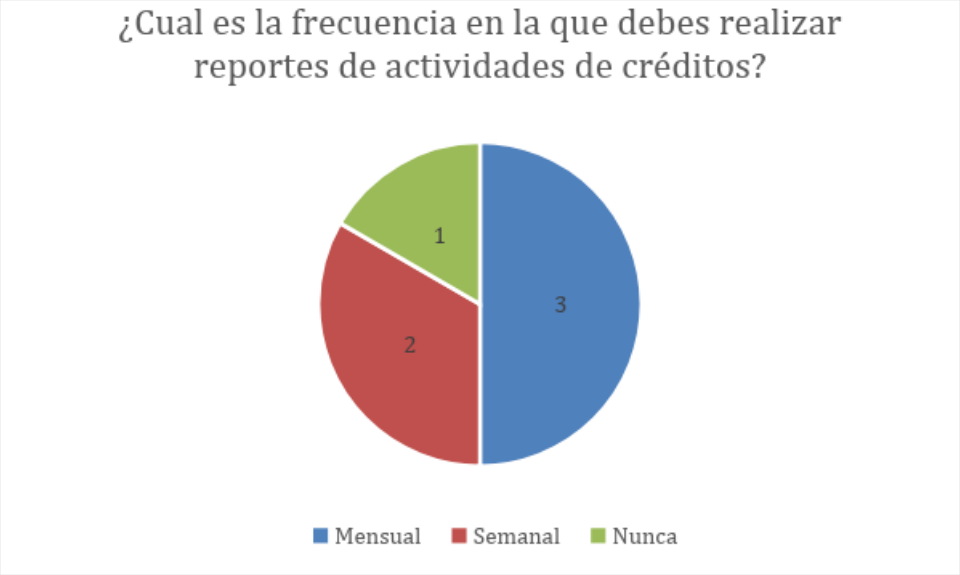


Figura 9 Gráfico con la frecuencia de realización de reportes de actividades.

Utilización de la tecnología.

Se observa que el conocimiento en ciertas herramientas de inteligencia de negocios no es tan alta. A nivel general con el uso de las tecnologías se observa que:

- Todos los encuestados reconocen que para extraer los datos se tienen dependencias.
- Además todos los que respondieron las preguntas, indican que si conocen acerca de PowerBI pero solo uno de ellos tiene idea de lo que es Tableau.
- Por otro lado, herramientas como los servicios de análisis de SQL Server no es conocido por ninguno de los encuestados.

Necesidad de datos

Se comprueba que los encuestados si consideran necesario o muy necesario la implementación de un sistema de inteligencia de negocios que facilite las tareas diarias de los encargados. Indican que el modo como se utiliza la base de datos no es confortable y que la misma no es rápida para poder hacer reportes rápidamente. Por último, se logra

observar que lo que ellos requieren no son datos “en tiempo real”, más bien datos históricos para ver la efectividad del departamento en un lapso de tiempo anterior.

Tamaño de la base de datos

Al momento del análisis de la base de datos, se confirma los siguientes factores:

- Que el tamaño de la misma asciende a más de 40GB.
- Acerca de la existencia de más de 1900 índices a lo largo de todas las tablas.
- Con respecto a la cantidad de registros, se logra determinar que en el momento del levantado de las características, la tabla con mayor cantidad de registros tenía alrededor de 16329083 filas.

Esto no significa que todos los elementos de datos serán utilizados, tal como se muestra en el apartado “Arquitectura de datos”, se seleccionaron una serie de objetos iniciales para dar respuesta al problema planteado.

Arquitectura de datos.

En esta sección se hace referencia a las distintas entidades que se han tomado en cuenta de la parte transaccional, su descripción y el tamaño del mismo. A partir de la observación inicial de la base de datos:

1. CUST_Customer: Información de clientes.
2. CUST_CustomerStatusType: Información de estados de tipos de clientes.
3. CUST_CustomerType: Información de tipos de clientes.
4. CUST_Region: Información acerca de la región del registro de los clientes.
5. DATA_CurrencyType: Información acerca del tipo de moneda.
6. DATA_PersonAssetVehicle: Información de registros de vehículos del cliente.
7. DEDU_Periodicity: Información de periodicidad.
8. LOAN_GuaranteeCategory: Información de la categorías de garantías.
9. LOAN_GuaranteeType: Información del tipo de garantías.
10. LOAN_InvestmentPlan: Información de planes de inversión.
11. LOAN_LoanOperation: Información de los préstamos.
12. LOAN_LoanOperationGuarantee: Información de las garantías de préstamos.

13. LOAN_LoanOperationTH: Información de las transacciones sobre los préstamos.
14. LOAN_LoanType: Información de los tipos de préstamos.
15. LOAN_LoanTypeCategory: Información de la categoría de préstamos.
16. LOAN_MonthlyPaymentType: Información acerca del tipo de pago mensual del préstamo.
17. LOAN_TransactionType: Información de los tipos de transacciones.
18. PERS_IDType: Información de los tipos de personas.
19. PERS_Person: Información de clientes/personas.
20. PERS_PersonJob: Información del oficio de las personas.
21. PERS_PersonSalary: Información del salario de las personas.

Capítulo quinto. Propuesta de la Solución

En esta sección se hace un análisis profundo de la solución planteada para el prototipo. Desde la elaboración de los procesos de carga de datos, hasta la estructura de la base de datos multidimensional y sus visualizaciones.

Tecnologías a utilizar

SQL Server. Este es el producto ofrecido por Microsoft, el cual es basado principalmente en el manejo de bases de datos al estilo relacional. Este motor de base de datos cuenta con su lenguaje, TSQL, el cual incluye características adaptadas del lenguaje base SQL y es la base de datos con la que cuenta el departamento para sus necesidades de datos. Este fue el componente principal durante la implementación de la propuesta.

Servicios de integración de SQL Server. Es la plataforma para poder realizar los diversos procesos de extracción, transformación y carga de datos, tanto a nivel de "Staging" como propiamente ya en la base de datos multidimensional.

Tableau. Como parte de las visualizaciones de la solución, se estuvo usando como parte de la propuesta. Este fue utilizado en las partes avanzadas del proyecto, en los esfuerzos de visualización de los datos.

Power BI. Igual que en el apartado anterior, de Tableau, este fue utilizado como parte de las visualizaciones de la solución y a su vez como comparativa. Este fue utilizado en las partes avanzadas del proyecto, en los esfuerzos de visualización de los datos.

Servicios de Análisis de SQL Server. Cuando ya los datos estaban cargados en la base de datos multidimensional, utilizando esta tecnología, se hizo la elaboración en formato tabular. Aquí se realizaron el formato y cálculos necesarios para el uso de la misma por parte de los usuarios finales.

Diseño de la arquitectura de la solución

Con la implementación de la solución, se pretende que el flujo de datos sea modificado de modo que las bases de datos transaccionales se vean utilizadas de manera más efectiva y no como al día de hoy que implica mucha lectura contra base de datos transaccionales, con consultas que consume mucho procesamiento.

A continuación se presenta el flujo de datos a seguir con esta propuesta y luego se indica de modo gráfico.

1. Los datos son producidos por clientes, que a través de sus laptops se conectan a realizar sus respectivos casos de uso. Esto puede ser desde el registro de un nuevo usuario hasta el registro de una transacción.
2. Estos datos son enviados a la base de datos transaccional.
3. A partir de este punto “todo es nuevo”. Los datos serán enviados a través de un proceso de extracción, transformación y carga (ETL) hacia el área de Staging.
4. Allí y luego de ser limpiados y optimizados, utilizando otro ETL para el almacén de datos se enviaran al repositorio de datos.
5. Desde este punto de datos multidimensional, eso sí en formato tabular, se crearán diversos tableros utilizando tecnología de Tableau y de PowerBI.

El proceso de carga se realizará de manera automatizada a través de tareas programadas las cuales permitan diariamente hacer el proceso de adición de nuevos elementos a la base de datos multidimensional. Al contar con servidores de plataforma Microsoft, implica que esas tareas programadas se configurarán con el “Task Scheduler”.

Existirá una tarea “padre” la cual permita ejecutar en primera instancia el proceso de staging y posterior el proceso para el almacenamiento multidimensional.

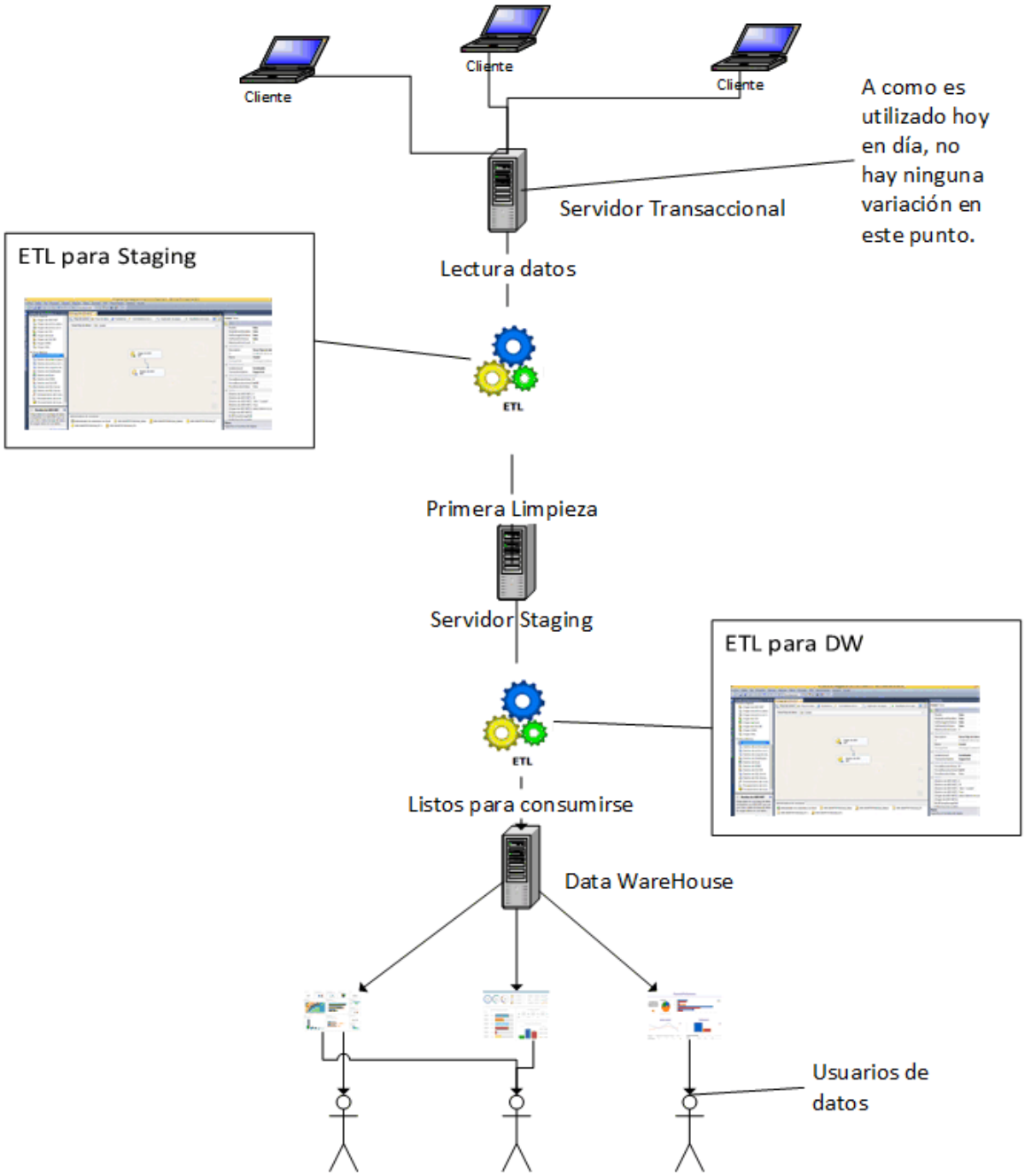


Figura 10 Diseño de la arquitectura de la solución (Fuente: Elaboración propia).

Diseño del área de Staging.

Conexiones. Existen un origen y un destino, el flujo de datos va desde la parte trasnacional hasta la parte de Staging. En la imagen siguiente se muestran tres objetos de conexión, pero es por motivos de los componentes de transferencia que están siendo utilizados.

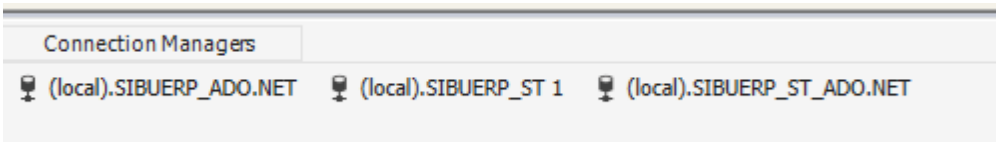


Figura 11 Conexiones para el área de Staging. (Fuente: Elaboración propia).

Se ha dado prioridad a las conexiones del tipo OLEDB con respecto a su contrapartida, ADO.NET debido a la flexibilidad que representan estas últimas para poder parametrizar los procesos. Especialmente en la extracción de la parte transaccional. Esto no significa que no se haga uso de conexiones del tipo ADO.NET.

Estructura de la base de datos. En la imagen siguiente se muestra la estructura de la base de datos del área de staging. Esta contiene las mismas estructuras de las tablas utilizadas del lado operacional, con las siguientes diferencias

- No existe ningún tipo de índice. Son tablas heap.
- Por tema de compatibilidad se decidió no hacer en tablas de memoria. Pero la idea inicial era esto para poder facilitar el proceso de carga (y acelerarlo también).
- No tienen ninguna relación entre ellas, aunque en la parte transaccional si existe.
- Se ha dejado todos los campos por tema de facilidad pero se ignoran los que no son copiables como la versión de la fila.

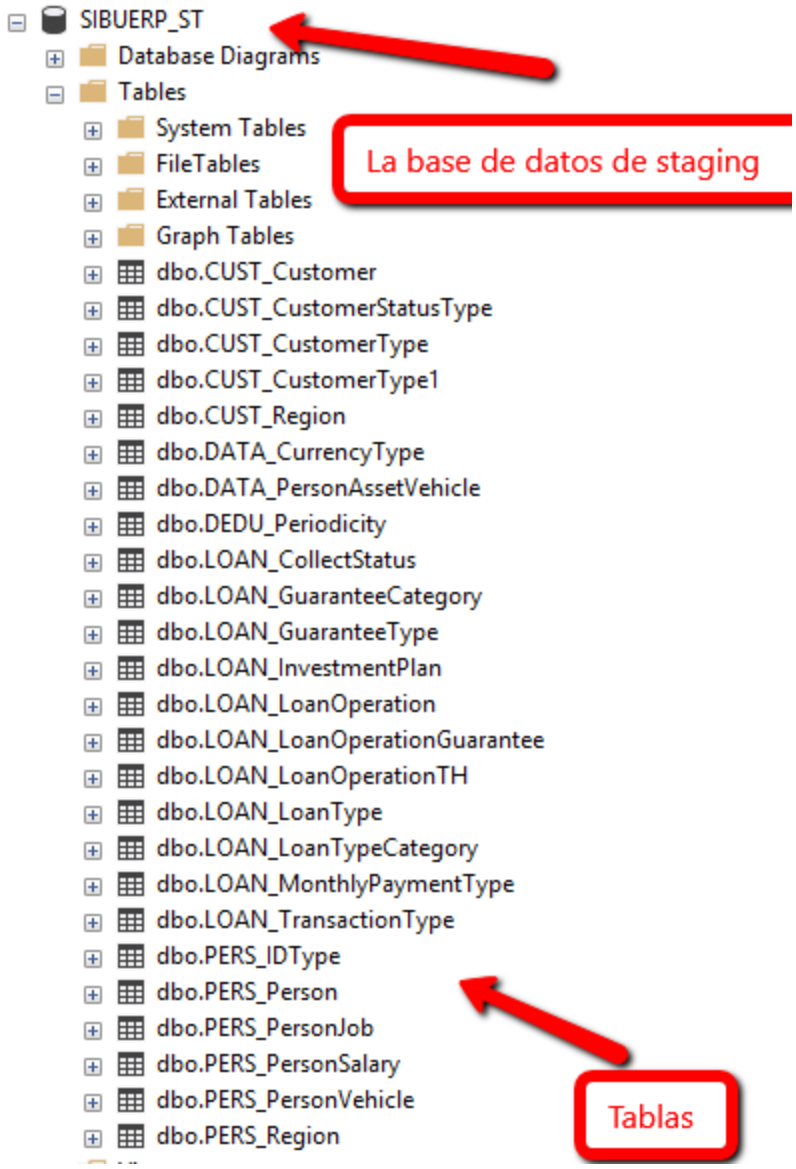


Figura 12 Objetos de la base de datos de Staging. (Fuente: Elaboración propia).

Proceso de extracción, transformación y Carga. Este consiste en dos procesos importantes, la eliminación o limpieza de las filas ya existentes en todas las tablas de esta área y la carga de cada una de ellas. Se pretende tener el proceso de validación de los registros cargar en esta parte, para que cuando se llegue al almacén de datos se pueda hacer la carga completa de lo que existe en esta parte.

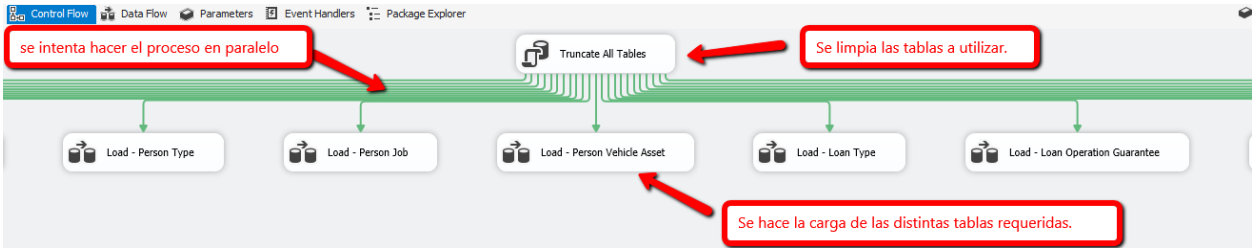


Figura 13 Proceso de carga de staging. (Fuente: Elaboración propia).

El proceso hará uso de la parte transaccional, sin embargo no hará lecturas completas de datos, debido a la capacidad de procesamiento disponible y la cantidad de datos existentes. Para poder mantener la trazabilidad de los procesos ETL, en el data warehouse, se mantendrá el detalle de los procesos y a partir de la última fecha de ejecución. Un ejemplo sería el siguiente.

- Se inicia el proceso ETL.
- Se consulta la última fecha de ejecución del ETL.
- Se limpia el área de almacenamiento temporal.
- Se carga del área transaccional aquellos registros elaborados posteriormente.
- Se trabaja sobre los registros extraídos.

Detalle del proceso de carga en el área de almacenamiento temporal. El proceso consiste, tal y como se muestra en la imagen siguiente, en primeramente tener un script que cargue del lado transaccional las filas correspondientes a lo que se requiere.

Luego debido a la configuración por defecto del SSIS, se deben realizar una serie de conversiones, especialmente con el juego de caracteres y valores booleanos, esto se hace en esa parte y luego finalmente, se hace la inserción en la parte de almacenamiento temporal.

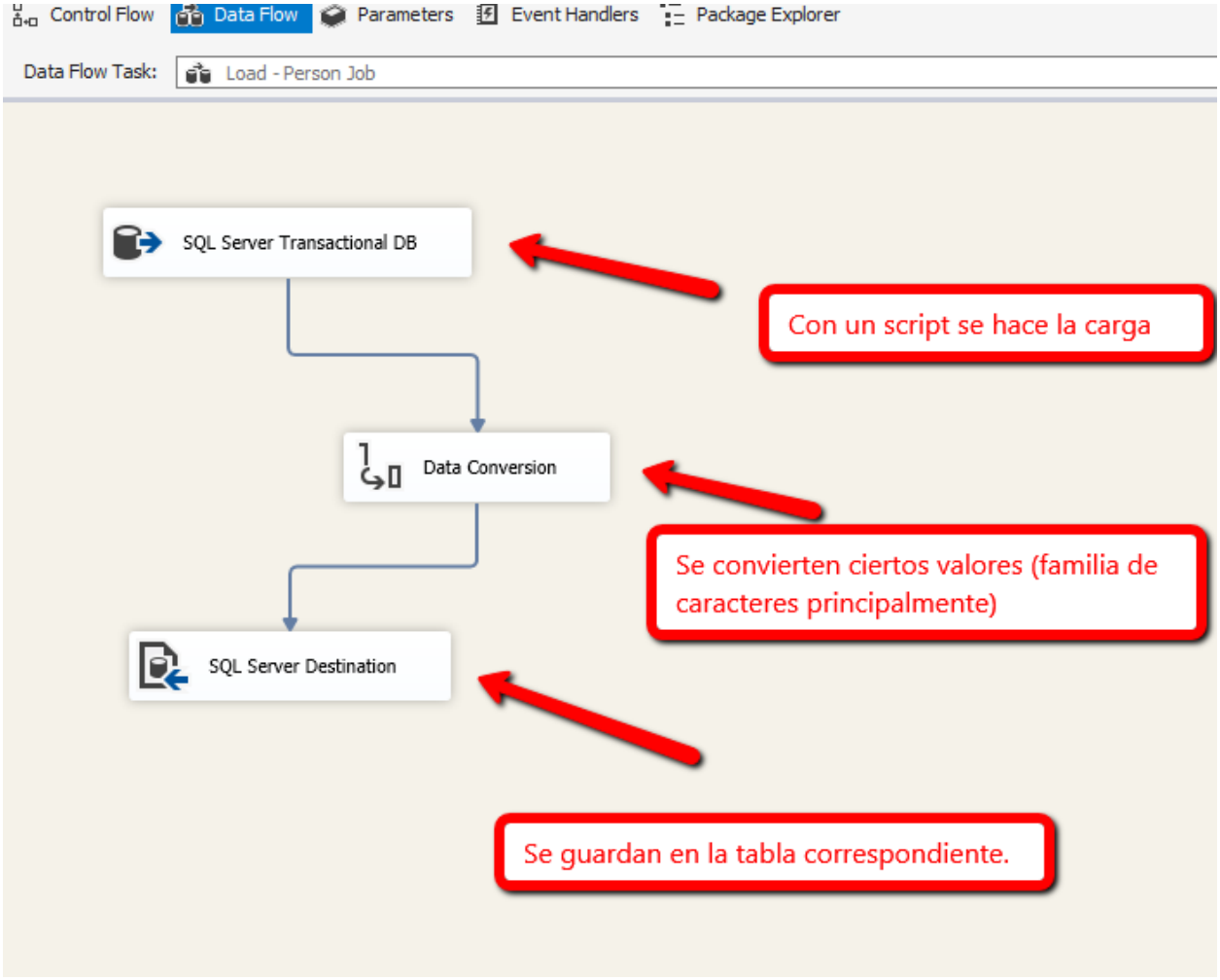


Figura 14 Especificación del proceso de carga de dimensión en el área de almacenamiento temporal. (Fuente: Elaboración propia).

Diseño del área de almacenamiento multidimensional.

El proceso principal que es manejado por el almacén de datos son las transacciones sobre los préstamos y ahorros de las personas del sistema. Esto hace que el proceso sea alrededor de la tabla “loan_loanoperationth”.

Conexiones. Se optó por usar solo un tipo de origen y de destino. El origen va ser las tablas del área temporal y el destino será el almacén de datos. En la imagen siguiente se muestran ambas conexiones, las cuales solo requieren actualizarse de acuerdo a la máquina donde se va a correr.

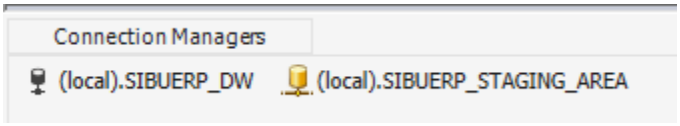


Figura 15 Conexiones para el proceso de carga al almacén de datos. (Fuente: Elaboración propia).

Estructura del almacén de datos. Se presentan parte de los objetos existentes en esta base de datos y que se consideran relevante para la comprensión del mismo.

Se ha utilizado el modelo estrella, el cual implica tener una tabla de hechos y alrededor de ella todas las tablas de dimensión que se consideren necesarias o relevantes. En el caso de SIBUERP_DW se tiene la tabla de hechos de las operaciones de préstamos y esta se ha denominado "loan_operation_fact". (Ver diagrama completo en el anexo B)

Tabla de hechos. Contiene los siguientes campos

- Loan_operation_id_or: que es el identificador de la operación del lado transaccional.
- Loan_operation_id: es el identificador de la operación de préstamo o ahorro per-se, no la transacción. Esta tiene relación con la tabla de dimensión loan_operation.
- Collect_status_id: hace referencia a la tabla de dimensión de collection_status_dim.
- Currency_type_id: es la relación con la tabla de dimensión que representa el tipo de moneda de la operación.
- Montly_payment_type_id: esto fue para efectos de des normalizar la tabla, va ligado principalmente con la operación per-se.
- Periodicity_type_id: parecido con lo anteriormente dicho, indica la periodicidad del préstamo y ahorro.

- Guarantee_category_id: relacionado con el tipo de garantía ofrecida para el préstamo.
- Ordinary_payment_interest_amount: el monto de pago relacionado con el interés ordinario.
- Transaction_amount: el monto de la transacción.
- New_balance: el nuevo balance del proceso de préstamo y ahorro.
- Amortization_amount: el monto de amortización.
- Date_id: el identificador de la fecha que está asociado con la transacción.
- Customer_id: el identificador del usuario con el que se asocia este préstamo.
- Customer_region_id: el identificador de la región del usuario con el que se asocia con el préstamo.
- Customer_occupation_id, el identificador de la ocupación con la que está asociado el usuario de este préstamo.
- Has_Additional_fees, es un campo calculado que indica si hubo pagos extras realizados.
- Additonal_fees_amount, el monto de pagos adicionales hechos.
- Loan_type_id: el tipo de préstamo con el que se asocia el préstamo.
- Late_payment_interest_amount: el monto de pagos extra por intereses debido a atrasos.

Tabla de dimensión de préstamos. Esta está denominada como "loan_operation_dim" y cuenta con las siguientes columnas

- Loan_operation_id: el identificador de la operación de préstamo del almacén de datos.
- Loan_operation_id_or: es el identificador que está en la base de datos transaccional.
- Guarantee_count: el número de garantías asociadas a este préstamo.
- Balance: el balance del préstamo.
- Late_pay_annual_interest_rate: el monto de pago atrasado que se debe hacer en dichos casos.
- Principal_amount: es el monto principal de la operación.

- Month_therms: es la cantidad de meses en el que se planea va a durar este préstamo.
- Montly_payment_flat: el monto plano a pagar por mes.
- Disbursed_amount: es el monto dado.
- Interest_rate_type: es el tipo de interés.
- Loan_type_name: el nombre del tipo de interés.
- Currency_name: el nombre del tipo de moneda del préstamo.
- Periodicty: el nombre del tipo de periodicidad.
- Payment_type: el nombre del tipo de pago.

Esta tabla cuenta con valores que son particulares al préstamo y no a las distintas transacciones.

Dimensión de fechas. Esta tabla de dimensión tiene una granularidad hasta horas. Además cuenta con dos columnas extras calculadas para indicar si es mañana o si es fin de semana.

Tabla de dimensión de clientes. En esta tabla de dimensión, se busca tener información general acerca del usuario. Como se planteó para el proyecto, este no debía tener datos personales. Sin embargo se ha incluido información acerca de que si tiene vehículo, salario neto y total además de su estado. Específicamente

- Customer_id: el identificador a nivel de almacén de datos.
- Customer_id_or: el identificador sacado de la parte transaccional.
- Total_salary: el salario total del cliente.
- Net_salary: el salario neto del cliente, ya con las deducciones.
- Has_vehicle: es un indicador de que si tiene o no vehículo.
- Status_name: el estado actual del usuario.

Dimensiones de filtrado. Se han denominado así puesto que de lo pretendido en el proyecto, estas actuarían como elementos de filtrado. En esta categoría se tienen las tablas siguientes:

- Collection_status_dim
- Currency_type_dim
- customer_occupation_dim
- customer_region_dim
- guarantee_category_dim
- loan_type_dim
- paytment_type_dim
- periodicity_type_dim

Proceso de extracción, transformación y Carga. El proceso de la base de datos multidimensional se puede resumir de la siguiente manera:

- Cargar los datos del almacenamiento temporal necesarios y mapearlos según corresponda en las diferentes dimensiones.
- Hacer la carga del área temporal y calcular los distintos valores para las dimensiones.

En la siguiente imagen se detallan las partes para hacer la carga de una dimensión de la base de datos multidimensional. Estas siguen una serie de pasos que se repiten en la mayoría de las dimensiones.

Las tareas específicas necesarias para cargar una dimensión se resumen de la siguiente forma

- Se cargan los datos del área temporal a través de una consulta SQL.
- Se convierten columnas que en su gran mayoría son por problemas para reconocer caracteres Unicode, sin embargo se incluyen columnas booleanas que fueron reconocidas como enteros.

- Se remueven registros incorrectos, estos son los que tienen campos específicos nulos. Estos varían de tabla de dimensión en tabla de dimensión. Los que no cumplen estas condiciones se consideran como no correctos.
- Capturar registros que ya están registrados actualmente y descartarlos del proceso.
- Hacer el registro de los datos.

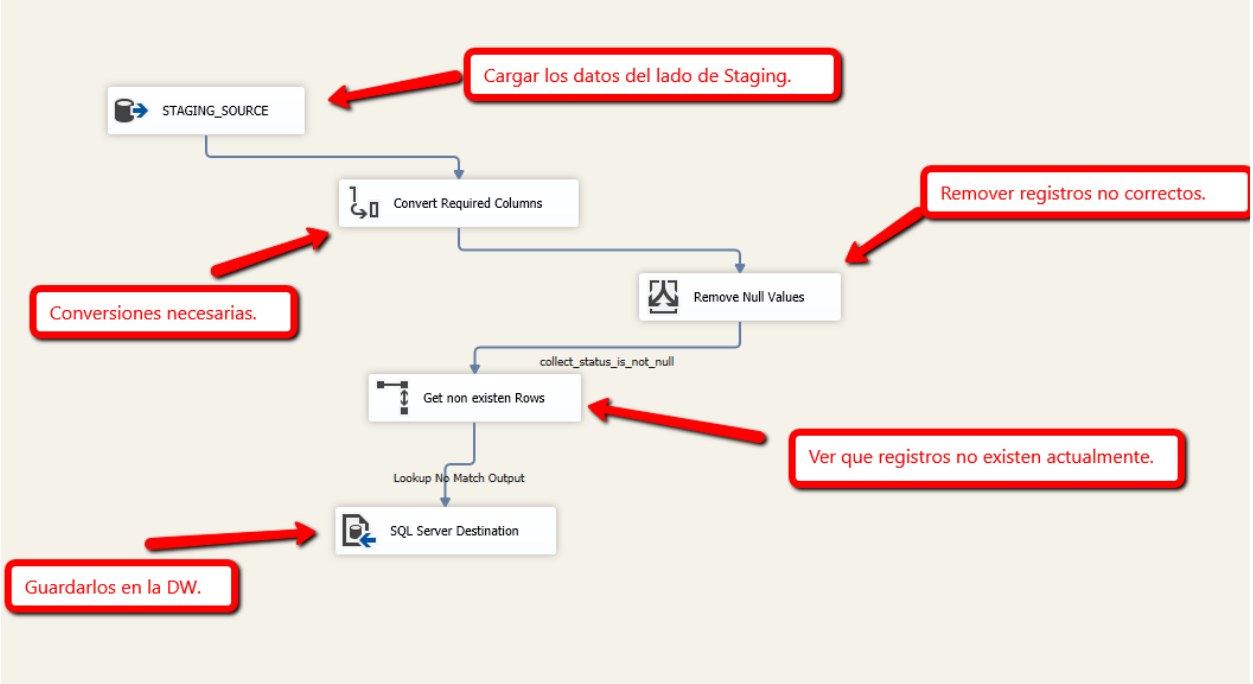


Figura 16 Carga de dimensión en la base de datos multidimensional. (Fuente: Elaboración propia).

Ya propiamente en el registro de la tabla de hechos el proceso varía de la siguiente forma.

- Se hace la carga de lo que se pretende cargar del temporal.
- Inicialmente se realiza una conversión de ciertas columnas que al igual que con la parte de las dimensiones se deben convertir para poder hacer uso de ellas.
- Luego de allí se hace un proceso de búsqueda de cada una de las llaves "foráneas" en las tablas de dimensiones. Ya para este momento estas deben ser ya existentes.
- Los registros que no sean encontrados en la tabla de dimensiones son descartados de la carga. Y serán analizados por otro proceso de carga aparte.

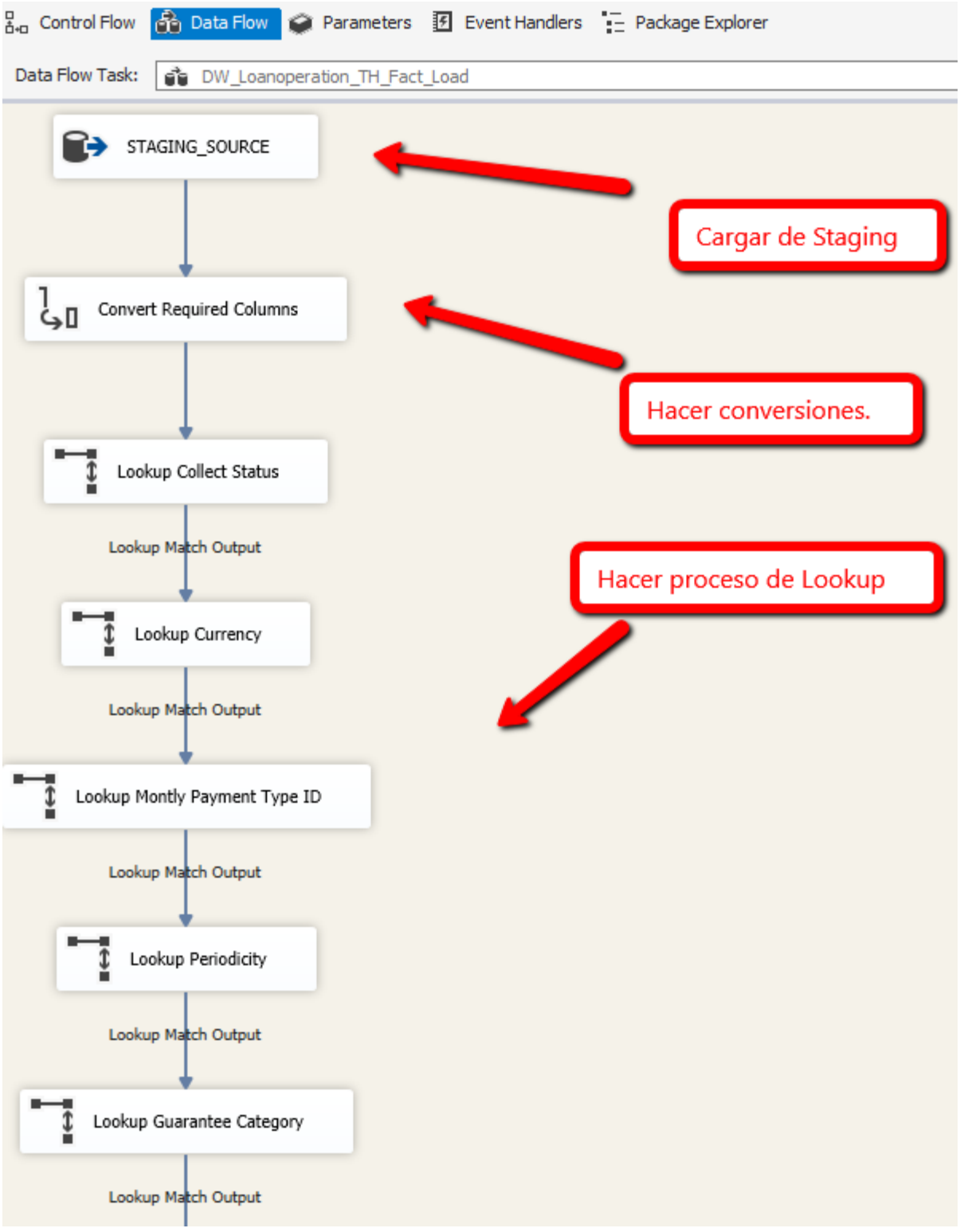


Figura 17 Proceso de Búsqueda para la carga de la base de datos multidimensional. (Fuente: Elaboración propia).

Esto ocurre por todas las dimensiones, hasta llegar propiamente al registro de los datos. Antes de esto se realiza el siguiente proceso.



Figura 18 Proceso de verificación y carga en la base de datos multidimensional. (Fuente: Elaboración propia).

Como se muestra en la imagen anterior, primeramente se valida que registros no existen y cuales si, los que no existen se guardan, el resto se descartan.

Diseño del ambiente analítico

Archivos de solución. Para efectos de SSAS se ha desarrollado una solución independiente dentro de la categoría de inteligencia de negocios que viene integrada al paquete de Microsoft.

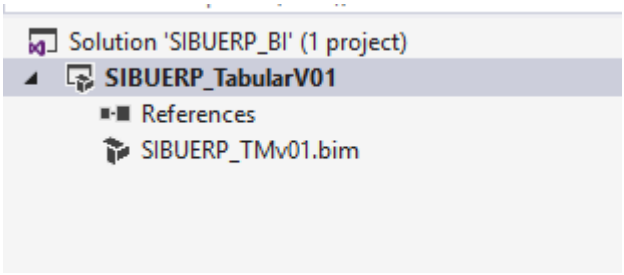


Figura 19 Archivos de solución para la parte analítica. (Fuente: Elaboración propia).

Configuración en formato tabular. Para poder soportar este Proyecto se ha tenido que configurar el motor de SSAS hacia el formato tabular. En la siguiente imagen se aprecia dicho cambio y como se visualiza.

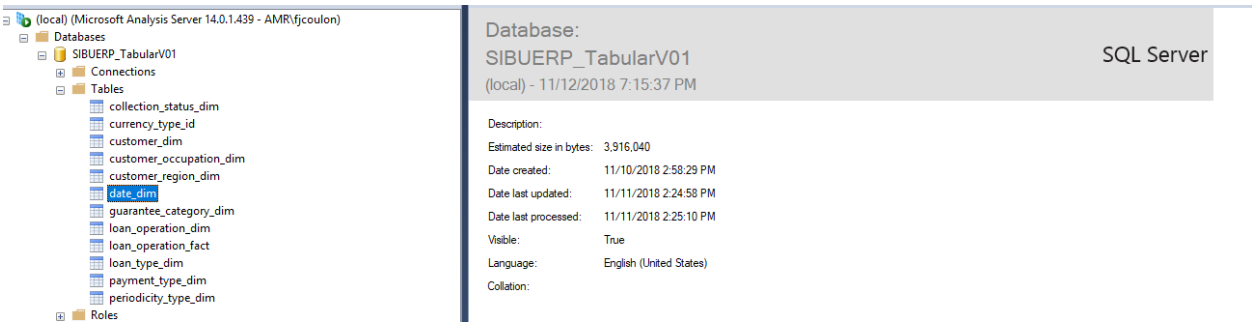


Figura 20 Servidor en formato tabular. (Fuente: Elaboración propia).

Luego de haber creado y cargado la fuente de datos, la cual es el almacén de datos, se visualiza las siguientes secciones en el software.

- En la primera parte, se aprecia una barra de fórmula, la cual cambia de acuerdo en la columna donde se esté posicionado.
- Se tiene los registros que se encuentran en la tabla.

- La solución y sus archivos, que para efectos de proyecto solo tiene un archivo .bim., por debajo se crea una serie de archivos más para guardar los metadatos y datos debido al modo seleccionado.
- En la parte inferior se observa las “tablas” como si fueran pestañas.
- Las columnas que se han ocultado, están con un color gris como si estuviera “deshabilitado”.
- Igualmente las columnas se han renombrado a un formato legible.
- En la imagen siguiente, como es la dimensión de fechas, se puede ver una serie de columnas calculadas.

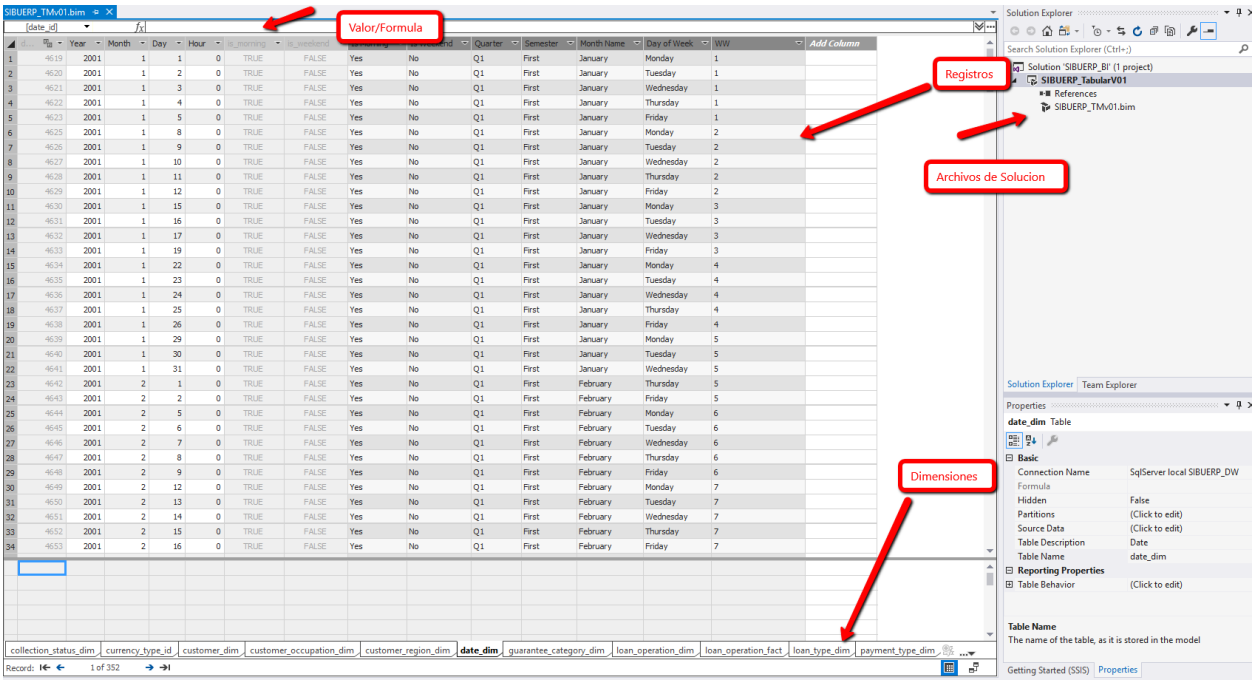


Figura 21 Estructura de los registros dentro del proyecto analítico. (Fuente: Elaboración propia).

Dimensiones. Para este caso, se decidió hacer dos jerarquías en la dimensión de fechas, una para lo que son valores numérico y otro para lo que es formato textual.

Igualmente, se pueden apreciar las columnas que han sido ocultadas, de hecho se observa una denominada “is_morning”, la cual ha sido ocultada y ahora la denominada “Is Morning” en formato textual es la que está disponible para los clientes.

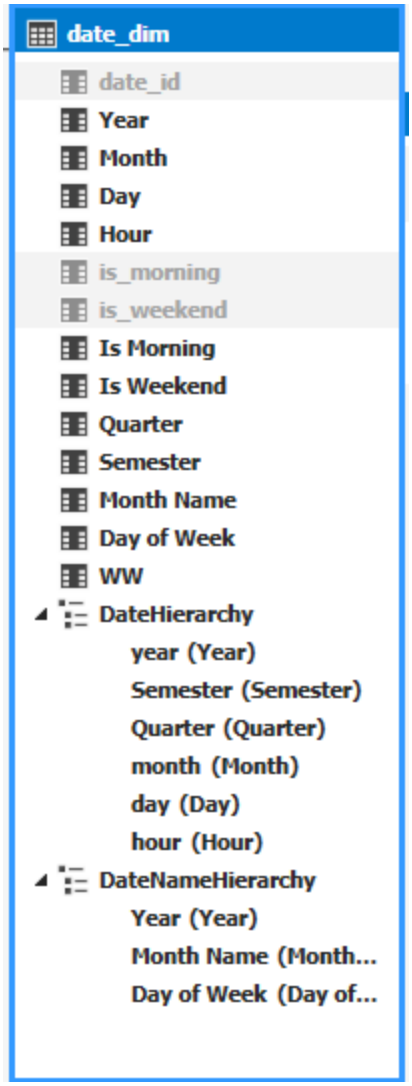


Figura 22 dimensión de fechas para la solución. (Fuente: Elaboración propia).

Tabla de hechos. En la tabla de hecho hay varias columnas relevantes para lo que son cálculos. Esto basado en el objetivo de la base de datos analítica realizada. Se ha dejado lo que tiene que ver con el monto, sus extras y relaciones con respecto a la deuda principal. Además se muestra varias columnas calculadas pero sobre todo, hay medidas a partir de todos los registros.

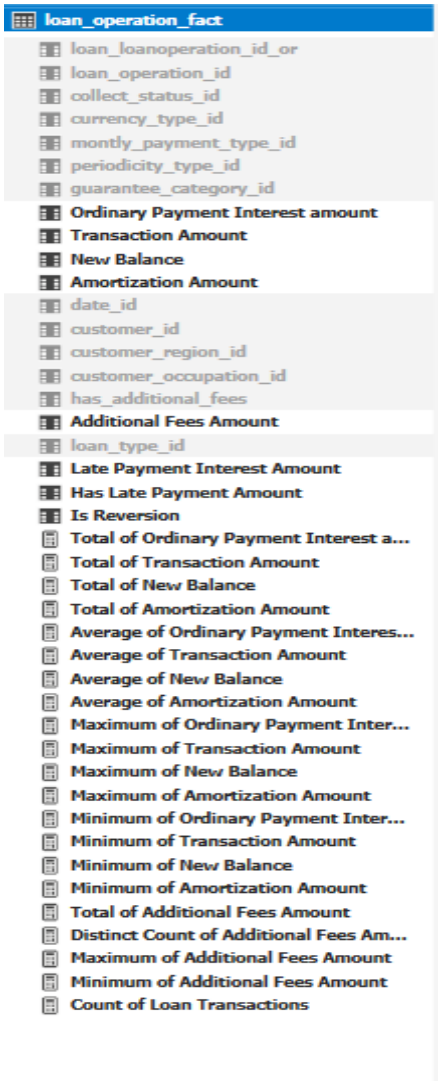


Figura 23 Tabla de hechos. (Fuente: Elaboración propia).

Traducciones. La traducción hecha fue de inglés para español, esta se usó usando BIDS, el complemento para Visual Studio y su objetivo es facilitar al usuario el uso de las visualizaciones.

BIDS Helper Tabular Translations Editor

Table	Object Type	Property	English (United States)	Spanish (Costa Rica)
customer_dim	Column	Caption	Has Vehicle	Tiene Vehículo
customer_dim	Measure	Caption	Sum of Total Salary	Total de Salario Total
customer_dim	Measure	Caption	Average of Total Salary	Promedio del Salario Total
customer_dim	Measure	Caption	Maximum of Total Salary	Máximo de Salario Total
customer_dim	Measure	Caption	Minimum of Total Salary	Mínimo de Salario Total
customer_dim	Measure	Caption	Count of Total Salary	Número de Salario Total
customer_dim	Measure	Caption	Distinct Count of Total Salary	Número de distintos Salario Total
customer_dim	Measure	Caption	Total of Net Salary	Total de Salario Neto
customer_dim	Measure	Caption	Average of Net Salary	Promedio de Salario Neto
customer_dim	Measure	Caption	Maximum of Net Salary	Máximo de Salario Neto
customer_dim	Measure	Caption	Minimum of Net Salary	Mínimo de Salario Neto
customer_occupation_dim	Table	Caption	customer_occupation_dim	Ocupación
customer_occupation_dim	Table	Description	Customer Occupation	Ocupación
customer_occupation_dim	Column	Caption	Name	Nombre
customer_region_dim	Table	Caption	customer_region_dim	Región del cliente
customer_region_dim	Table	Description	Customer Region	Región del cliente
customer_region_dim	Column	Caption	Name	Región
date_dim	Table	Caption	date_dim	Fechas
date_dim	Table	Description	Date	Fecha
date_dim	Column	Caption	Year	Año
date_dim	Column	Caption	Month	Mes
date_dim	Column	Caption	Day	Dias
date_dim	Column	Caption	Hour	Hora
date_dim	Column	Caption	Is Morning	¿Es Mañana?
date_dim	Column	Caption	Is Weekend	¿Es Fin de Semana?
date_dim	Column	Caption	Quarter	Trimestre
date_dim	Column	Caption	Semester	Semestre
date_dim	Column	Caption	Month Name	Mes
date_dim	Column	Caption	Day of Week	D. Semana
date_dim	Column	Caption	WW	WW
date_dim	Hierarchy	Caption	DateHierarchy	Jerarquía Fechas
date_dim	Level	Caption	year	Año
date_dim	Level	Caption	Semester	Semestre
date_dim	Level	Caption	Quarter	Trimestre
date_dim	Level	Caption	month	Mes
date_dim	Level	Caption	day	Día
date_dim	Level	Caption	hour	Hora
date_dim	Hierarchy	Caption	DateNameHierarchy	Niveles de Fecha
date_dim	Level	Caption	Year	Año
date_dim	Level	Caption	Month Name	Mes
date_dim	Level	Caption	Day of Week	Día de la semana
guarantee_category_dim	Table	Caption	guarantee_category_dim	Categoría de garantía
guarantee_category_dim	Table	Description	Guarantee	Garantía
guarantee_category_dim	Column	Caption	Name	Código
guarantee_category_dim	Column	Caption	Code	Nombre
loan_operation_dim	Table	Caption	loan_operation_dim	Operación de crédito
loan_operation_dim	Table	Description	Loan Operation	Operación de crédito
loan_operation_dim	Column	Caption	Guarantee Count	Cantidad de garantías
loan_operation_dim	Column	Caption	Balance	Balance
loan_operation_dim	Column	Caption	Late payment anual interest rate	Cargo por pago tardío anual

Figura 24 Traducciones para la solución analítica con el uso del complemento BIDS. (Fuente: Elaboración propia).

Diseño de las visualizaciones

En las siguientes líneas se hace referencia a la elaboración de los distintos reportes iniciales para el proyecto. Las visualizaciones aquí mostradas conforman el conjunto inicial de reportes requeridos para la solución.

Se ha decidió utilizar dos tecnologías diferentes para poder comprarlas. De un lado se tuvo Tableau y por otro PowerBI. De aquí que esta sección se divide en tres partes importantes. La primera se expone lo que se quiere mostrar, la segunda y tercera parte indican lo hecho utilizando las tecnologías de visualización seleccionadas.

Las visualizaciones están relacionadas a los resultados de las operaciones del departamento principalmente. Se busca mostrar resultados de los créditos realizados de acuerdo al tipo, igualmente esos mismos créditos pero por tipo de moneda y tipo de inversión. De allí que se tenga un primer tablero que se ha denominado "Resumen de operaciones", el cual tiene como objetivo responder a preguntas tales como

- ¿Cuáles son los tipos de préstamos más registrados?
- ¿Cuál es el número de préstamos que hay por cantidad de garantías presentadas?
- ¿Cuántos préstamos se hace por tipo de interés?
- ¿Cuántos préstamos se hacen por tipo de periodicidad?
- ¿Cuántos préstamos se hacen por tipo de moneda?

En un segundo tablero, lo que se busca es poder visualizar las transacciones desde un punto de vista holístico. De allí que se ha denominado "resumen de transacciones" y tiene como objetivos responder a preguntas tales como:

- ¿Cuántas transacciones hay por tipo de ocupación?
- ¿Cuántas transacciones hay por fechas (año, mes)?
- ¿De todas estas transacciones, cuántas son reversiones y cuáles montos representan cada uno de ellas?
- ¿Monto de amortizaciones en dichas transacciones?

En un tercer tablero el objetivo es ya ver cantidades de tiempo y de dinero. De allí que se optó por un nombre como "resumen de monto de créditos". En él se busca responder preguntas como

- ¿El término máximo de meses para pagar el préstamo y su tipo?
- ¿Cantidad de dinero entregado por cantidad de garantías?
- ¿Cuál es el pago mensual plano por periodicidad?

Cabe destacar que las tres visualizaciones, se van a realizar usando las dos tecnologías, estas se describen en las siguientes dos subsecciones.

Visualizaciones con Tableau. Con esta herramienta, se logró hacer múltiples tableros. Se utilizó una estrategia en la cual se hacen primero los reportes particulares, posterior a ello se pasa a una segunda etapa en donde se unen esas visualizaciones en tableros y por último a estos tableros se asignan/agregan acciones, que resultan en filtros.

Resumen de operaciones. Este primer tablero consiste en mostrarle al usuario la cantidad de créditos realizados por tipos, número de garantías, la moneda, periodicidad y tipo de inversión. El mismo está dirigido para mostrar cuales son los tipos de créditos con más auge durante todo el tiempo.

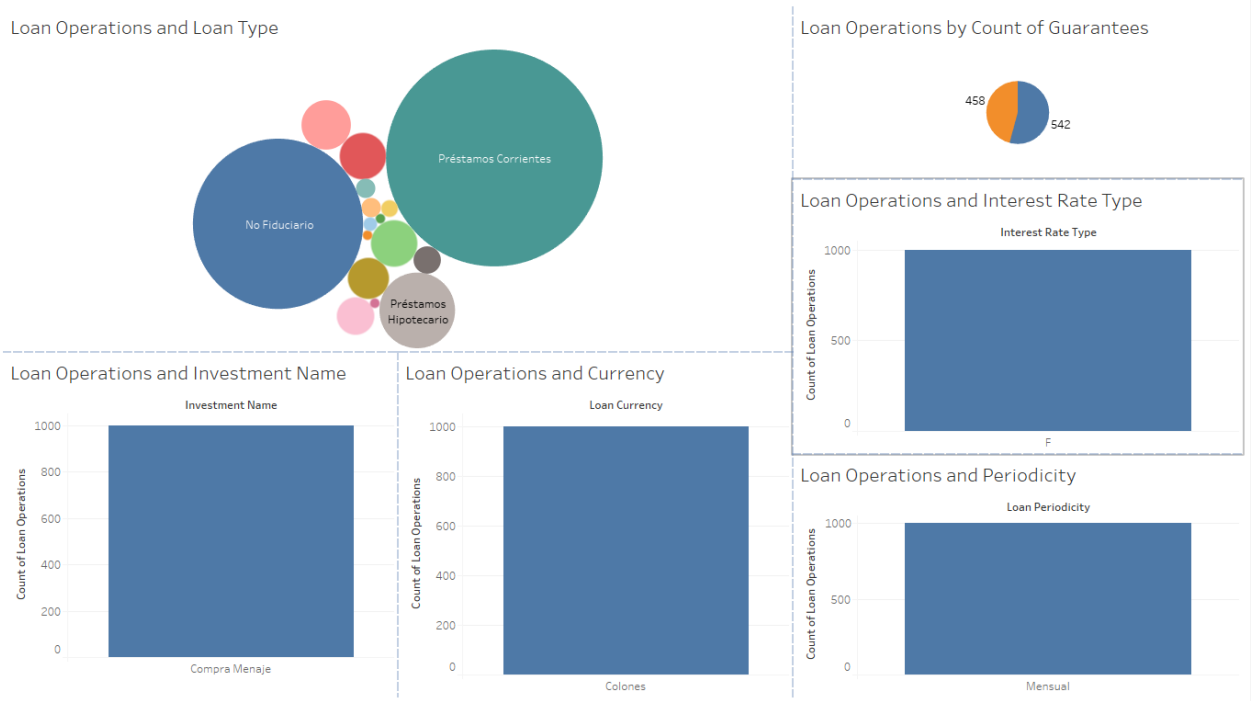


Figura 25 Tablero resumen de operaciones. (Fuente: Elaboración propia).

Resumen de transacciones. Este tablero está enfocado en las operaciones del departamento de crédito. El mismo busca mostrar las transacciones realizadas por tipo de ocupación, por años y pretende mostrar la cantidad de reversiones realizadas en esos periodos de tiempo.

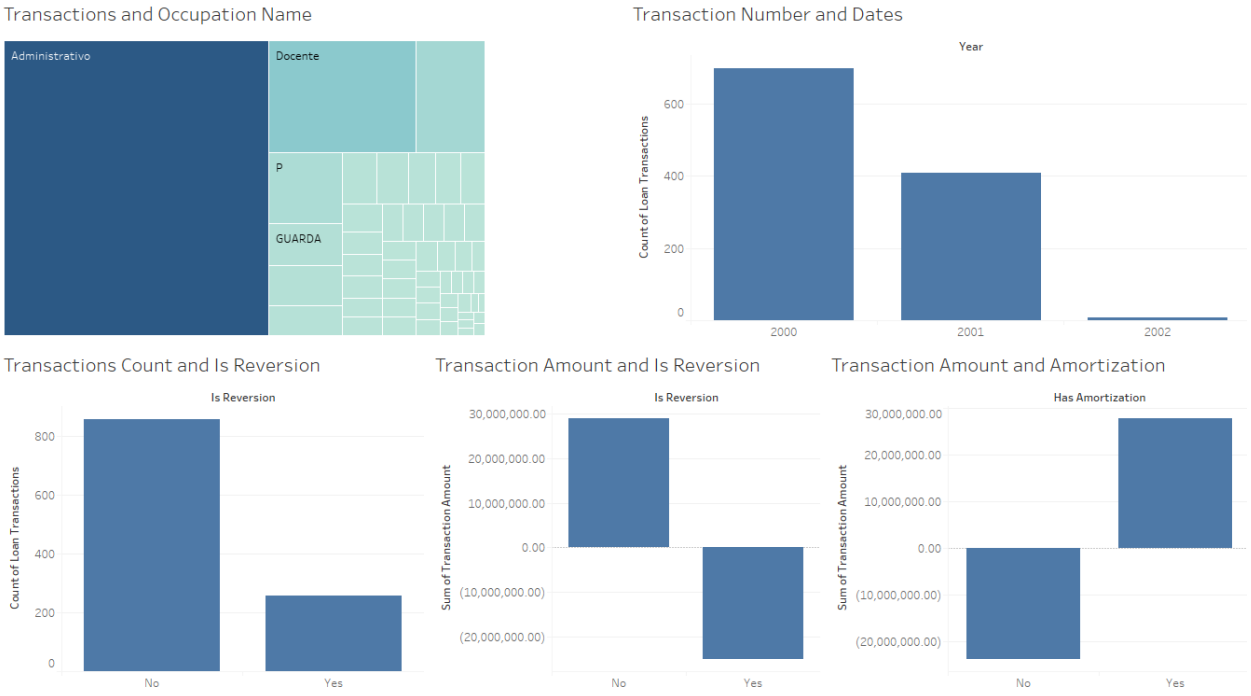


Figura 26 Tablero resumen de transacciones. (Fuente: Elaboración propia).

Resumen de monto de créditos. Este está conformado por cuatro distintas visualizaciones. El objetivo del mismo es poder mostrar la cantidad de dinero entregado por tipos garantía, además cual es el mayor espacio de tiempo que se da para poder pagar los préstamos.

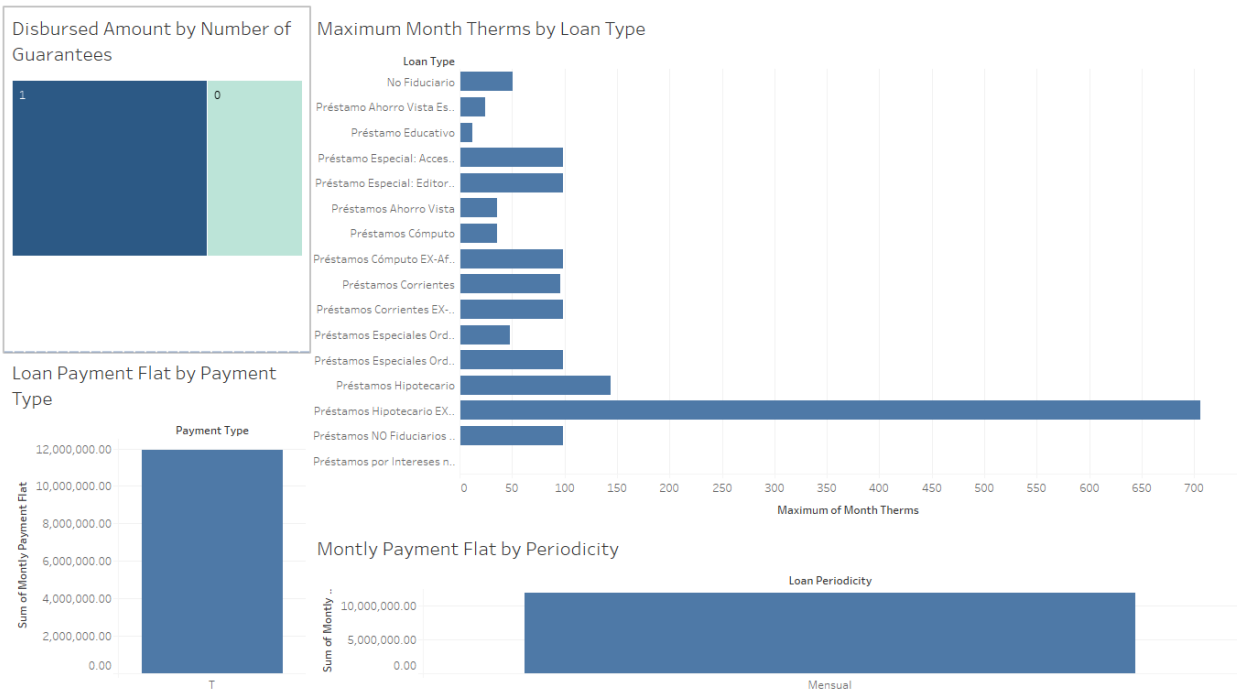


Figura 27 Tablero resumen de los montos por tipo de transacción. (Fuente: Elaboración propia).

Visualizaciones con PowerBI. Utilizando esta herramienta, a diferencia de Tableau, se pudo generar los diversos tableros de manera directa, es decir, no se debió hacer ningún tipo de reporte y unirlos. En las siguientes secciones, se muestran las diversas salidas generadas, que se asimilan a lo realizado con las visualizaciones de Tableau, debido al objetivo de comparación.

Resumen de operaciones. Este primer tablero consiste en mostrarle al usuario la cantidad de créditos realizados por tipos, número de garantías, la moneda, periodicidad y tipo de inversión. El mismo está dirigido a mostrar cuáles son los tipos de créditos con más auge durante todo el tiempo.

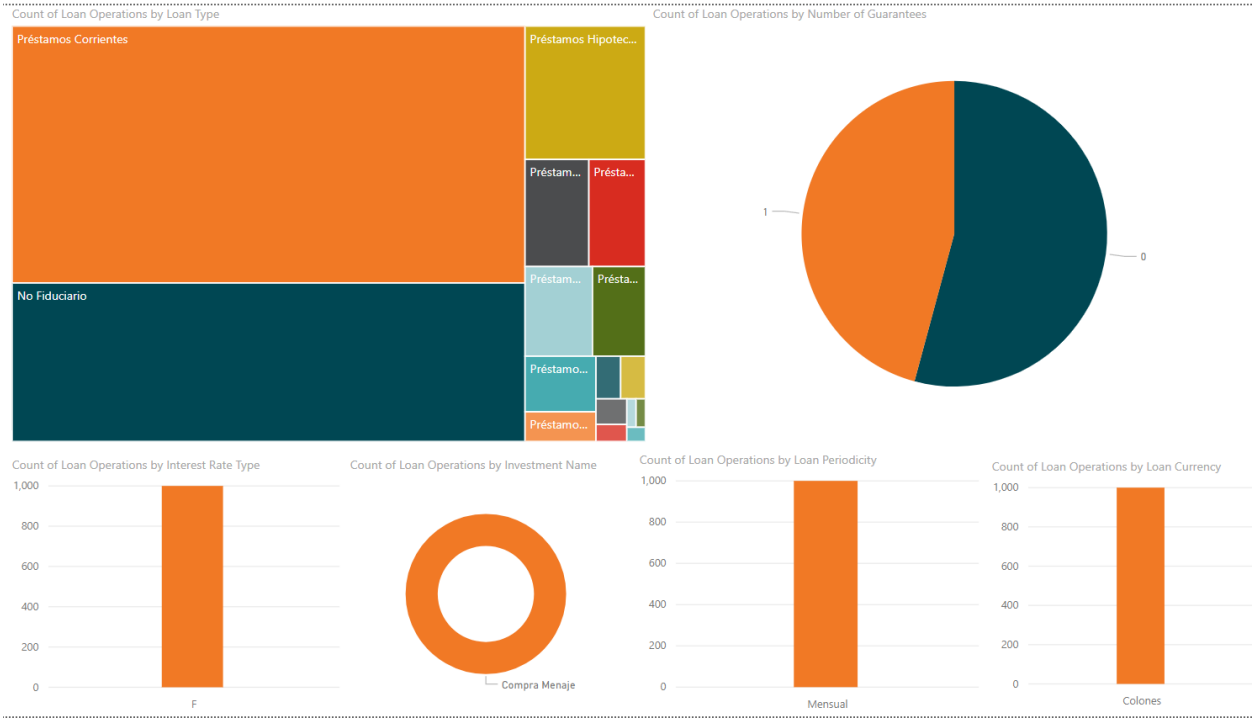


Figura 28 Resumen de operaciones con PowerBI. (Fuente: Elaboración propia).

Resumen de transacciones. Este tablero está enfocado en las operaciones del departamento de crédito. El mismo busca mostrar las transacciones realizadas por tipo de ocupación, por años y pretende mostrar la cantidad de reversiones realizadas en esos periodos de tiempo.

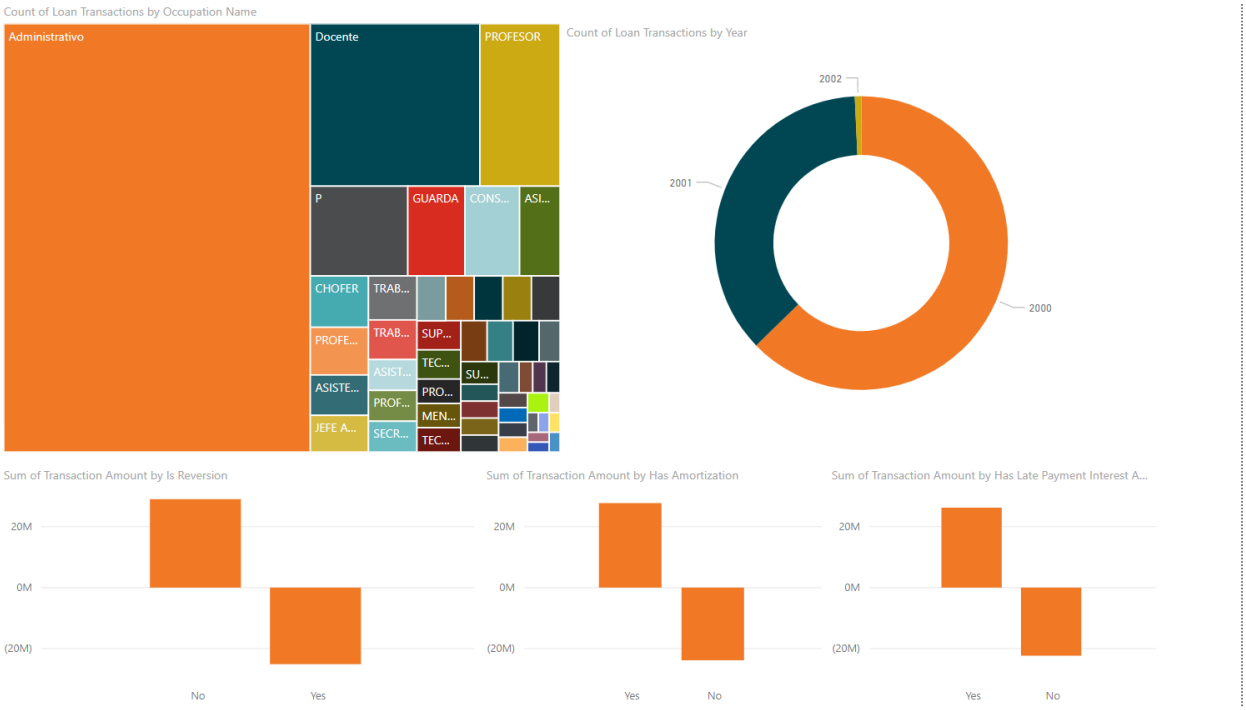


Figura 29 Resumen de transacciones usando PowerBI. (Fuente: Elaboración propia).

Resumen de monto de créditos. Este está conformado por cuatro distintas visualizaciones. El objetivo del mismo es poder mostrar la cantidad de dinero entregado por tipos garantía, además cual es el mayor espacio de tiempo que se da para poder pagar los préstamos.

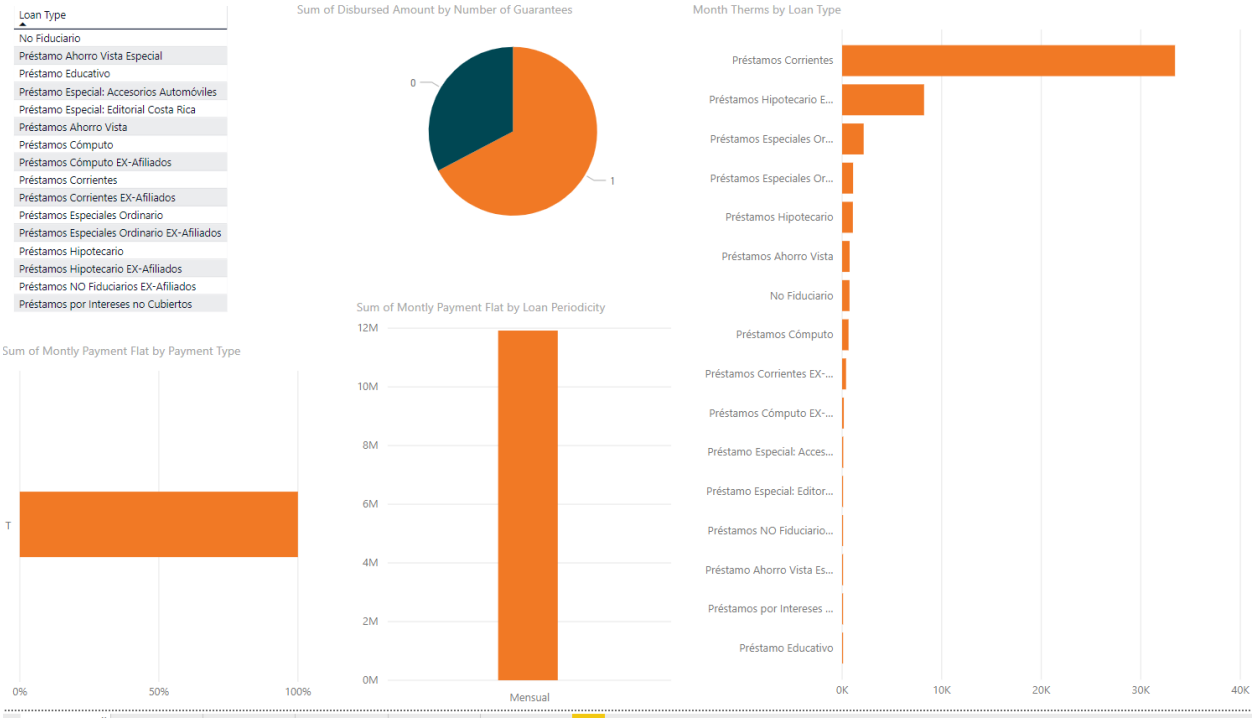


Figura 30 Tablero resumen de monto de créditos utilizando PowerBI. (Fuente: Elaboración propia).

Capítulo sexto. Conclusiones y Recomendaciones

En esta sección se presentan las conclusiones de la investigación aplicada; se indica cómo se han alcanzado los objetivos del objetivo general y específicos. En la segunda parte se expone las recomendaciones, que se orientan los aspectos que habrían favorecido al proceso como tal.

Conclusiones

Sobre el objetivo específico "Revisar el estado actual de los objetos de la base de datos existente que permitan la detección de oportunidades de mejora en su estructura.", se concluye que

- Después de un análisis de la base de datos, la misma cuenta con varias oportunidades de mejora. Una de ellas es existen registros muy antiguos, que forman parte del área transaccional y podrían ser desplazados o archivados en otro lugar, lo que se traduciría en mejoras del rendimiento. El no poder hacer esto implica que la base de datos multidimensional va a contener datos que son antiguos y poco relevantes para la actualidad que además requieren más capacidad de procesamiento.
- Existen objetos de datos que cuentan con falta de estandarización. La base de datos ha sufrido modificaciones a lo largo de tiempo, haciendo que tablas tengan columnas con nombres en diferente idioma y no haya estándares. Esto para el trabajo implicó ciertos atrasos especialmente con la comprensión de ciertos objetos.
- Se comprobó que existen una gran cantidad de tablas que no están siendo utilizadas. Esto se debe a que la estructura no es original para este tipo de departamentos, sino es una adaptación. Una oportunidad de mejora es la eliminación de objetos de datos que no son usados.
- Las encuestas realizadas a los miembros del departamento demuestran que el rendimiento de la base de datos en ocasiones no es el mejor y además que el proceso para extraer datos es complicado, lento y requiere coordinación, aspecto que en varias situaciones es difícil de alcanzar.

- El considerar un proyecto de reestructuración en conjunto con inteligencia de negocios, hace de un proyecto, algo muy costoso. Esto es en ocasiones requeridos debido al análisis de la calidad de los datos existentes.
- Además, si bien que esta asociación es de unos 10 mil afiliados, la infraestructura actual hace que haya procesos bastantes largos, que toman alrededor de dos horas. Al ser una organización pública, algún cambio de esta índole es realmente difícil de ser realizado debido al tema burocrático. Esto implica que con los recursos disponibles se debe determinar qué está fallando de manera manual, analizando procesos y esta situación no se visualiza que cambie a corto plazo.

Sobre el objetivo específico "Documentar los objetos de datos existentes que sean compatibles con la solución de inteligencia de negocios planteada." Se concluye que

- Solo una pequeña parte de las tablas se requiere para dar solución de inteligencia de negocios del proyecto. En la sección de diseño de la arquitectura de la solución se detalló la lista de elementos utilizados.
- Existen objetos de datos con nombres similares, los cuales tienden a confundir a los desarrolladores.

Para con el objetivo "Elaborar la documentación base para la solución de inteligencia de negocios para que pueda ser utilizado en toma de decisiones, reestructuraciones y solución de problemas de manera eficiente." Se concluye que

- La documentación realizada para este proyecto es suficiente para poder hacer extensiones a lo hecho.
- Existen objetos de bases de datos que son complejos de documentar debido a que no son conocidos por los que lo usan o no son utilizados.
- El problema con que se contó durante la mayor parte del proyecto fue el hecho que las personas no conocen la base de datos, la documentación es nula o mínima, y esto reduce la capacidad de poder avanzar con la rapidez que uno se plantea.

Sobre el objetivo “Desarrollar un cubo multidimensional que permita dar respuestas a múltiples consultas planteadas por los usuarios.”

- Tal y como se ha planteado en la sección de alcance del proyecto, se ha realizado un cubo multidimensional con una serie de visualizaciones mínimas requeridas para la prueba de concepto.
- La implementación de un proyecto de este tipo, implicaría múltiples beneficios para los usuarios del área funcional. Ayudaría específicamente a la toma de decisiones y sobre todo la visualización de las actividades actuales del departamento. Tal como se mostró durante este proyecto, uno de los activos más importantes con los que cuenta una organización es la información ya que se va recolectando con el tiempo y el uso de los mismos va a decir que tantos beneficios se puede encontrar.
- El proceso realizado no requerido de ningún equipo especial, por lo que desde un punto de vista económico, no representa gastos para la organización y los recursos que implican mayor inversión fue el tiempo y la implementación del mismo.

Sobre el objetivo de “Evaluar el prototipo elaborado en un ambiente de desarrollo, con datos reales de la organización para medir el nivel de efectividad y el rendimiento.”

- Se comprobó que la técnica de deshabilitar los índices de la base de datos durante el procesamiento, colabora con el rendimiento en procesos largos, porque evita muchas tareas de validación y creación innecesarias durante la ejecución del mismo. Esto no significa que ha de evitarse el uso de los índices.
- La utilización de particiones en la base de datos analítica, colaboraría con el procesamiento de datos requeridos, esto especialmente debido al tamaño de la base de datos que se debería manejar. Esto permite que el gestor pueda cargar los datos de manera más adecuado, siempre dependiendo del modo cómo los usuarios ingresen sus consultas.
- La utilización de PowerBI y Tableau ayudó a comprobar las diferencias entre ambos sistemas y hacia dónde se orienta cada una. Por un lado, se tiene la de PowerBI, la cual es conocida por la mayoría del departamento, que su uso es

menos complejo y el costo, dependiendo del modo deseado es menor (a través de la utilización del sistema Desktop).

- Se comprobó que los usuarios no conocen la herramienta de Tableau, lo que hace de este sistema el menos factible para efectos de proyecto. Por otro lado PowerBI es la más adecuada debido a que la versión desktop es gratuita y se requiere publicar en la nube, al tener una cuenta office 365 en el departamento ya se cuenta con una licencia de PowerBI.
- Se comprobó que en términos económicos que Tableau representa un costo mayor a PowerBI debido al modo de licenciamiento y tipos de aplicaciones existentes.
- La complejidad de la evaluación del sistema, se vio reducida en gran medida debido al poco control o nivel de rigidez en el control de datos. La limitante principal fue el tamaño de la base de datos en sí, que ascendía a las 40 GB.

Recomendaciones

Después de la elaboración del presente proyecto y su respectivo análisis de los datos recolectados, se plantean las recomendaciones y acciones estratégicas siguientes para los que están interesados en proyectos similares o relacionados.

La elaboración de proyectos en ambientes públicos, es complejo y burocrático. El presente proyecto solo es una propuesta, debido a que la implementación del mismo implica mucho más tiempo debido a las distintas aprobaciones que se necesitan. De aquí se recomienda entonces que si es posible hacer proyectos en otros ámbitos, sea privado o particular, este brindará más opciones de progreso y mejor adaptabilidad con lo planteado en fases iniciales.

Se recomienda la utilización de la herramienta PowerBI, esto debido a que Tableau implica pago de licencias. Si bien es cierto, existe la licencia de estudiantes, está para la elaboración de proyecto resulta complicado, debido a la coordinación con los usuarios clientes. En cambio PowerBI, cuenta con una versión completamente gratuita, que si bien tiene limitaciones y requieren ciertos trabajos manualmente, permite hacer

pruebas de manera más realista desde el punto de vista del usuario. Por otro lado, en temas de complejidad, Tableau termina siendo mayor que PowerBI.

También es importante no olvidar lo que son metodologías ágiles. En este proyecto, si bien no se indica explícitamente, las técnicas que este brinda permitió que el trabajo se pudiera realizar en los tiempos indicados y sin mayor inconveniente. Además se comprobó que no es solo para proyectos de codificación pero también los de cualquier área.

Se recomienda hacer una separación y traslado de datos antiguos de la base de datos, esto se podría complementar con particiones para mejorar el rendimiento de la base de datos. Por ejemplo, los datos generados previo al 2007 podrían ser enviados a un almacenamiento aparte, que solo es accedido cuando haya actividades de auditorías específicas. Con eso alrededor de 4 millones de filas podrían ser reducidas del procesamiento principal.

Con los objetos de datos que cuentan con falta de estandarización, se recomienda que estos sean estandarizados en idioma y en estilo. Si bien la base de datos ha sufrido modificaciones a lo largo de tiempo y continuará con ello, con un nivel de estandarización alto, futuros desarrollos podrán ser manejados con mayor facilidad.

Se recomienda hacer uso de la herramienta PowerBI para las visualizaciones de la base de datos multidimensional. Esto primeramente por el conocimiento de la misma, la cual es superior a Tableau. Desde el punto de vista económico y como objetivo del sistema, también se recomienda PowerBI a través del sistema Desktop.

Capítulo séptimo. Reflexiones finales.

Al inicio de esta investigación aplicada, se visualizó el producto final como un prototipo. Lo anterior debido a que el ambiente donde se iba a aplicar es sumamente burocrático y requiere de etapas más amplias para ser puestos en ambiente de producción.

Dada la naturaleza aplicada del proyecto, a lo largo de todo el trabajo se buscó poner en práctica diversas técnicas y buenas prácticas aprehendidas a lo largo de los diversos cursos de la maestría. Esto va desde el punto de vista de bases de datos, a través del análisis de los objetos, uso del lenguaje estructurado y la identificación de oportunidades de mejor hasta el uso de técnicas para la elaboración de proyectos de inteligencia de negocios.

Respecto a los propósitos del proyecto

Desde el inicio, si bien el proyecto estuvo orientado a una solución de inteligencia de negocios, y de lo analizado al inicio, estas terminan con la implementación de una base de datos multidimensional, para este proyecto se consideró ir un paso más adelante y se ha incluido la elaboración de visualizaciones iniciales. La razón principal es el hecho de que se considera que este tipo de soluciones deben tener una parte que incluya cómo los usuarios finales van a utilizar los componentes desarrollados. Estos pueden ser desde aplicaciones desde cero o utilizando tecnologías como PowerBI.

Respecto a la solución propuesta

Se considera que la solución propuesta está enmarcada en las necesidades de la organización. La misma ya después de haberla completado se pudo comprobar que los recursos disponibles en muchas ocasiones son suficientes para implementar o incluir nuevas ideas y cumplir nuevos objetivos. Esto es contrario a pensamientos que implica que una renovación ha de comenzar con la compra de nuevos equipos o nuevo personal. Si bien la solución está a nivel de prototipo, este es un producto mínimo aceptable para que se tenga una idea de los beneficios que traería este tipo de soluciones.

También, analizando lo que es el tipo de negocio y su misión, con el uso de la solución de inteligencia de negocios, estos se podrán cumplir de mejor manera, con información a la mano y conocimiento que pueda ser utilizado en la toma de decisiones.

Respecto a los retos del área

El incremento de los componentes móviles hace pensar acerca de nuevos tipos de proyectos, no solo los que tienen como objetivo finales los computadores. El poder obtener visualizaciones, información y conocimiento útil a través de dispositivos móviles hacen pensar que en un futuro cercano este tipo de aparatos han de ser incluidos de manera obligatoria en los proyectos de inteligencia de negocios.

El estar en contacto con la inteligencia de negocios y sus diversas tecnologías disponibles, hacen pensar que en un futuro cercano los retos principales no va ser el encontrar tecnología para dichos proyectos sino el poder hacer uso efectivo del mismo. Esto implica tener las habilidades técnicas, los medios de procesamiento y de conexión necesarios. También y no menos importante, las capacidades de planeación para llevar a cabo proyectos de este tipo. Si bien existen marcos de trabajo o estilos como el de Kimball que permiten dar respuestas a diversos aspectos de proyectos de inteligencia de negocios.

Capítulo octavo. Trabajos a futuro

En este capítulo muestra diversos posibles trabajos a futuro que se derivan de lo que se acaba de realizar. Estos van dirigidos principalmente a aquellos que se hacen investigaciones aplicadas o de campo y también los que prefieren estar desde un punto de vista más conceptual. En esta sección se hace análisis a partir de distintas áreas de trabajo o conocimiento.

Afinamiento de la base de datos

El proyecto no incluye el refinamiento de la base de datos. Sin embargo es una oportunidad de mejor que se visualizó desde el inicio del proyecto. Este sería importante debido a que hoy en día el rendimiento no es el más óptimo para los usuarios. Por ejemplo

- Se podría hacer un proyecto que incluya la reestructuración de los objetos de bases de datos para eliminar aquellos que no son utilizados hoy en día.
- Incluiría el análisis de los índices y creación de particiones de datos para la mejora de los mismos.
- Creación de tareas programadas para el mantenimiento de la base de datos, desde sus objetos hasta el reporte temprano de posibles fallos en motor o equipos dedicados a dicha labor.
- Además se puede incluir la documentación respectiva para futuros trabajos.

Con este tema de afinamiento, también existen otros posibles proyectos que se fueron vislumbrando a lo largo del tiempo. Algunos de ellos se enfocan más en documentación, otros se enfocan más en la estructura de la base de datos como tal. En esta sección se han mostrado solo algunas, las posibilidades son muchas y con la combinación de ideas el número crece de manera importante.

Minería de datos

La herramienta propuesta en el trabajo no incluye el hacer el proceso de minería de datos debido a las implicaciones del mismo y el nivel del desarrollo de este tipo de tecnologías en el departamento. Sin embargo, se logró visualizar que el proceso de negocios manejado puede ser utilizado en proyectos de minerías de datos. Más específicamente sería:

- Tener una solución que logre manejar y limpiar de acuerdo a lo necesario en el área de transacciones de préstamos.
- A partir de ello, sacar características que ayuden a tomar mejores decisiones a futuro con respecto al manejo de negocio.

Esto abriría puertas a nuevas oportunidades de herramientas y procesos de negocio, porque con el conocimiento generado y el apoyo que brindaría al negocio surgirían nuevas necesidades de datos y por ende soluciones.

Rendición de cuentas

El departamento forma parte de una institución pública. Por ende, está siempre la limitante en que la rendición de cuentas tiene oportunidades de mejora. Desde este punto de vista, una idea que estaría orientada más al área de los que analizan aspectos legales en el manejo de datos sería evaluar maneras que las visualizaciones puedan ser publicadas o accesibles a través de la web. Esto implicaría tareas como:

- Revisar el tipo de datos. Si bien, en la actualidad, se conoce que existen conjuntos de datos no sensibles y sensibles no se tiene una clasificación exacta.
- Definir los datos que pueden ser clasificados y cuáles no. Esto para efectos legales.
- Crear una interfaz en la cual se pueda exponer los datos al público y puedan ser descargables para su análisis.
- Limpiar o preparar los datos para ser consumidos por terceros. Haciendo un tipo de maquillaje en ciertas columnas de acuerdo a lo requerido.
- Documentar los diversos procedimientos adoptados a lo largo de la ejecución del proyecto en cuestión.

Con esta idea se pretende que otros interesados en datos puedan utilizar los mismos y presentar nuevas ideas de análisis de datos.

Visualizaciones móviles

Utilizando las visualizaciones generadas con este proyecto, combinada con las futuras visualizaciones agregadas, se puede crear una aplicación web que utilice dichos gráficos y datos y hacerlo disponible a través de una interfaz móvil. Esto puede ser un

celular o una tableta. Con ello se podría disponibilidad datos de la junta a los asociados y mejorar el sistema de rendición de cuentas, tal como se indicó en secciones anteriores.

Con ello se pretende facilitar el acceso a los datos sin necesidad de requerir un computador y una aplicación. Lo que sí se debe considerar es que por el tipo de tecnologías, se deben buscar maneras de crear visualizaciones que no impliquen generación de gastos para el departamento ya que estos recursos probablemente no sean asignados a un lapso de tiempo corto.

Futuras líneas de investigación

Una de las razones por las cuales se adoptó este trabajo es que, en la actualidad, no existe mucho con respecto a la inteligencia de negocios y las instituciones públicas. En Costa Rica, este tipo de entidades son conocidas en el aspecto tecnológicas como las que se quedan atrás. Un tema que también queda de lado y que la idea surge luego de comenzado este proyecto es el estado actual de la minería de datos en el ámbito público.

Otras instituciones como el Instituto Meteorológico Nacional (IMN) no brindan datos a cualquier particular que soliciten; lo mismo sucede con entidades como la Universidad de Costa Rica. Una investigación puede ser el fundamento legal que existe para basar este tipo de pensamientos o decisiones. Además las posibles ventajas y desventajas que implicaría abrir los datos hacia el público en general.

Referencias

Batini, C., & Scannapieca, M. (2006). *Data quality dimensions. Data Quality: Concepts, Methodologies and Techniques*. Berlin New York: Springer.

Boselli, R., Cesarini, M., & Mezzanzanica, M. (2011). *Framework Guidelines to Measure the Impact of Business Intelligence and Decision Support Methodologies in the Public Sector.* De Proceedings of the European Conference on e-Government (pp. 107-115).

Boselli, R., Cesarini, M., & Mezzanzanica, M. (2011). Public service intelligence: evaluating how the public sector can exploit decision support systems. De XXI International RESER Conference "Productivity of Services NextGen–Beyond Output/Input". Fraunhofer Verlag, Hamburg, Germany.

Daneshvar Kakhki, M., & Palvia, P. (2016). *Effect of Business Intelligence and Analytics on Business Performance* [Archivo PDF]. Recuperado de https://pdfs.semanticscholar.org/bf07/d9d68769506e912342dd546f301e33dd5473.pdf

Ferrari, A., Russo, M., & Webb, C. (2012). Microsoft SQL Server 2012 Analysis Services: The BISM Tabular Model. Pearson Education. Sebastopol, CA: O'Reilly Media

Fink, L., Yogev, N., & Even, A. (2017). *Business intelligence and organizational learning: An empirical investigation of value creation processes* [Archivo PDF]. Information & Management, 54(1), 38-56. Disponible en https://parsproje.com/tarjome/modiriyat/638.pdf

Gangadharan, G. R., & Swami, S. N. (2004, June). *Business intelligence systems: design and implementation strategies.* [Archivo PDF] In Information Technology Interfaces, 2004. 26th International Conference on (pp. 139-144). IEEE. Disponible en http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.477.1449&rep=rep1&type=pdf

Ghazanfari, M. J. S. R. M., Jafari, M., & Rouhani, S. (2011). *A tool to evaluate the inteligencia de negocio of enterprise systems* [PDF]. Scientia Iranica, 18(6), 1579-1590. Disponible en https://www.sciencedirect.com/science/article/pii/S102630981100215X

Junta de Ahorro y Préstamo de la Universidad de Costa Rica (s.f.) *Conózcanos. Recuperado* el 07 de octubre del 2018, de http://www.jafapucr.com/Inicio/Con%C3%B3zcanos.aspx

López, R. G. (06 de 22 de 2015). *Business Performance Management: ¿Cómo este sistema puede beneficiar a tu empresa?* Conexionesan. Recuperado el 07 de octubre del 2018, de https://www.esan.edu.pe/conexion/actualidad/2015/06/22/business-performance-management-como-este-sistema-puede-beneficiar-a-tu-empresa/

Marjanovic, O. (2010, Enero). *Business value creation through business processes management and operational inteligencia de negocio integration* [Archivo PDF]. In System Sciences (HICSS), 2010 43rd Hawaii International Conference on (pp. 1-10). IEEE. Disponible en https://www.computer.org/csdl/proceedings/hicss/2010/3869/00/09-02-02.pdf

Real Academia Española. (2014). Diccionario de la lengua española (23.a ed.). Consultado en https://dle.rae.es/?id=AoexZZK

Silahtaroğlu, G., & Alayoglu, N. (2016). *Using or Not Using Business Intelligence and Big Data for Strategic Management: An Empirical Study Based on Interviews with Executives in Various Sectors* [PDF]. Procedia-Social and Behavioral Sciences, 235, 208-215. Disponible en https://www.sciencedirect.com/science/article/pii/S1877042816315506

Sinnexus. (2018). *Business Intelligence Informatica*. Estratégica. Recuperado el 15 de 06 de 2018, de https://www.sinnexus.com/business_intelligence/

Twum Amoako, Benjamin. (2013). *The Importance Of Business Intelligence As A Decision-Making Tool: –Case Study Electricity Company Of Ghana (E.C.G)* (Tesis de maestría). Universidad de Boras, Suecia.

Bibliografía

Elmasri, R., & Navathe, S. (2016). *Fundamentals of database systems*. London: Pearson.

Flores, J. (08 de 12 de 2015). *Tomando decisiones acertadas con BI*. Recuperado el 20 de Julio de 2018, de Grupo CONSISA. Recuperado el 27 de 06 de 2018, de http://www.consisa.com/index.php/es/tomando-decisiones-acertadas-con-bi/

Guerrero, M. (21 de 11 de 2016). *Business Process Management y su relación con SOA y BI* MegaPractical. Recuperado el 07 de 23, de https://www.megapractical.com/blog-de-arquitectura-soa-y-desarrollo-de-software/business-process-management-y-su-relacion-con-soa-y-bi

INCAE. (12 de 07 de 2017). *Principales retos al implementar Business Intelligence* INCAE. Recuperado el 07 de 2018, de https://www.incae.edu/es/blog/2017/07/12/principales-retos-al-implementar-business-intelligence.html

ITnow. (20 de marzo de 2015). *¿Cómo optimizar el BPM a partir de BI?* ITnow. Recuperado el 07 de 2018, de https://revistaitnow.com/como-optimizar-el-bpm-a-partir-de-bi/

Juan, A. (08 de Febrero de 2017). *5 claves para la toma de decisión empresarial*. Recuperado el 26 de Junio de 2018, de Cultura CRM: https://culturacrm.com/business-intelligence/toma-decisiones-business-intelligence/

Calvo, M. (22 de 03 de 2018). Algunos ejemplos prácticos de uso de inteligencia de negocio. Captio. [Entrada de blog]. Recuperado el 07 de 07 de 2018, de https://www.captio.net/blog/algunos-ejemplos-practicos-de-uso-de-business-intelligence

Martínez, S. (25 de 05 de 2018). *¿Reporting o Business Intelligence?* Mundo.erp. Recuperado el 17 de 07 de 2018, de https://www.mundoerp.com/blog/reporting-o-business-intelligence/

Meneses, H. (2013 de 06). *El Enfoque de Procesos como Requisito para el Éxito de un Proyecto de BI.* CiclusGroup. Recuperado el 07 de 2016, de https://ciclusgroup.wordpress.com/2013/06/13/el-enfoque-de-procesos-como-requisito-para-el-exito-de-un-proyecto-de-bi/

Proceso BI [Imagen En línea]. (2011). Obtenido en Julio 23, 2018 de https://www.gestiopolis.com/wp-content/uploads/2011/09/inteligencia-de-negocios-en-que-ayuda-a-tu-negocio-y-cual-es-su-proceso.gif

Rubens, P. (03 de 2014). *6 Big Business Intelligence Mistakes* Enterprise Apps Today. Recuperado el 19 de 07 de 2018, de Big Business Intelligence: http://www.enterpriseappstoday.com/business-intelligence/6-big-business-intelligence-mistakes.html

Soto, M. (30 de 11 de 2017). *7 Beneficios Estratégicos de utilizar Inteligencia de Negocios* wivo. Recuperado el 10 de 07 de 2018, de https://blog.wivoanalytics.com/7-beneficios-estrategicos-de-utilizar-inteligencia-de-negocios

Stotecheck. (07 de 2018). *3 aspectos a evaluar para elegir una herramienta de Business Inteligence.* Stotecheck. Recuperado el 24 de 07 de 2018, de http://blog.storecheck.com.mx/antes-de-elegir-business-inteligence

Versiones y Costos del Motor SQL Server [Imagen en Línea]. (2017) Obtenido el 23 de Julio 2018 de http://blogs.askcts.com/2017/01/17/introduction-microsoft-sql-server-2016-service-pack-1/

Anexos

Anexo A. Encuesta diagnóstica del departamento de contabilidad.

Intención. La presente encuesta busca descubrir un estado inicial del departamento con respecto al uso de datos. La misma se ha realizado en el departamento objetivo del proyecto y cuenta con la participación de seis integrantes del mismo.

Estructura de las preguntas. Las preguntas se orientaron a dos ámbitos principales. La primera es el uso de los datos actuales, su experiencia con los mismos y los pasos que deberían seguir cada vez. En una segunda parte se analiza lo que serían nuevas tecnologías, el conocimiento de los integrantes y el uso que posiblemente le darían.

Descripción de las preguntas. En esta sección se analizan cada una de las preguntas realizadas a los participantes. ¿Cuál es su intención?

Pregunta #1 ¿Cuál es la frecuencia en la que debes realizar reportes de actividades de créditos? La idea de esta pregunta es poder comprender el tipo de reportes que normalmente realizan. Las opciones son semanal, mensual, nunca. Si es semanal, los datos utilizados probablemente son más "nuevos", si fuese mensual significa el informe de labores normal que hay en distintas organizaciones cada mes y si no requiere hacer reportes implica entonces para otras preguntas subsecuentes.

Pregunta #2 ¿Qué tan confortable es el acceder a los datos? La idea de esta pregunta es poder mostrar el sentimiento actual de cada uno de los implicados con respecto al acceso a los datos. Las opciones son mucho, normal, poco y muy poco.

Pregunta #3 ¿Tienes dependencias para la extracción de datos? Esta pregunta busca conocer acerca del proceso de acceso a datos. Normalmente en las organizaciones no se tiene una base de datos directamente accesible a todos los usuarios y esto implica que se requiere preguntar a terceros acerca de los mismos. La pregunta entonces es conocer o confirmar si esto es efectivamente el caso, solo hay dos opciones: si o no.

Pregunta #4 ¿Qué tan actual requieren los datos? Esto es una pregunta interesante para nuestra solución, debido a que de acuerdo lo que respondan implicaría así la actualización de datos en el almacén de datos. Las opciones son diarias, mensual, anual.

Pregunta #5 ¿Qué tan necesario visualizan tener una base de datos analítica? Esta pregunta es para descubrir acerca de cómo visualizan las bases de datos analíticas. No se entra en detalles pero se consulta acerca de si es necesario, o poco necesario la inclusión de este tipo de tecnologías de análisis.

Pregunta #6 ¿Conocen acerca de PowerBI? Esta es una pregunta cerrada para ver qué tanto conocen acerca de esta tecnología de bases de análisis de datos. La respuesta es cerrada: Si o No.

Pregunta #7 ¿Conocen acerca de Tableau? Esta es una pregunta cerrada para ver qué tanto conocen acerca de esta tecnología de bases de análisis de datos. La respuesta es cerrada: Si o No.

Pregunta #8 ¿Conocen acerca de SSAS? Esta es una pregunta cerrada para ver qué tanto conocen acerca de esta tecnología de bases de análisis de SQL Server. La respuesta es cerrada: Si o No.

Pregunta #9 ¿Sienten que la base de datos es rápida para consultas? Esta pregunta es para confirmar una de los sentimientos que se lograron captar al inicio del proyecto. Es una confirmación. Según el conocimiento actual se espera que esta sea no para la mayoría de los casos.

Pregunta #10 ¿Requieren generar gráficos para representar los datos?

Esta pregunta busca descubrir qué tan necesario son los gráficos para los diversos usuarios potenciales de la tecnología de inteligencia de negocios.

Resultados de la encuesta. Con respecto a la pregunta #1 se logró comprobar que la mitad de los que hacen reportes lo realizan de manera mensual. Esto representa tres de las seis personas que hicieron la encuesta. Las que respondieron que de manera semanal indicaron que esto lo hacían de manera casual, no en todos los momentos del año y generalmente iniciados por una solicitud de datos.

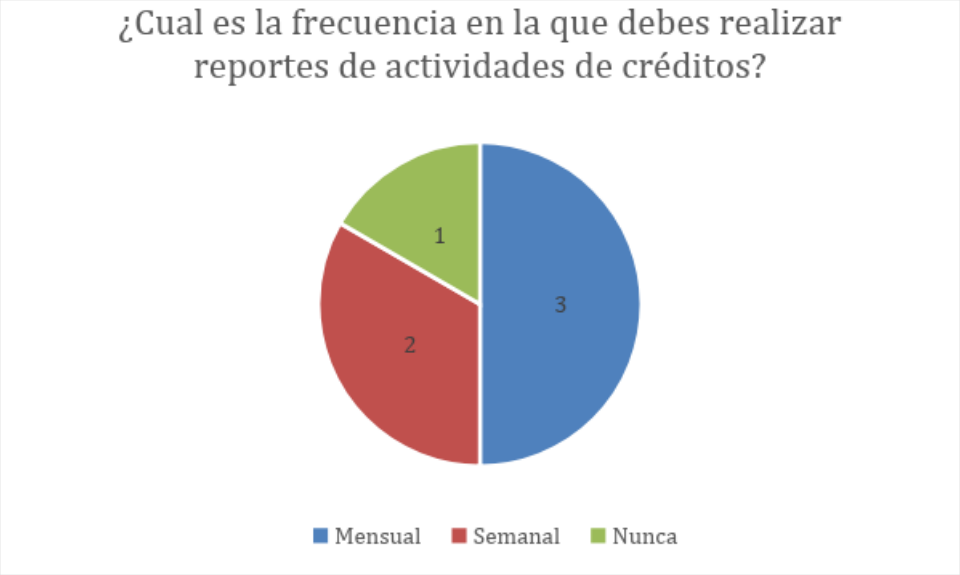


Figura 31 Gráfico de la frecuencia con lo que se debe analizar reportes en el diagrama.

Con respecto a la pregunta #2 se logró comprobar que todos están disconformes con la forma en que se acceden los datos. Esto si lo asociamos con la pregunta siguiente, la #3 se confirma que gran parte de esto es la dependencia del departamento de tecnologías de información, todos indicaron que si se tienen dependencias. Esto produce atrasos, malos entendidos. Esta es una de las situaciones que más oportunidades de mejora presenta de acuerdo a la investigación.

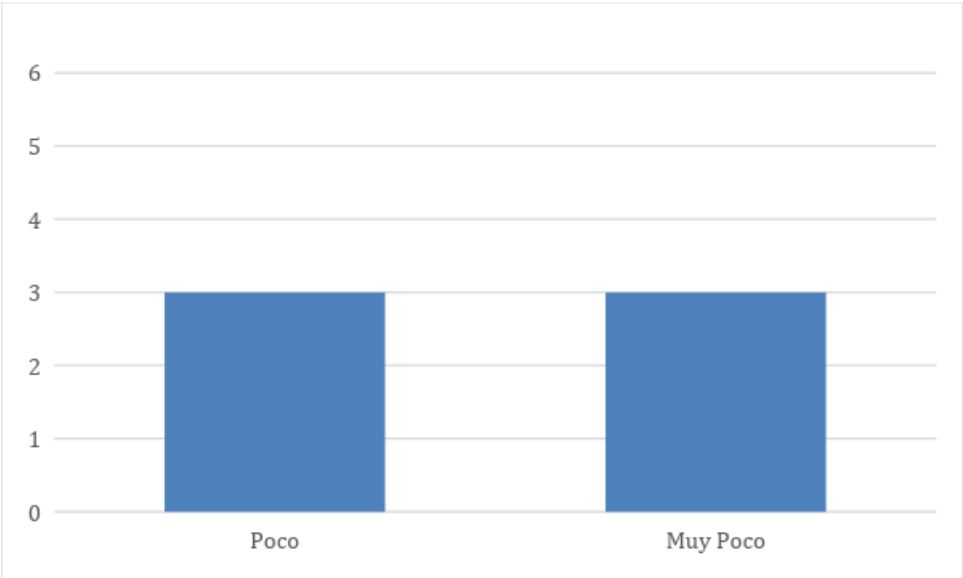


Figura 32 Gráfico con respecto a lo confortable del acceso de datos.

Con respecto a la pregunta de ¿qué tan actual requieren los datos?, el total de los entrevistados coinciden que ha de ser las operaciones del mes. Igualmente con la pregunta 5, todos están de acuerdo de que es necesario o muy necesario el poder contar con una base de datos analítica.

Ya en el área más técnicas, orientado al conocimiento de PowerBI, Tableau y SSAS, el 83% de los entrevistados indican que no conocen acerca de Tableau, mientras todos si conocen o tienen la idea de lo que es PowerBI. Algo muy parecido sucede con los servicios de Análisis de SQL server.

Ya en la última parte de preguntas, en la pregunta 9 consideran que a la base de datos no es rápida para realizar consultas. Por su parte el 83% coinciden que si requieren gráficos para representar los datos.

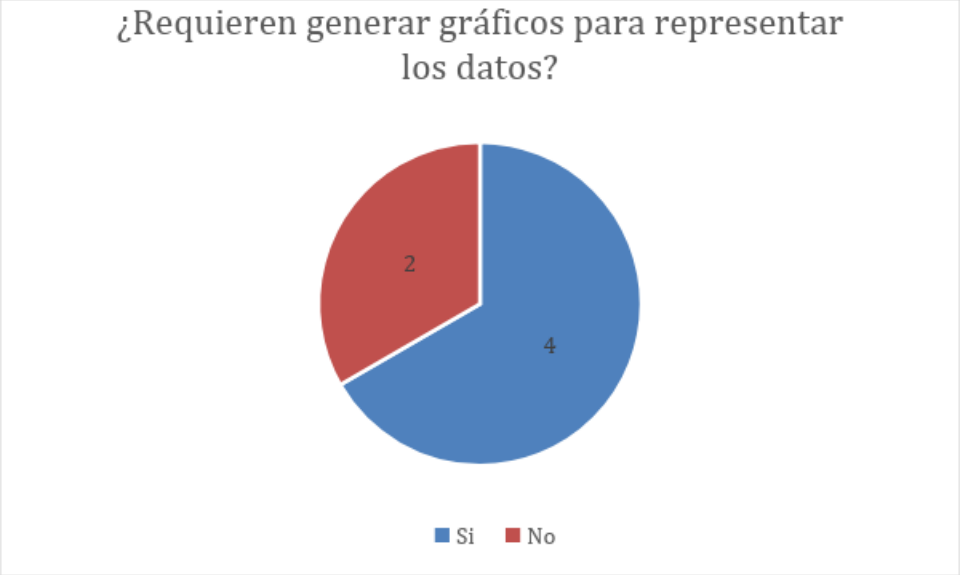


Figura 33 Gráfico con respecto a la necesidad de utilizar elementos gráficos para presentar datos

Anexo B. Diagrama de base de dato multidimensional.

En este anexo se presenta la estructura seleccionada para la base de datos multidimensional. Se ha seleccionado el modelo estrella del mismo.

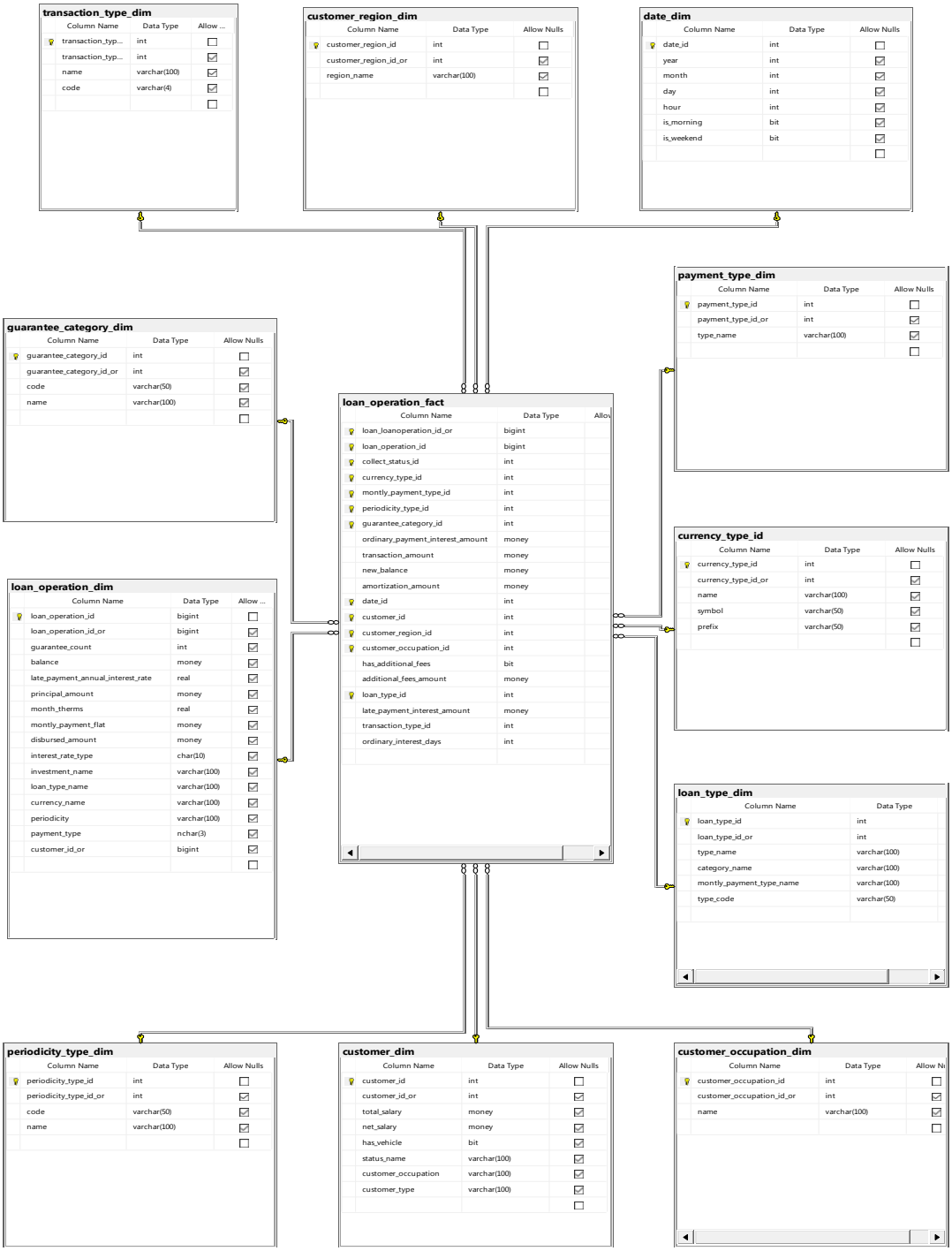


Figura 34 Anexo 1. Diagrama de bases de datos multidimensional.

Anexo C. Diagrama de base datos analítica.

En este anexo se presenta la estructura seleccionada para la base de datos analítica utilizada en el proyecto.

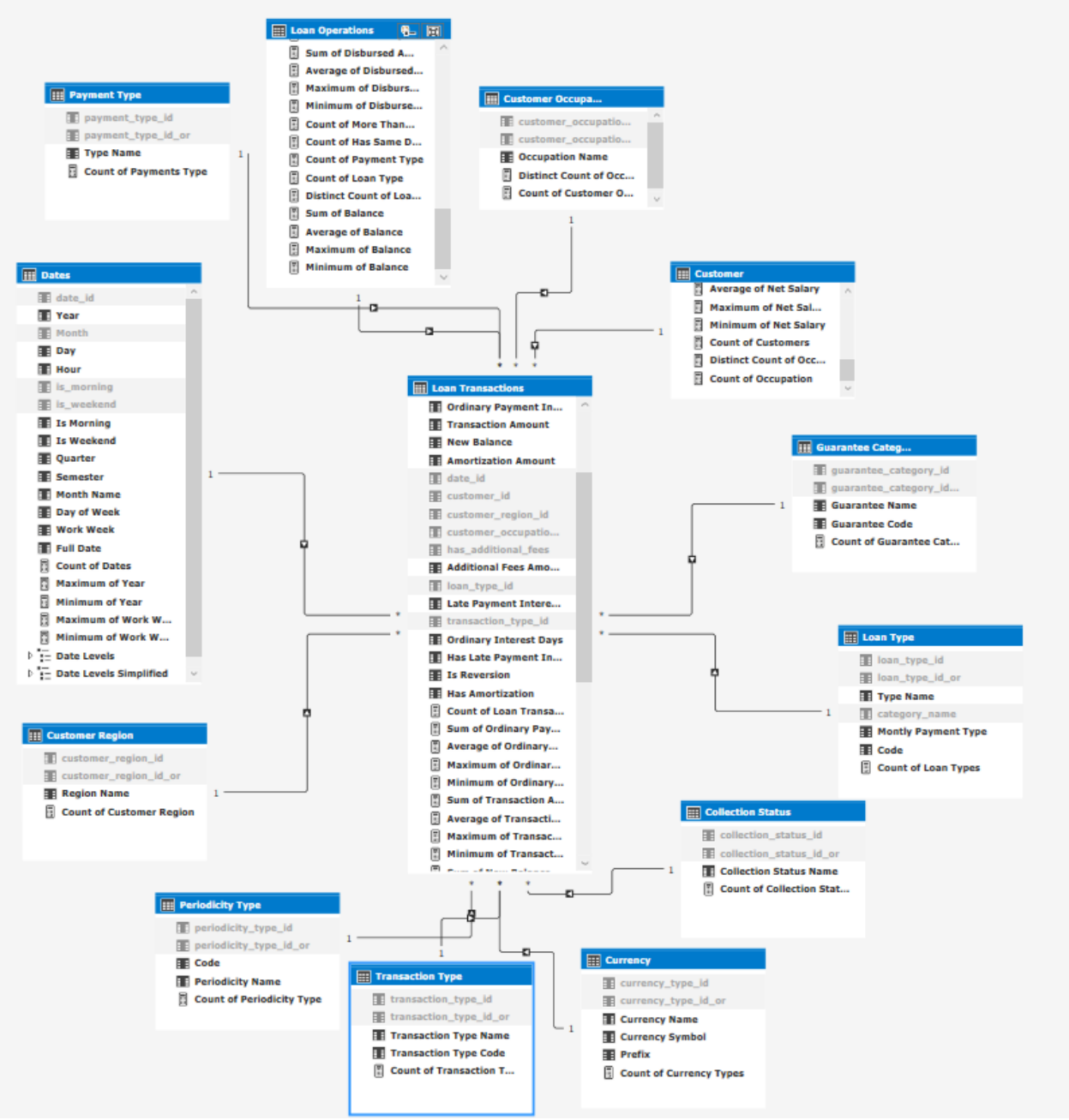


Figura 35 Anexo B. Diagrama de base de dato analítica.