



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

FACULTAD DE INGENIERÍA

CARRERA DE INGENIERÍA EN CIENCIAS DE LA COMPUTACIÓN

TEMA:

Desarrollo de prototipo de sistema de Business Intelligence para la toma de decisiones estratégicas en una camaronera localizada en la provincia del Guayas.

AUTORA:

Sánchez Naranjo Eileen Belen

**Trabajo de titulación previo a la obtención del título de
INGENIERO EN CIENCIAS DE LA COMPUTACIÓN**

TUTOR:

Yong Yong Byron Severo, Ing.

Guayaquil, Ecuador

08 de febrero del 2024



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA

CARRERA DE INGENIERÍA EN CIENCIAS DE LA COMPUTACIÓN

CERTIFICACIÓN

Certificamos que el presente trabajo de titulación fue realizado en su totalidad por **Sanchez Naranjo Eileen Belen**, como requerimiento para la obtención del título de **Ingeniero en Ciencias de la Computación**.

TUTOR

f. _____

Yong Yong Byron Severo, Ing.

Guayaquil, a los 8 días del mes de febrero del año 2024



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA

CARRERA DE INGENIERÍA EN CIENCIAS DE LA COMPUTACIÓN

DECLARACIÓN DE RESPONSABILIDAD

Yo, **Sánchez Naranjo Eileen Belen**

DECLARO QUE:

El Trabajo de Titulación: **Desarrollo de prototipo de sistema de Business Intelligence para la toma de decisiones estratégicas en una camaronera localizada en la provincia del Guayas**, previo a la obtención del título de **Ingeniero en Sistemas Computacionales**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guayaquil, a los 8 días del mes de febrero del año 2024

EL AUTOR

f. _____

Sánchez Naranjo Eileen Belen



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA

CARRERA DE INGENIERÍA EN CIENCIAS DE LA COMPUTACIÓN

AUTORIZACIÓN

Yo, **Sánchez Naranjo, Eileen Belen**

Autorizo a la Universidad Católica de Santiago de Guayaquil a la **publicación** en la biblioteca de la institución del Trabajo de Titulación: **Desarrollo de prototipo de sistema de Business Intelligence para la toma de decisiones estratégicas en una camaronera localizada en la provincia del Guayas**, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

Guayaquil, a los 8 días del mes de febrero del año 2024

EL AUTOR:

f. _____

Sánchez Naranjo, Eileen Belen



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERIA

CARRERA DE INGENIERIA EN CIENCIAS DE LA COMPUTACIÓN

REPORTE COMPILATIO

 CERTIFICADO DE ANÁLISIS
mgster

3 ES_Trabajo de Titulación ICC 24-01-26

1%
Textos
sospechosos

Nombre del documento: 3 ES_Trabajo de Titulación ICC 24-01-26.docx	Depositante: Byron Severo Yong Yong
ID del documento: 8d01cc7b23287dde11ccb8d66f6fb64b48967498	Fecha de depósito: 26/1/2024
Tamaño del documento original: 7.79 MB	Tipo de carga: Interface
	fecha de fin de análisis: 26/1/2024

f. _____

Ing. Byron Yong Yong
Tutor de Trabajo de Integración Curricular – UCSG
Carrera en Ciencias de la Computación

Agradecimiento

Quisiera expresar mi agradecimiento en primer lugar con Dios por cada una de las oportunidades brindadas para culminar mi carrera, con mi familia por su apoyo constante en cada nivel de mi etapa universitaria, por los valores inculcados ya que con ellos logre ser la persona que soy ahora.

A mis padres que a pesar de los obstáculos que se me presentaron en esta etapa jamás dudaron de mí, a quienes les agradezco por su insistencia en la culminación de mis estudios, ellos que siempre estaban presente con su apoyo y su soporte en cualquier situación que se me presentaba, Gracias.

A mis abuelos que sin duda alguna me han apoyado incondicionalmente y confiaron en la culminación de mis estudios con éxito, a quienes que a pesar de estar en otro país siempre han estado presente, gracias a ustedes que con sus valores y llamados de atención ayudaron a mi formación como persona.

A mis verdaderos amigos, quienes que desde su presencia en mi vida han vivido cada angustia y celebrado cada logro de esta etapa universitaria, agradezco que hayan sido un soporte más en mi vida en momentos que no he podido sola.

Y, por último, agradezco a los docentes y administrativos de la universidad quienes me han permitido avanzar en mi formación como profesional, apoyándome en cada nivel, y dándome la oportunidad de demostrarles la excelente profesional que llegare a ser; a mi tutor Byron Yong, por su paciencia y su excelente guía en el desarrollo exitoso de este trabajo.

Dedicatoria

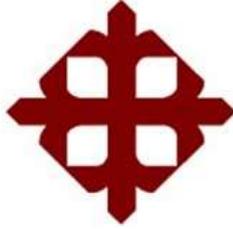
Dedico el presente trabajo de titulación a mis padres Lizandro Sánchez y Carmen Naranjo, a mi hermano Emiliano Carpio quienes son el pilar fundamental en mi vida, fuente de inspiración y superación a pesar de las adversidades.

A mis abuelos: Epifanía de los Ángeles Oquendo, Aquilino Sanchez, Carmen Sanchez y Rafael Naranjo quienes han estado presentes con su apoyo incondicional en cada etapa de mi vida y más aún en la universitaria.

A mis mejores amigas: Lissette Ávila y Verónica León, quienes son el empujoncito de confianza en cada momento de duda que he tenido.

A mi pareja y futuro compañero de vida, Jean Paul Suarez quien me recordaba día a día mi meta profesional y me impedía rendirme sea cual sea el obstáculo.

A mi primo hermano en el cielo Rommel Ramos quien en vida fue y será un ejemplo a seguir de liderazgo y superación.



UNIVERSIDAD CATÓLICA DE SANTIAGO DE GUAYAQUIL

FACULTAD DE INGENIERÍA

CARRERA DE INGENIERÍA EN CIENCIAS DE LA COMPUTACIÓN

TRIBUNAL DE SUSTENTACIÓN

f. _____

ING. ANA CAMACHO CORONEL, MGS

DIRECTORA DE CARRERA

f. _____

ING. GALO ENRIQUE CORNEJO GOMEZ

DOCENTE DE LA CARRERA

f. _____

ING. EDISON JOSÉ TOALA QUIMI, MGS

OPONENTE

ÍNDICE

Resumen	XIV
Abstract	XV
Introducción	2
Capítulo I: El Problema	3
Análisis Del Problema	3
Ubicación del Problema en un Contexto	3
Causas y Consecuencias del Problema.....	4
Delimitación del Problema	5
Formulación del Problema	6
Evaluación del Problema	6
Objetivos.....	7
Objetivo General.....	7
Objetivos Específicos	7
Alcances Del Problema	7
Justificación e Importancia	8
Hipótesis O Pregunta De Investigación	8
Variables De La Investigación	8
Capítulo II: Marco Teórico	9
Fundamentación Teórica.....	9
Base de Datos	9
Gestores de Bases de Datos (DBMS):.....	10
Tipos de Base de Datos.....	10
Framework.....	11
Lenguaje de Programación	12
Python	12
ETL (Extract, Transform, Load) con Python	13
Data Warehouse.....	13
Business Intelligence	14
Microsoft Power BI:.....	15
KPI's (Indicadores Clave de Desempeño):	15
Indicadores de Desempeño Financieros	16

Capítulo III: Metodología De La Investigación	18
Método De Investigación Cuantitativa	18
Población y Muestra.....	19
Población.....	19
Muestra	20
Recolección de Información.....	20
Metodología de Construcción de Software	25
Capítulo IV: Propuesta Tecnológica	28
Situación Actual	28
Origen de Datos	28
Herramientas De Desarrollo	28
Selección de Herramientas.....	28
Fase 1: Análisis de Requerimientos.....	29
Fase 2: Análisis de Data Sources (OLPT)	32
Fase 3: Modelo Lógico del Data Warehouse (DW)	35
Fase 4: Integración de Datos	35
Puesta en Producción	38
Conclusiones	44
Recomendaciones	45
Referencias Bibliográficas.....	46
Anexos.....	47
Anexo 1 Informe Reunión de Requerimientos con Alta Gerencia	47
Anexo 2 Reportes Contables de Diciembre.....	47
Anexo 3 Script de Conexión a la Base de Datos	48
Anexo 4 Script de Carga Histórica	48
Anexo 5 Script de Actualización	48
Anexo 6 Proyecto BI	48
Glosario	49

ÍNDICE DE ILUSTRACIONES

Ilustración 1 Flujo de Toma de Decisiones Gerenciales.....	4
Ilustración 2 Síntesis de la problemática.....	5
Ilustración 3 Componentes de Base de datos.....	9
Ilustración 4 Procesos de Business Intelligence	14
Ilustración 5 Tipos de Indicadores	16
Ilustración 6 Formato de Encuesta	20
Ilustración 7 Experiencia con Dashboard.....	22
Ilustración 8 Indicadores Financieros.....	22
Ilustración 9 Nuevas propuestas de Indicadores.....	23
Ilustración 10 Preferencias de Visualización.....	23
Ilustración 11 Alertas automáticas	23
Ilustración 12 Accesibilidad.....	24
Ilustración 13 Integración de Sistemas	24
Ilustración 14 Capacitación.....	24
Ilustración 15 Importancia del Proyecto	25
Ilustración 16 Metodología HEFESTO.....	26
Ilustración 17 Tareas Metodología Hefesto.....	27
Ilustración 18 Query Balance General	30
Ilustración 19 Query Estado Pérdidas y Ganancias	31
Ilustración 20 Modelo Conceptual.....	32
Ilustración 21 Dataframe de Balance General.....	33
Ilustración 22 Dataframe de Estado de Pérdidas y Ganancias	33
Ilustración 23 Mapeo de Datos	34
Ilustración 24 Modelo Conceptual Ampliado	35
Ilustración 25 Modelo Lógico DW	35
Ilustración 26 Extracción de Balances Generales	36
Ilustración 27 Extracción de los Estados de Pérdidas y Ganancias	36
Ilustración 28 Script de Actualización	37
Ilustración 29 Carga de Información	38
Ilustración 30 Tablas Creadas del Data Warehouse	38
Ilustración 31 Pantalla de Menú de Dashboard.....	39
Ilustración 32 Pantalla de Indicadores de Liquidez	39

Ilustración 33 Pantalla de Indicadores de Endeudamiento.....	40
Ilustración 34 Pantalla de Indicadores de Rentabilidad.....	40
Ilustración 35 Series de Tiempo de Indicadores de Liquidez	41
Ilustración 36 Series de Tiempo de Indicadores de Endeudamiento	41
Ilustración 37 Series de Tiempo de Indicadores de Rentabilidad	41
Ilustración 38 Gráficos de barra comparativos por meses de los Indicadores de Liquidez	42
Ilustración 39 Gráficos de barra comparativos por meses de los Indicadores de Endeudamiento.....	43
Ilustración 40 Gráficos de barra comparativos por meses de los Indicadores de Rentabilidad.....	43

ÍNDICE DE TABLAS

Tabla 1 Detalle tiempo de proceso de obtención de información	3
Tabla 2 Delimitación del problema	6
Tabla 3 Indicadores Financieros	16
Tabla 4 Población Reportes Financieros	19
Tabla 5 Población Alta Gerencia.....	19
Tabla 6 Comparación de Metodologías	26
Tabla 7 Requerimientos obtenidos	29
Tabla 8 Campos del Dataframe	34

Resumen

El presente trabajo de titulación tiene como objetivo la implementación de un prototipo de Inteligencia de Negocios para ayudar en la toma de decisiones de la alta gerencia de una camaronera localizada en la ciudad de Guayaquil, esto simplificaría la obtención de información efectiva y fundamentada en datos históricos, necesarios para el proceso. Para ellos se hará uso de tecnología que permita realizar el análisis, diseño, consolidación y extracción de los datos transaccionales de la organización. Los reportes desactualizados y los reprocesos solicitados hacen que la toma de decisiones no se realice de manera efectiva, la aplicación de BI permite obtener indicadores financieros y de gestión de manera fácil y eficiente en cualquier momento sin necesidad de un proceso manual, para la construcción de la solución informática se hará uso de: la metodología de Hefesto, la herramienta Power BI y la base de datos MySQL permitiendo la creación de un propio Data Warehouse que permitirá la consulta de información efectiva en cualquier momento.

Palabras Clave: *Inteligencia de Negocio, Power BI, MySql, Indicadores Financieros, Hefesto.*

Abstract

The objective of this degree work is the implementation of a Business Intelligence (BI) prototype to help in the decision-making process of the top management of a shrimp farm located in the city of Guayaquil, this would simplify the obtaining of effective information based on historical data, necessary for the process. For this purpose, technology will be used to analyze, design, consolidate and extract the transactional data of the organization. The BI application allows to obtain financial and management indicators easily and efficiently at any time without the need of a manual process. For the construction of the IT solution, we will make use of the Hefesto methodology, the Power BI tool and the MySQL database allowing the creation of a Data Warehouse that will allow the consultation of effective information at any time.

Key words: *Business Intelligence, Power BI, MySQL, Financial Indicators, Hefesto.*

Introducción

En 1976 en la Provincia del Guayas la camaronera inicia sus actividades como productora, empackadora y exportadora de camarón. Con 47 años de experiencia, se dedica a producir productos del mar y exportar más de 900 contenedores al año; sus principales productos de exportación son: camarón (*Penaeus Vannamei*) y Tilapia Marina (*Oreochromis Niloticus*) fresca y congelada, siendo este su principal objetivo comercial.

Sus productos son de calidad garantizada, por lo que cuenta con laboratorios propios, fincas y plantas industriales modernas, mismos que cumplen con las normas internacionales establecidas por La Administración de Alimentos y Medicamentos de USA (FDA¹), todo bajo control e instructivos del plan Análisis de puntos críticos de control (HACCP²) y con certificación de la Asociación de Minoristas Británicos (BRC³).

La camaronera se encuentra en constante innovación tecnológica, esto permite la optimización y digitalización de procesos generando grandes cantidades de datos de sus actividades diarias. Debido al avance de la industria se hace demandante fuentes de información precisas y exactas para la toma de decisiones estratégicas.

La toma de decisiones en los modelos de negocios por parte de la alta gerencia (Gobierno Corporativo), es crucial y debe ser respaldada en información confiable, efectiva, integra y disponible, a efectos de ser precisos al decidir alguna situación del negocio; aprovechando las actividades de mejoramiento continuo e innovación tecnológica de la empresa, nace la necesidad de mejorar el proceso de toma de decisiones de la alta gerencia mediante el uso de herramientas de TI enfocadas en el Business Intelligence.

Con este proyecto se pretende brindar a la empresa una herramienta de Business Intelligence, para brindar información de indicadores financieros en un Dashboard, usando el lenguaje de programación Python para el proceso ETL de los datos, los cuales serán almacenados en una estructura de Data Warehouse para luego visualizarlos en la aplicación Power BI.

¹ U.S. Food and Drug Administration.

² Hazard Analysis and Critical Control Points.

³ British Retail Consortium.

Capítulo I: El Problema

A continuación, se describe la situación problema de la toma de decisiones estratégicas por parte de la alta gerencia que demanda tiempo operativo del área contable requiere de una solución óptima que permita evitar los cálculos manuales y la obtención de información exacta en menor tiempo.

Análisis Del Problema

Ubicación del Problema en un Contexto

La camaronera se mantiene en constante innovación tecnológica, esto demanda a la alta gerencia tomar decisiones más eficientes para mejores resultados; al enfocarse en su constante crecimiento tecnológico, optimiza procesos y mejora la estructura de costos; no obstante, los procesos gerenciales utilizan reportes de diversas fuentes y cálculos manuales, haciendo que la información no sea 100% confiable y no disponible inmediatamente.

La ejecución de todas sus actividades genera grandes cantidades de datos los cuales son únicamente almacenados y usados para generar reportería contable la cual toma un tiempo aproximado de 4 horas debido al periodo de información que se necesita, luego de la obtención de la información de reportería, el personal contable de la camaronera procede a elaborar el informe de indicadores financieros requerido por la alta gerencia; esto se realiza manualmente, ocupando tiempo adicional de sus labores diarias correspondiente a 10.75 horas, tal como se observa en la siguiente tabla:

Tabla 1 Detalle tiempo de proceso de obtención de información

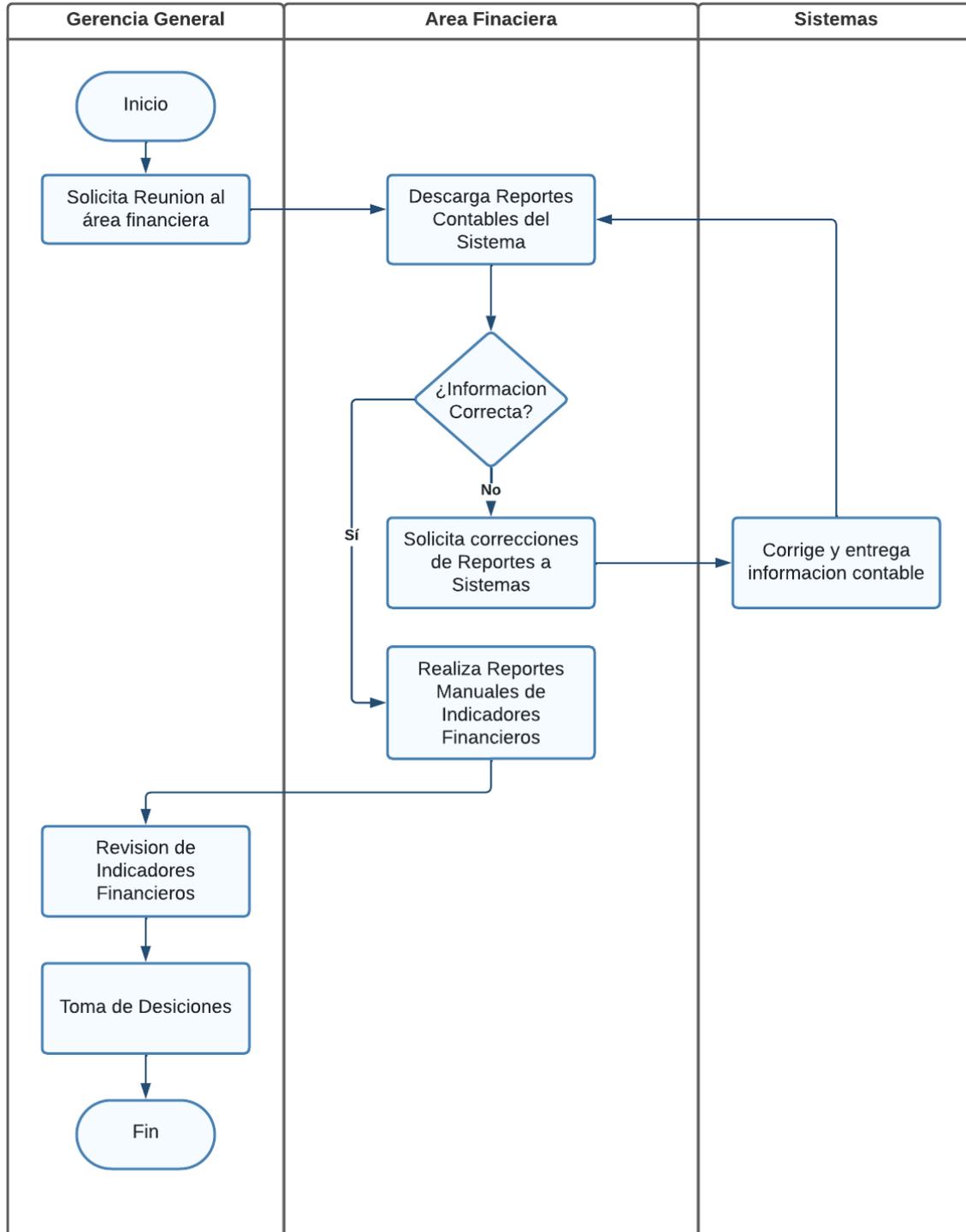
Laborales	Horas Revisión	Horas Reproceso	Reprocesos Solicitados	Horas Totales
Personal Contable 1	1.00	0.50	2.00	3.00
Personal Contable 2	0.75	0.50	3.00	3.75
Personal Contable 3	1.50	0.50	2.00	4.00
Total				10.75

Nota: Las horas totales son obtenidas de la suma de las horas de revisión y las horas de reproceso multiplicadas por la cantidad de reprocesos solicitados.

Si se considera los tiempos de: elaboración de los reportes contables y de generación de los indicadores financieros el tiempo total utilizado para la preparar la información a la alta gerencia para el proceso de toma de decisiones es de 14.75 horas. Analizando este proceso se observa que está definido correctamente, sin embargo, al con contar con un sistema que extraiga, depure, cargue, procese y muestre la información se tiene una subutilización del

talento humano; a continuación, podemos observar el flujo del proceso:

Ilustración 1 Flujo de Toma de Decisiones Gerenciales



Nota: Este proceso involucra el área gerencial, financiera y de sistemas los dos últimos son responsables de brindar la información requerida para la gerencia para el proceso.

Causas y Consecuencias del Problema

La toma de decisiones estratégicas demanda al personal contable alrededor de 4

horas estar listos con la información de indicadores financieros para la alta gerencia, esto se debe a que no cuentan con una herramienta de información de acceso inmediato y con datos actualizados, esto hace que personal contable intervenga para la gestión estadística manual.

La camaronera no cuenta con presupuestos de gastos de las áreas establecidos, se maneja por solicitudes emergentes de acuerdo a la demanda, la falta de la información de indicadores financieros no permite establecer presupuestos, es decir de acuerdo a la situación de la empresa destinar presupuestos establecidos para los diferentes departamentos, ni tomar decisiones estratégicas, la falta de reportería actualizada y no tener proyecciones son la causa de que la toma de decisiones estratégicas no sea oportuna y con información efectiva; usando el modelo de Ishikawa se muestra a continuación de manera sintetizada las causas del problema.

Ilustración 2 Síntesis de la problemática



Nota: Las causas de problema en este proyecto de titulación es la falta de herramientas de inteligencia de negocio que ayuden a la toma de decisiones estratégicas.

Este trabajo se enfoca en la automatización del proceso manual de generación de información financiera que demanda el proceso de toma de decisiones, es decir los cálculos de indicadores realizados manualmente, se realizaran por medio del proceso ETL de esta manera no necesitara sobrecargar al personal contable permitiéndoles seguir con sus funciones normales y tendría la información lista para consultarla en cualquier momento.

Delimitación del Problema

El problema se refleja en el proceso de toma no oportuna de decisiones estratégicas, la generación de reportes, la intervención de personal del área contable para general reportes gerenciales, el tiempo perdido en este proceso, para obtener información no tan exacta para

las estrategias gerenciales. A continuación, podemos delimitar el problema enfocándonos en 3 aspectos.

Tabla 2 Delimitación del problema

Campo		Tecnología
Área	Inteligencia de Negocios	
Aspectos	Haciendo uso de la Inteligencia de Negocios se proporcionará herramientas exactas para la ayuda a la toma de decisiones, esto será elaborado en Power BI	
Tema	Desarrollo de Prototipo de sistema de Business Intelligence para la toma de decisiones estratégicas en una camaronera localizada en la Provincia del Guayas	

Nota: Se detallan el área, los aspectos y tema que ayudan a la delimitación del Problema.

Formulación del Problema

¿Cómo la implementación del Business Intelligence, mediante la herramienta Power BI, ayudará a la toma de decisiones estratégicas de manera oportuna, apropiada y con información efectiva?

Evaluación del Problema

Delimitado: La problemática es identificada en una camaronera de la Provincia del Guayas, a pesar de que vive en constante innovación tecnológica y almacenando datos importantes de todas sus actividades, no tiene implementado un dashboard con un Cuadro de Mando Integral (CMI). La propuesta que se plantea en el trabajo de titulación se enfoca en la optimización del proceso de toma de decisiones estratégicas a nivel gerencial.

Claro: La camaronera a pesar de su constante innovación tecnológica no cuenta con la implementación de un dashboard con un cuadro de mando integral, debido a esto resulta ser una demanda de alta prioridad.

Evidente: Gracias al análisis de los tiempos de ejecución del proceso de toma de decisiones estratégicas podemos evidenciar que es necesario automatizar este proceso.

Relevante: La camaronera no cuenta con una herramienta de Business Intelligence a pesar de ser una gran industria, la aplicación de este proyecto puede servir como el primer paso en otras camaroneras de la provincia.

Original: Al no contar con una herramienta de inteligencia de negocios la camaronera implementará un proyecto desde el inicio y podrá a futuro ser optimizado y adaptado a las condiciones cambiantes del modelo de negocio.

Concreto: El problema de la camaronera se centra en el proceso de obtención de información de calidad para la toma de decisiones estratégicas.

Factible: El problema es factible de ser resuelto gracias a su análisis y obtención de conclusiones a corto plazo durante el estudio del proyecto, sin embargo, también sirve para análisis a largo plazo.

Objetivos

Objetivo General

Mejorar el proceso de toma de decisiones de la alta gerencia mediante el uso de herramientas de TI. enfocadas en el Business Intelligence que permitan la extracción, depuración y carga de datos transaccionales a cubos de información con su respectiva generación de reportes y consultas financieras.

Objetivos Específicos

- Diseñar el Almacén de datos (Data Warehouse) para el fácil acceso a la información corporativa de carácter financiera.
- Construir los procesos de extracción, transformación y carga de datos (ETL⁴) desde los sistemas transaccionales hacia la base de datos DW.
- Implementar el prototipo en el ambiente de producción de la camaronera.
- Elaboración de dashboard informativo y reportes gerenciales.

Alcances Del Problema

La implementación del prototipo de Business Intelligence a través de un dashboard, permitirá a la alta gerencia tener una mejor perspectiva sobre los indicadores financieros para la toma de decisiones estratégicas.

Se analizará e identificará los datos necesarios para la creación del Dashboard, enfocándose únicamente en la parte financiera de la empresa.

El desarrollo de este prototipo incluirá la creación de cubos OLAP⁵ para contener

⁴ Extract, transform, and load.

⁵ Online analytical processing.

información consolidada con respecto a lo que el gobierno corporativo necesita acceder.

La información se generará en dashboard y/o reportes dinámicos en el cual se aplicarán los indicadores CMI.

En aplicaciones de dashboard se debe considerar la información recolectada y consolidada, para permitir la toma de decisiones oportunas en base a datos procesados y ser mostrados de manera clara. Esto será posible con la implementación del proceso ETL usando el lenguaje de programación Python sobre los datos de la empresa, procediendo a la creación de una estructura de Data Warehouse en MySQL en donde se almacenará la información procesada para la obtención de reportes y en Dashboard informativo creados en la herramienta Power BI.

Justificación e Importancia

La camaronera almacena cada actividad que realiza, desde los procesos de producción, movimientos de inventarios, hasta las exportaciones que es la principal actividad de la empresa, esto es posible gracias a la automatización de procesos en los distintos departamentos.

Desde que se prioriza la innovación tecnológica no se ha tomado en consideración la implementación de un Dashboard, sin embargo, la alta gerencia demanda la necesidad de la misma. El presente trabajo de titulación buscará mediante la implementación del prototipo resolver la necesidad de obtener información financiera consolidada en formato de indicadores de gestión, necesarios para la toma de decisiones estratégicas en tiempo oportunos para el negocio.

Hipótesis O Pregunta De Investigación

¿Permitirá el prototipo de Business Intelligence proporcionar información confiable y de manera oportuna para el proceso de toma de decisiones estratégicas de la alta gerencia?

VARIABLES DE LA INVESTIGACIÓN

- **Variable independiente:** Prototipo de Business Intelligence
- **Variable dependiente** Indicadores Financieros.

Capítulo II: Marco Teórico

En este capítulo se describe conceptos teóricos y definiciones que forman parte del marco teórico para el desarrollo del prototipo de BI con información financiera para una camaronera.

Fundamentación Teórica

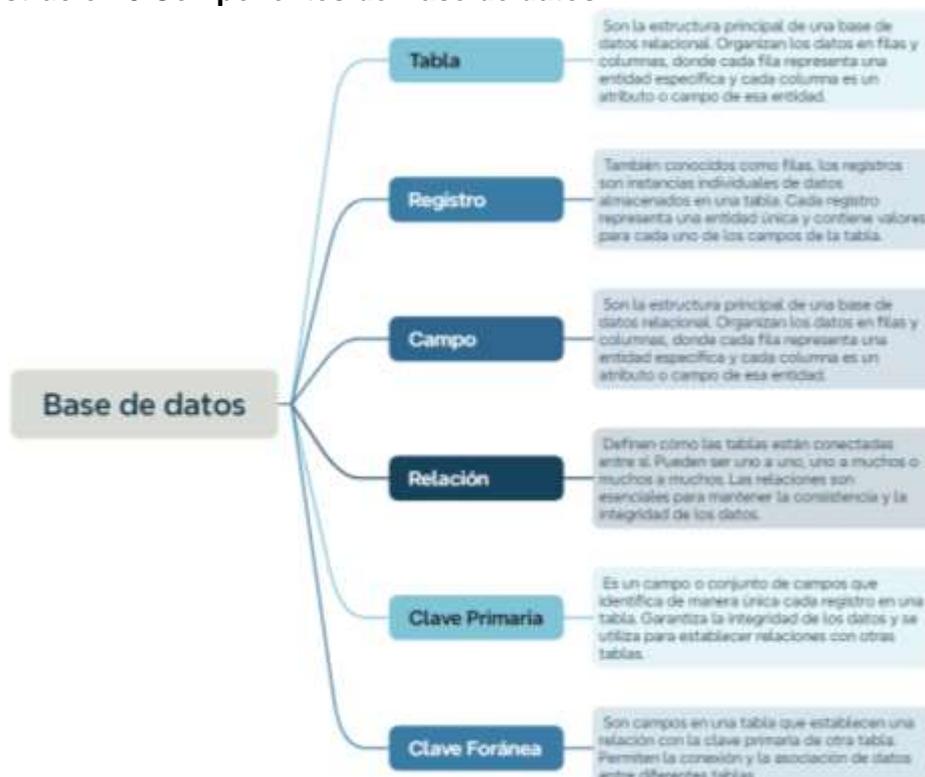
A continuación, se detallan los soportes teóricos que reforzaran el desarrollo del proyecto de titulación.

Base de Datos

Una base de datos es un conjunto estructurado de datos, que se relacionan entre sí. Se diseña, construye y rellena de datos con propósito específico. También es definida como almacén de datos relacionados que representan algunos aspectos del mundo real, del interés del diseñador, (Coronel & Morris, 2018).

Aquí están los componentes fundamentales:

Ilustración 3 Componentes de Base de datos



Nota: Una base de datos se compone de tablas las que a su vez tienen registros y los registros tienen campos los cuales se relacionan entre sí por medio de claves como lo son las primarias y foráneas.

Gestores de Bases de Datos (DBMS):

Un Gestor de Bases de Datos (DBMS) es un software que facilita la creación, el mantenimiento y el uso de bases de datos. Hay diversos tipos de DBMS, y cada uno está diseñado para satisfacer diferentes necesidades y requisitos. (Hueso Ibáñez, 2016)

Algunos de los gestores de bases de datos más conocidos incluyen:

- **Oracle Database:** Es un sistema de gestión de bases de datos relacional (RDBMS⁶) ampliamente utilizado en empresas. Ofrece una amplia gama de funcionalidades y es conocido por su rendimiento y escalabilidad.
- **Microsoft SQL Server:** Es un RDBMS desarrollado por Microsoft. Se utiliza comúnmente en entornos empresariales y se integra bien con otras tecnologías de Microsoft, como el sistema operativo Windows y el entorno de desarrollo .NET.
- **MySQL:** Es un sistema de gestión de bases de datos relacional de código abierto. Es conocido por ser rápido, confiable y fácil de usar. Es ampliamente utilizado en aplicaciones web y en entornos de código abierto.
- **PostgreSQL:** Es un sistema de gestión de bases de datos relacional de código abierto que se destaca por su conformidad con los estándares SQL ⁷y su extensibilidad. Es conocido por su robustez y soporte para características avanzadas.
- **MongoDB:** MongoDB es un sistema de gestión de bases de datos NoSQL orientado a documentos. Almacena datos en formato BSON⁸ (formato binario JSON⁹) y es adecuado para aplicaciones que manejan grandes cantidades de datos no estructurados.

Tipos de Base de Datos

Existen varios tipos de bases de datos, pero la de interés en el proyecto son las del tipo OLTP y OLAP.

OLAP (Online Analytical Processing):

Las bases de datos OLAP están diseñadas para facilitar el análisis y la consulta

⁶ Relational Database Management System.

⁷ Structured query language.

⁸ Binary Javascript Object Notation.

⁹ JavaScript Object Notation.

eficiente de grandes conjuntos de datos. Son ideales para entornos de Data Warehousing y soportan operaciones de consulta complejas. Los datos en una base de datos OLAP suelen estar organizados de manera multidimensional para permitir un rápido análisis de tendencias y agregaciones.(Wrembel & Koncilia, 2011).

Características:

- Estructuras multidimensionales (cubos OLAP).
- Optimizadas para consultas analíticas.
- Apoyo a la agregación y resumen de datos.
- Rendimiento eficiente para análisis de datos.

Ejemplos de bases de datos OLAP: Microsoft SQL Server Analysis Services, IBM Cognos, Oracle OLAP.

OLTP (Online Transaction Processing):

Las bases de datos OLTP están diseñadas para procesar transacciones en tiempo real, lo que significa que son eficientes para insertar, actualizar y eliminar registros en una base de datos. Se utilizan en entornos donde la prioridad es la captura y modificación rápida de datos, como en sistemas de gestión de ventas, inventarios y sistemas de reservas.(Arias, 2015)

Características:

- Diseñadas para transacciones rápidas y concurrencia.
- Estructuras normalizadas para garantizar la integridad de los datos.
- Soporte para transacciones ACID (Atomicidad, Consistencia, Aislamiento, Durabilidad).
- Rendimiento eficiente para operaciones de inserción, actualización y eliminación.

Ejemplos de bases de datos OLTP: MySQL, PostgreSQL, Microsoft SQL Server (en su configuración OLTP).

Framework

Es un marco de trabajo, en la programación es una estructura conceptual y lógica que brinda al usuario herramientas y pautas para el desarrollo y gestión de la creación de un

software; sirve para mejorar la eficiencia, acelerar los tiempos de desarrollo, facilita la colaboración, mejora la calidad de código y aumenta la flexibilidad.

Lenguaje de Programación

Un lenguaje de programación es un conjunto de reglas y convenciones utilizadas para escribir programas informáticos. Estos permiten a los programadores comunicarse con las computadoras y darles instrucciones para realizar tareas específicas. Pueden variar en complejidad y propósito, desde lenguajes de bajo nivel que están más cerca del lenguaje de la máquina hasta lenguajes de alto nivel que son más comprensibles para los humanos.

Python

Python es un lenguaje de programación de alto nivel, interpretado y de propósito general. Es conocido por su sintaxis clara y legible, lo que lo hace ideal para principiantes, pero también es poderoso y versátil, lo que lo convierte en una opción popular entre profesionales de la programación y la ciencia de datos. (Toro Bonilla, 2022)

Algunas características destacadas de Python incluyen su amplio conjunto de bibliotecas y frameworks, la facilidad de aprendizaje, la compatibilidad con múltiples plataformas y su activa comunidad de desarrollo. Se utiliza en una variedad de aplicaciones, desde desarrollo web hasta análisis de datos, inteligencia artificial, automatización de procesos.

Python también se utiliza ampliamente en el contexto de Data Warehousing, aunque no es un framework en sí mismo, sino un lenguaje de programación versátil que puede ser empleado en diversas fases del desarrollo de un Data Warehouse. (Ellison, 2021)

Aquí hay algunas formas en que Python puede contribuir al desarrollo de Data Warehouses:

- ETL (Extract, Transform, Load)
- Conexión a Bases de Datos
- Análisis de datos
- Automatización de Tareas
- API's de Conexión a Servicios en la Nube
- Desarrollo de Modelos Analíticos

- Frameworks Web para Construir Interfaces de Usuario

ETL (Extract, Transform, Load) con Python

ETL es un proceso clave en el desarrollo de un Data Warehouse, donde se extraen datos de diversas fuentes, se transforman según las necesidades y se cargan en el almacén de datos para su análisis.(Kishore Pandey & Ro Schoof, 2023). Python ofrece varias herramientas y bibliotecas que facilitan la implementación del proceso ETL de manera efectiva. Algunas de las bibliotecas y herramientas populares para ETL en Python incluyen:

- **Pandas:** Es una biblioteca de análisis de datos que proporciona estructuras de datos flexibles y eficientes, así como herramientas para trabajar con conjuntos de datos estructurados. Es esencial en la fase de transformación de ETL.
- **Apache Airflow:** Es una plataforma para la orquestación de flujos de trabajo. Permite definir, programar y monitorear flujos de trabajo de ETL de manera eficiente. Además, es extensible y se puede integrar con otras herramientas.
- **Dask:** Ofrece capacidades de procesamiento paralelo y distribuido para operaciones en grandes conjuntos de datos, lo que puede ser útil para la fase de transformación de ETL en entornos distribuidos.
- **Bonobo:** Es un marco de ETL para Python que facilita la definición y ejecución de flujos de trabajo de ETL. Su sintaxis simple y declarativa lo hace fácil de entender y utilizar.
- **Petl:** Es otra biblioteca de ETL para Python que simplifica la manipulación de datos. Proporciona funciones para extraer, transformar y cargar datos de manera eficiente.

Data Warehouse

Se denomina Data Warehouse a un almacén de datos usado para grandes cantidades de información (histórica) sobre una empresa permitiendo a los usuarios un fácil acceso a ella por medio de herramientas de visualización de datos; tiene aplicabilidad en la recolección de datos que respaldaran la toma de decisiones gerenciales; se aplican en diversos sectores como el Análisis de Negocios y Reporting, Planificación Empresarial, gestión de Relaciones con Clientes (CRM) Análisis de Rendimiento Empresarial (BPA) y Análisis de Big Data. (Kimball & Ross, 2013)

Existen herramientas y frameworks que facilitan el desarrollo de Data Warehouse,

estos ofrecen funcionalidades para la extracción, transformación y carga de la data (ETL), almacenamiento y consulta eficiente de datos, A continuación, se nombran los más conocidos:

- Apache Hadoop
- Apache Spark
- Amazon Redshift
- Microsoft Azure Synapse Analytics (anteriormente SQL Data Warehouse)
- Pentaho Data Integration

Pero adicionalmente a los Frameworks tenemos un lenguaje de programación que nos brinda muchas funcionalidades necesarias para el ETL como lo es Python.

Business Intelligence

Según (Pérez Marqués, 2015, pp. 1-22) Business Intelligence es un conjunto de metodologías, aplicaciones, estrategias, prácticas y capacidades que se enfocan en crear y administrar la información para su utilidad en el sector financiero de una organización, permitiendo tomar mejores decisiones a los usuarios.

La implementación exitosa de un proyecto de Business Intelligence implica varios procesos los cuales se describen a continuación:

Ilustración 4 Procesos de Business Intelligence



Nota: En la imagen se detallan cada uno de los procesos de Business Intelligence y sus principales características.

Microsoft Power BI:

Power BI es una suite integral de herramientas de análisis empresarial que permite a los usuarios conectar, visualizar y compartir insights desde sus datos. (Deckler et al., 2022)

Características:

- Conectividad con diversas fuentes de datos.
- Creación de informes y dashboards interactivos.
- Integración con otras herramientas de Microsoft.

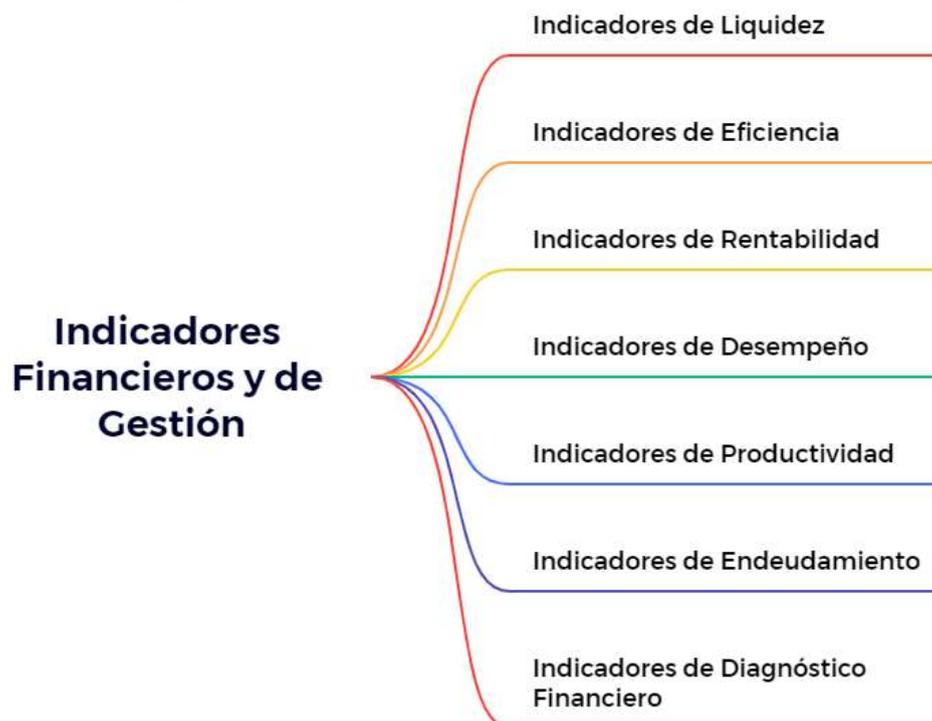
KPI's (Indicadores Clave de Desempeño):

(Parmenter, 2015, pp. 2-24) en su libro indica que los KPI's son métricas cuantitativas utilizadas para evaluar el rendimiento de una organización, departamento o proceso; representan medidas críticas alineadas con los objetivos estratégicos o metas de la empresa. Sirven para la medición de rendimiento proporcionando una medida clara y cuantificable y para la toma de decisiones basada en datos y en el seguimiento del progreso hacia metas estratégicas. A continuación, se indican características relevantes:

- Medición cuantitativa del rendimiento.
- Relacionados con objetivos y metas.
- Facilitan la toma de decisiones.

En la gráfica siguiente se muestran tipos de indicadores que se aplican en los modelos de negocios:

Ilustración 5 Tipos de Indicadores



Nota: En la imagen se enuncian los diferentes indicadores que se pueden ser útiles en una organización

Para este proyecto se tomará en consideración el cálculo de tres tipos de indicadores los cuales son necesarios en esta primera entrega del proyecto del BI para la camaronera.

Indicadores de Desempeño Financieros

Los indicadores financieros son métricas cuantitativas que evalúan el rendimiento financiero de una organización. (S. Kaplan & P. Norton, 2016); estos proporcionan información sobre la salud financiera, la rentabilidad y la eficiencia operativa de la empresa; a continuación, se detallan las fórmulas de los indicadores que se mostrarán en el dashboard.

Tabla 3 Indicadores Financieros

Factor	Indicadores	Fórmulas
Liquidez	Razón Circulante	$RC = \frac{\text{Activo Corriente}}{\text{Pasivo Corriente}}$
	- Si $RC \leq 1$ → incapacidad para cumplir con obligaciones a corto plazo.	
	- Si $RC > 2$ → capacidad para cumplir con obligaciones a corto plazo.	
	Prueba Acida	$PA = \frac{\text{Activo Corriente} - \text{Inventario}}{\text{Pasivo Corriente}}$
	- Si $PA > 1$ → capacidad para cumplir con obligaciones a corto plazo.	
	- Si $PA < 1$ → incapacidad para cumplir con obligaciones a largo plazo.	
	Fondo de Maniobra	$FM = \text{Activo corriente} - \text{Pasivo corriente}$
	- Si $FM \geq 0$ → equilibrio financiero.	

Factor	Indicadores	Fórmulas
	- Si FM ≤ 0 → desequilibrio financiero.	
Endeudamiento	Total	$ET = \frac{\text{Pasivo}}{\text{Activo}}$
	- Si ET > 100%* → las deudas de la empresa superan sus activos. *Debe analizarse de acuerdo con el tipo de empresa.	
	A Corto Plazo	$ECP = \frac{\text{Pasivo corriente}}{\text{Patrimonio neto}}$
	ECP ≥ 20-30% → buen poder de negociación a corto plazo.	
	A Largo Plazo	$ELP = \frac{\text{Pasivo No corriente}}{\text{Patrimonio neto}}$
	- Si ELP > 1 → liquidez en riesgo. - Si ELP < 1 → mayor solvencia.	
	Apalancamiento Total	$AT = \frac{\text{Pasivo}}{\text{Patrimonio neto}}$
	Si AT < 1 → capacidad de cobertura de obligaciones con terceros.	
Rentabilidad	Margen de Utilidad Bruta	$MUB = \frac{\text{Utilidad Bruta}}{\text{Ventas Netas}}$
	Permite conocer la eficacia operativa de la empresa	
	Sobre las Inversiones (Return to Assets)	$ROI = \frac{\text{Utilidad Neta}}{\text{Activos}}$
	Si ROI > 5% → la empresa es rentable	
	Sobre el Patrimonio (Return to Equity)	$ROE = \frac{\text{Utilidad Neta}}{\text{Patrimonio}}$
	- Si ROE > ROI → mayor rentabilidad financiera. - Si ROE = ROI → la empresa no tiene deudas. - Si ROE < ROI → el coste medio de la deuda es superior a la rentabilidad.	
	Sobre Ventas	$RV = \frac{\text{Utilidad Neta}}{\text{Ventas Netas}}$
	Permite conocer la eficacia operativa con respecto a las ventas	

Nota: La tabla presenta las fórmulas de los indicadores y su interpretación financiera.

Capítulo III: Metodología De La Investigación

Para la realización de este trabajo de titulación se hará uso del método cuantitativo de investigación para la obtención de datos y de la metodología Hefesto para el desarrollo del proyecto, lo cual se destalla a continuación:

Método De Investigación Cuantitativa

La aplicación de este método lleva un proceso ordenado que asegura la validez y confiabilidad del proyecto a realizar; a continuación, se detallan los pasos a seguir para su implementación:

- **Definición del Problema de Investigación:** Este método nos permite identificar claramente el problema o pregunta de investigación que se va a resolver con el desarrollo del proyecto, asegurando que sea específico, medible y relevante.
- **Revisión de fundamentación teórica:** Nos indica que para el desarrollo de este proyecto se debe tener contenido literario que respalde la ejecución del proyecto.
- **Formulación de Hipótesis o Pregunta de Investigación:** De acuerdo con lo revisado y con su efectiva argumentación se procede con base al problema la formulación de una pregunta de investigación como se lo realizó en el capítulo I del presente trabajo.
- **Diseño de la Investigación:** Es este paso podemos elegir el medio de recolección de datos respecto a un experimento controlado, para este proyecto se aplicará encuestas.
- **Identificación de Variables:** Al hacer uso de este método debemos tener identificada cual sería la variable dependiente e independiente, asegurando que las mismas sean medibles y susceptibles de ser cuantificadas. Para el desarrollo de este proyecto se obtuvo como Variable Independiente: Prototipo de Business Intelligence y como Variable Dependiente: Indicadores Financieros.
- **Selección de la Muestra:** La aplicación de este método nos permite determinar una población objetivo y una muestra medible y suficiente para permitir generalizaciones significativas.
- **Recopilación de Datos:** Al ser un método cuantitativo, se procede con la realización de encuestas que brinden respuestas medibles para su posterior interpretación

- **Análisis e Interpretación de Resultados:** Haciendo uso de los datos recolectados en la etapa previa, este método nos exige un análisis e interpretación de los resultados obtenidos para poder indicar recomendaciones futuras en la realización de proyectos similares.

Población y Muestra

Población

Para el desarrollo de este proyecto se identificó como población a los reportes contables que son usados para la gestión manual de los indicadores financieros necesarios para la toma de decisiones gerenciales. Sin embargo, para la comprobación de esta gestión se tiene presente una muestra de personas a quienes se les realizara una encuesta estructurada.

La población de reportes identificados adjuntos en el Anexo 2 hace referencia al periodo 2021-2022 de manera mensual, que forman parte de la información histórica de la organización y la del año en curso que será ingresado para la actualización de datos a medida que se desarrolla el proyecto.

Tabla 4 Población Reportes Financieros

Objetos de Estudio	Históricos	Año Actual 2023	Total de Reportes
Balance General	24	11	35
Estado de Pérdidas y Ganancias	24	11	35
Indicadores Financieros	24	11	35
Total	72	33	105

Nota: Se detalla la población de reportes contable seleccionados.

La encuesta se aplicó a los integrantes de alta gerencia que constituyen 8 personas que interactuarán directamente con el prototipo; a continuación, el detalle:

Tabla 5 Población Alta Gerencia

Sujetos de Estudio	# de personas
Gerentes	3
Socios Mayoritarios	5
Total	8

Nota: Se detalla la muestra a la cual se le aplicara la encuesta.

Muestra

La muestra de reportes a utilizar en el proyecto será de los años 2021 y 2022 generados de manera mensual estos serán parte de los datos históricos, considerando el año 2023 como la actualización de datos de los indicadores, un total de 105 reportes (ver tabla # 4).

Recolección de Información

Para la recolección de datos se realizó una encuesta estructurada a la alta gerencia con la cual podemos concluir que no tienen experiencias previas con herramienta de Business Intelligence, pero si conocen de los indicadores financieros que son los que se calculan de forma manual.

A continuación, se muestra el formato de la encuesta aplicada.

Ilustración 6 Formato de Encuesta

1. ¿Tiene experiencia previa en el uso de dashboards con indicadores financieros? *

Sí

No

2. ¿Cuáles son los indicadores financieros más críticos para su departamento? *

Liquidez

Endeudamiento

Rentabilidad

Gestión o Eficiencia

3. Considera algún otro indicador primordial para la toma de decisiones, menciónelo.

Tu respuesta

4. ¿Tiene preferencias específicas en cuanto a la visualización de datos en el dashboard? *

Gráficos

Tablas

Mapas

5. ¿Prefiere recibir alertas automáticas en caso de desviaciones significativas en los indicadores? *

Sí

No

6. ¿Requiere acceder al dashboard desde dispositivos móviles? *

Sí

No

7. ¿Requiere que el nuevo dashboard se integre con el software existente en su departamento? *

Sí

No

8. ¿Considera importante recibir capacitación para utilizar el nuevo dashboard? *

Sí

No

9. ¿Considera importante el uso del Dashboard para la toma de decisiones mas efectiva?

- Sí
- No

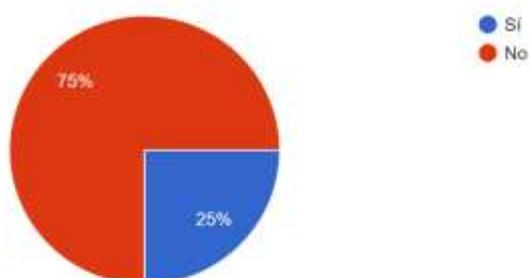
Nota: Preguntas realizadas a la muestra de personas de la alta gerencia de la camaronera.

A continuación, se observa la tabulación de los resultados obtenidos de la encuesta:

Ilustración 7 Experiencia con Dashboard

1. ¿Tiene experiencia previa en el uso de dashboards con indicadores financieros?

8 respuestas

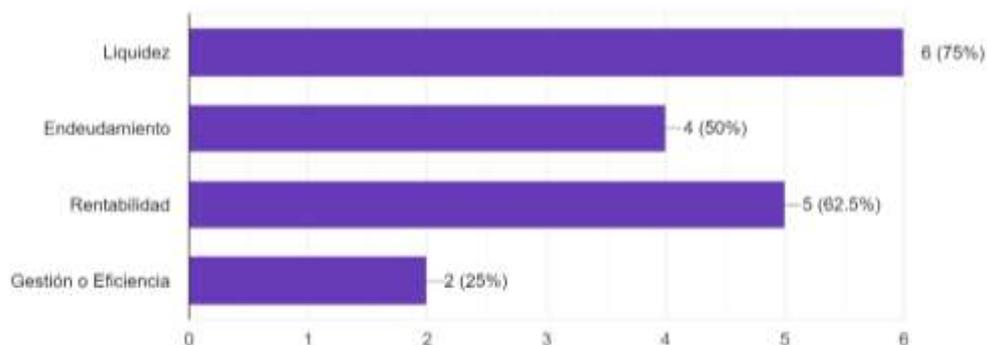


Nota: El gráfico circular muestra que la mayor parte de encuestados desconocen el uso de un dashboard

Ilustración 8 Indicadores Financieros

2. ¿Cuáles son los indicadores financieros más críticos para su departamento?

8 respuestas

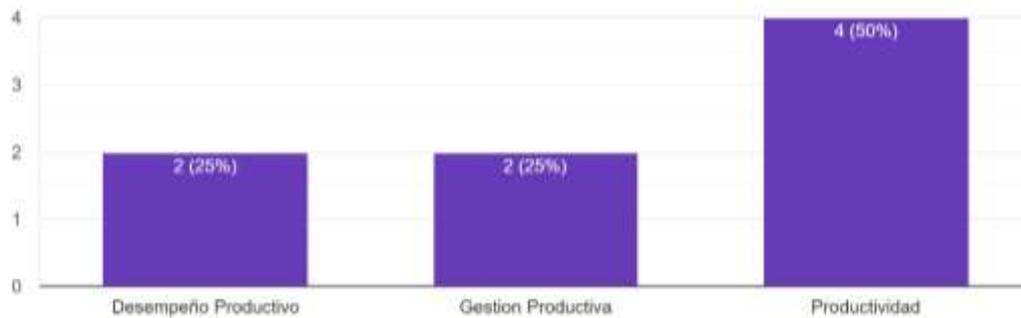


Nota: Podemos observar que son de vital importancia para la empresa los indicadores de rentabilidad, liquidez y endeudamiento.

Ilustración 9 Nuevas propuestas de Indicadores

3. Considera algún otro indicador primordial para la toma de decisiones, menciónelo.

8 respuestas

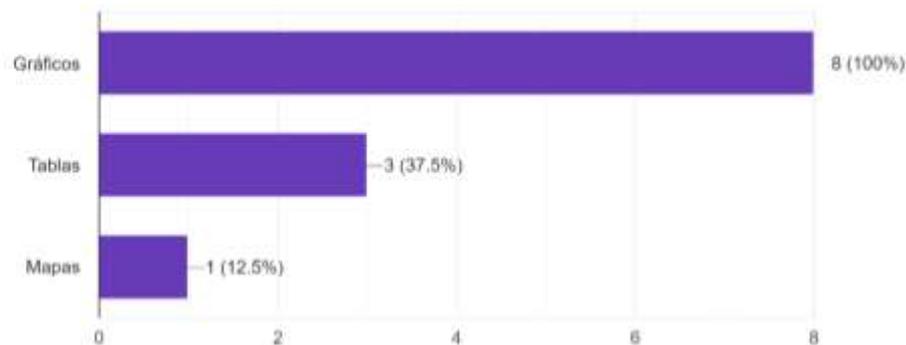


Nota: Podemos observar que la productividad es un indicador importante que se puede implementar a futuro.

Ilustración 10 Preferencias de Visualización

4. ¿Tiene preferencias específicas en cuanto a la visualización de datos en el dashboard?

8 respuestas



Nota: Podemos observar que de preferencia tienen los gráficos para visualizar la información

Ilustración 11 Alertas automáticas

5. ¿Prefiere recibir alertas automáticas en caso de desviaciones significativas en los indicadores?

8 respuestas



Nota: Aquí podemos apreciar que solicitan las alertas automáticas.

Ilustración 12 Accesibilidad

6. ¿Requiere acceder al dashboard desde dispositivos móviles?

8 respuestas



Nota: Los usuarios demandan accesibilidad de manera móvil.

Ilustración 13 Integración de Sistemas

7. ¿Requiere que el nuevo dashboard se integre con el software existente en su departamento?

8 respuestas



Nota: Es importante la integración de las herramientas de trabajo para los usuarios.

Ilustración 14 Capacitación

8. ¿Considera importante recibir capacitación para utilizar el nuevo dashboard?

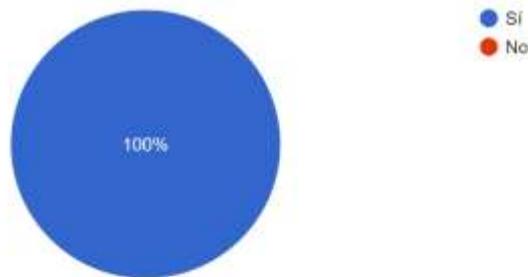
8 respuestas



Nota: Al tener una nueva herramienta es necesaria la capacitación a los usuarios que harán uso de ella.

Ilustración 15 Importancia del Proyecto

9. ¿Considera importante el uso del Dashboard para la toma de decisiones mas efectiva?
8 respuestas



Nota: Podemos observar que consideran muy importante el desarrollo de un Dashboard para la toma de decisiones efectivas.

Una vez culminada la realización de las encuestas se procede a el análisis de las respuestas obtenidas, de las cuales podemos concluir lo siguiente:

1. De la muestra de personas que van a hacer uso de la nueva herramienta solo el 25% tiene experiencia previa con un Dashboard, podemos indicar que será una herramienta de gran utilidad para ellos.
2. A pesar de no tener experiencias con Dashboard tienes conocimientos de los indicadores necesarios para el control financiero de la camaronera.
3. El usuario demanda accesibilidad móvil, alertas automáticas, y esto sería necesario para notificar de algún cambio que ponga en riesgo la estabilidad de la empresa.
4. A pesar de conocer de los indicadores demandan capacitación para el uso de la herramienta nueva a implementar.

Luego del Análisis de Resultados se concluye que ante estas necesidades y debido al modelo de negocios que maneja la camaronera, la solución planteada de BI deberá contener indicadores Financieros con sus respectivas alertas y retroalimentación. Además de la capacitación previa del desarrollo para el uso correcto de los usuarios finales.

Metodología de Construcción de Software

Existen diversas metodologías de desarrollo de proyecto que permiten la correcta implementación del BI, a continuación, se muestra comparativo de evaluación de metodologías que se pudieren aplicar para el desarrollo del proyecto.

Tabla 6 Comparación de Metodologías

Criterios	METODOLOGÍAS		
	Kimball	Hefesto	Inmon
Autor	Ralph Kimball	Ricardo Bernabéu	Bill Inmon
Alcance	Departamentos Individuales	Departamentos Individuales	Toda la Compañía
Modelado	Modelado Dimensional	Modelado Dimensional	Modelado Normalizado en tercera forma normal
Énfasis	Data mart	Data mart y Data Warehouse	Data Warehouse
Experiencia TH.	Especialización media	Especialización alta	Especialización alta
Tiempo de implementación	Rápido	Fases de desarrollo y de despliegue relativamente cortas	Demanda más tiempo, Por la implementación completa del Data Warehouse

Nota: Se muestra la comparativa de tres metodologías, para evaluar la mejor para el desarrollo del proyecto.

A partir de la comparación realizada, se selecciona la metodología Hefesto para el desarrollo del prototipo debido a que se requiere la construcción de un Data Warehouse enfocado en un área específica del modelo de negocio, en este caso sería la financiera.

Esta metodología, tiene como objetivo facilitar la construcción del DW desde cero, brindando información que permitirá mejorar su desempeño, inicia con la recolección de requerimientos del usuario y finaliza con la confección del esquema lógico y sus respectivos procesos ETL. (Bernabeu & García Mattio, 2018)

Hefesto tiene 4 fases, las que se indican a continuación:

Ilustración 16 Metodología HEFESTO



Nota: Características de cada paso de la metodología según la última actualización del 2018 del libro de Bernabeu.

En la siguiente gráfica se presentan las tareas que se realizarán en cada fase de la metodología:

Ilustración 17 Tareas Metodología Hefesto



Nota: En su última actualización del 2018 del libro de Bernabeu indica las tareas de cada paso de la metodología HEFESTO.

Capítulo IV: Propuesta Tecnológica

En este capítulo se documenta el proceso del desarrollo del proyecto a nivel tecnológico, se detallará la aplicación de la metodología y sus 4 fases y la creación del prototipo de Dashboard en el aplicativo de Power BI.

Situación Actual

Actualmente la camaronera para la toma de decisiones estratégica realiza de manera manual la obtención de indicadores financieros por medio de la descarga de reportes contables, realizando reprocesos en los mismos debido a la falta de actualización de información. A la alta gerencia le cuesta en factor tiempo y monetario la falta de efectividad al momento de tomar decisiones.

Origen de Datos

La fuente de datos será la base transaccional actual de la camaronera específicamente las tablas que contienen la información con la que se genera los reportes contables como lo son el balance general y los estados de pérdida y ganancia.

Herramientas De Desarrollo

Selección de Herramientas

De acuerdo al modelo del negocio de la camaronera, las herramientas seleccionadas serán de las que tiene experiencia la organización como lo es el motor de base de datos, el lenguaje de programación Python, a excepción del visualizador de datos sin embargo es el seleccionado por la alta gerencia.

Para mantener la integridad de los datos se hará uso de MySQL que es el motor actual de base de la camaronera, en el cual se tiene los datos transaccionales y se procederá a crear la base de DW.

Para la aplicación de los procesos ETL la empresa propuso como principal opción el lenguaje de Python por este motivo en conjunto con sus librerías como Pandas, MySQL, Datatime se propone la solución de creación de script de carga de datos tanto para la parte histórica como para la actualización.

Power BI es una herramienta de fabricación de Dashboard con licenciamiento, el cual es demandado por los gerentes de la empresa, por esta razón es seleccionado para el

desarrollo del proyecto; se procedió con la descarga desde la tienda oficial de Microsoft, y el pago de la licencia fue realizada por parte de la organización.

A continuación, se procede con la aplicación de las fases de la metodología de Hefesto para el desarrollo exitoso del proyecto de BI.

Fase 1: Análisis de Requerimientos

1.1. Preguntas

Con las entrevistas realizadas como se observa en el informe del Anexo 1 se procedió a analizar y entender el proceso manual de obtención de indicadores y se definió los requerimientos de acuerdo a la necesidad de la siguiente manera.

Tabla 7 Requerimientos obtenidos

Necesidad	Preguntas	Reportes	Filtros
Control de Liquidez	<ul style="list-style-type: none"> • ¿Desea saber acerca de la Liquidez de la empresa? • ¿Desea reportes de los indicadores de Liquidez de manera histórica? 	Balance General	Periodo y mes
Control de Endeudamiento	<ul style="list-style-type: none"> • ¿Desea saber acerca del Endeudamiento de la empresa? • ¿Desea reportes de los indicadores de Endeudamiento de manera histórica? 	Balance General	Periodo y mes
Control de Rentabilidad	<ul style="list-style-type: none"> • ¿Desea saber acerca de la Rentabilidad de la empresa? • ¿Desea reportes de los indicadores de Rentabilidad de manera histórica? 	Balance General Estado de Pérdidas y Ganancias	Periodo y mes

Nota: Se detallan las preguntas obtenidas necesarias para cubrir con los requerimientos solicitados por los usuarios, resultado de las reuniones con la alta gerencia.

1.2. Indicadores y Perspectivas

Una vez obtenidas las preguntas se procede con la descomposición para la obtención de las métricas que se usarán y las perspectivas analíticas que intervendrán en el proceso de toma de decisiones. Se plantearon el cálculo de 11 indicadores para responder las preguntas realizadas, estos se obtendrán de los 105 reportes contables extraídos del sistema de la camaronera.

1.3. Modelo conceptual

Para el desarrollo del proyecto se obtuvo los querys que generan el balance general y el estado de perdida y ganancias de los cuales nos guiaremos para la generación del modelo conceptual, se muestran a continuación.

Ilustración 18 Query Balance General

```

1 SELECT IF(Id=1,@ActivasD:=D,null),IF(Id=2,@PasivasD:=D,null),IF(Id=3,@CapitalD:=D,null),IF(Id=3,@CapitalH:=H,null)
2 ,IF(Id=4,@RentasD:=D,null)
3 FROM `a021t1` WHERE Id<5;
4 CREATE TEMPORARY TABLE p02350_Temp1 (formulario_101 VARCHAR(15),CC VARCHAR(15),SC VARCHAR(15),CTATR VARCHAR(15), SACT DECIMAL(18,2), KEY `Codigo`
5 INSERT INTO p02350_Temp1 ( formulario_101, CC, SC, CTATR, SACT )
6 SELECT ifnull(a02201.formulario_101,'') as formulario_101, trim(cast(ifnull(cc.siglas, '') as char)) as CC
7 , trim(cast(ifnull(divi.codigo_oc, '') as char)) as SC,CTATR, sum(if(substring(CTATR,1,1) = @ActivasD, VALTR,VALTR*-1)) AS SACT
8 FROM a02300 as cont
9 LEFT JOIN a02200 ON a02200.codct= cont.ctatr
10 LEFT JOIN a02201 on a02200.codct=a02201.codct and cont.CO_PERIODO = a02201.PERIODO
11 LEFT JOIN a06t9 as cc ON a02200.cc= cc.codigo
12 LEFT JOIN a06t1 as divi ON a02200.`DIV`= divi.codigo and divi.cod_tabla = 'DV'
13 WHERE (substring(CTATR,1,1) = @ActivasD OR substring(CTATR,1,1) = @PasivasD OR substring(CTATR,1,1) = @CapitalD)
14 AND ESTTR=0
15 And FECTR >= '2023-01-01' And FECTR <= '2023-12-31'
16 AND cheque_impreso = 1 AND ifnull(ajuste,0) = 0
17 GROUP BY CTATR
18 ORDER BY CTATR;
19 CREATE TEMPORARY TABLE p02350_Temp2
20 (CTATR VARCHAR(15), SACT DECIMAL(18,2), KEY `Codigo` ('CTATR'));
21
22 INSERT INTO p02350_Temp2 ( CTATR, SACT )
23 SELECT CTATR, SACT from p02350_Temp1;
24 SELECT a02200.nivel AS NIVCT, trim(cast(ifnull(cc.siglas, '') as char)) as CC, trim(cast(ifnull(divi.codigo_oc, '') as char)) as SC
25 , ifnull(a02201.formulario_101,'') as formulario_101,a02200.CODCT, DESCT, Tipo AS TIT, SACT
26 FROM a02200
27 LEFT JOIN p02350_Temp1 ON a02200.CODCT=p02350_Temp1.CTATR
28 LEFT JOIN a02201 on a02200.codct=a02201.codct and a02201.PERIODO = '2023'
29 LEFT JOIN a06t9 as cc ON a02200.cc= cc.codigo
30 LEFT JOIN a06t1 as divi ON a02200.`DIV`= divi.codigo and divi.cod_tabla = 'DV'
31 WHERE (substring(a02200.CODCT,1,1) = @ActivasD OR substring(a02200.CODCT,1,1) = @PasivasD Or substring(a02200.CODCT,1,1)= @CapitalD)
32 AND a02200.TIT = 'M'
33 union
34 SELECT a02200.nivel AS NIVCT, trim(cast(ifnull(cc.siglas, '') as char)) as CC ,trim(cast(ifnull(divi.codigo_oc,'') as char)) as SC
35 , ifnull(a02201.formulario_101,'') as formulario_101, a02200.CODCT, DESCT, Tipo AS TIT, SACT
36 FROM a02200
37 LEFT JOIN a02201 on a02200.codct=a02201.codct and a02201.PERIODO = '2023'
38 LEFT JOIN (select t.codct as CTATR, sum(d.sact) as SACT from p02350_Temp2 as d, a02200 as m, a02200 as t where m.codct = d.ctatr and t.tit = 'T'
39 and m.tit = 'M' and m.codct like concat(t.codct,'%') group by t.codct
40 ) as m ON a02200.CODCT=m.CTATR
41 LEFT JOIN a06t9 as cc ON a02200.cc= cc.codigo
42 LEFT JOIN a06t1 as divi ON a02200.`DIV`= divi.codigo and divi.cod_tabla = 'DV'
43 WHERE (substring(a02200.CODCT,1,1) = @ActivasD OR substring(a02200.CODCT,1,1) = @PasivasD Or substring(a02200.CODCT,1,1)= @CapitalD) and a02200.`
44 ORDER BY CODCT;
45
46 DROP TABLE p02350_Temp1;
47 DROP TABLE p02350_Temp2;
48 SELECT Sum(VALTR) AS Total1
49 FROM a02300
50 WHERE CTATR >= @RentasD AND ESTTR=0 AND cheque_impreso = 1 And FECTR >= '2023-01-01' And FECTR <= '2023-12-31' AND ifnull(ajuste,0) = 0 ;
51
52 SELECT Sum(VALTR) AS Total2
53 FROM a02300
54 WHERE CTATR >= @PasivasD AND CTATR <= @CapitalH
55 AND ESTTR=0
56 AND cheque_impreso = 1
57 And FECTR >= '2023-01-01' And FECTR <= '2023-12-31'
58 AND ifnull(ajuste,0) = 0 ;

```

Message	Result 1	Result 2	Result 3	Result 4	Profile	Status		
	NIVCT	CC	SC	formulario_101	CODCT	DESCT	TIT	SACT
	1				1	ACTIVO	Titulo	1249389,82
	2				11	ACTIVO CORRIENTE	Titulo	1249389,82
	0				110601037	INV. MULTIPROYECTOS	Titulo	(Null)
	3				111	CAJA	Titulo	337987,14
	4				11101	CAJA GENERAL	Titulo	337987,14
	2 ECA	PC			11101001	CAJA GENERAL IPSP	Movimiento	337987,14
	1 MAT	OC			11101002	CAJA GENERAL DOLAR (SR)	Movimiento	(Null)

Nota: Query extraído de la generación del reporte balance general en el sistema actual de la camaronera

Ilustración 19 Query Estado Pérdidas y Ganancias

The screenshot shows a SQL query execution window with the following query:

```

1 SELECT MAX(IF(Id=4,@RentasD:=D,null)) AS Codigo4,MAX(IF(Id=5,@CostosD:=D,null)) AS Codigo5
2 ,MAX(IF(Id=6,@GastosD:=D,null)) AS Codigo6,MAX(IF(Id=7,@ORentasD:=D,null)) AS Codigo7,MAX(IF(Id=8,@OGastosH:=H,null))
3 FROM `a02t1` WHERE Id>3 AND Id<=8;
4 CREATE TEMPORARY TABLE p02360_Temp1 (CTATR VARCHAR(15), SACT DECIMAL(18,2), KEY `Codigo` (`CTATR`));
5 INSERT INTO p02360_Temp1 ( CTATR, SACT )
6 SELECT CTATR,Sum(IF(CTATR Like concat(@RentasD,'%') OR CTATR Like concat(@ORentasD,'%'),VALTR*-1,VALTR)) AS SACT
7 From a02300
8 WHERE (CTATR >= @RentasD AND CTATR <= @OGastosH)
9 AND ESTTR=0
10 And FECTR >= '2023-01-01' AND FECTR <= '2023-12-31'
11 AND cheque_impreso = 1
12 AND ifnull(ajuste,0) = 0
13 GROUP BY CTATR;
14 ORDER BY CTATR;
15 CREATE TEMPORARY TABLE p02360_Temp2 (CTATR VARCHAR(15), SACT DECIMAL(18,2), KEY `Codigo` (`CTATR`));
16 INSERT INTO p02360_Temp2 ( CTATR, SACT )
17 SELECT CTATR, SACT from p02360_Temp1;
18 SELECT a02200.nivel AS NIVCT, ifnull(a02201.formulario_101, a02200.formulario_101) as formulario_101, cc.siglas as CC , divi.codigo_oc as SC, a02200.CODCT
19 , DESCT, Tipo AS TIT, SACT
20 FROM a02200
21 LEFT JOIN p02360_Temp1 ON a02200.CODCT=p02360_Temp1.CTATR
22 LEFT JOIN a02201 on a02200.CODCT = a02201.CODCT and a02201.PERIODO = '2023'
23 LEFT JOIN a06t9 as cc ON a02200.cc= cc.codigo
24 LEFT JOIN a06t1 as divi ON a02200.DIV = divi.codigo and divi.cod_tabla = 'DV'
25 WHERE (a02200.CODCT >= @RentasD AND a02200.CODCT <= @OGastosH)
26 AND a02200.TIT = 'M'
27 union
28 SELECT a02200.nivel AS NIVCT, ifnull(a02201.formulario_101, a02200.formulario_101) as formulario_101 , cc.siglas as CC , divi.codigo_oc as SC, a02200.CODCT
29 , DESCT, Tipo AS TIT, SACT
30 FROM a02200
31 LEFT JOIN a02201 on a02200.CODCT = a02201.CODCT and a02201.PERIODO = '2023'
32 LEFT JOIN a06t9 as cc ON a02200.cc= cc.codigo
33 LEFT JOIN (select t.codct as CTATR, sum(d.sact) as SACT
34 from p02360_Temp2 as d, a02200 as m, a02200 as t
35 where m.codct = d.ctatr
36 and t.tit = 'T'
37 and m.tit = 'M'
38 and m.codct like concat(t.codct,'%')
39 group by t.codct
40 ) as m ON a02200.CODCT=m.CTATR
41 LEFT JOIN a06t1 as divi ON a02200.DIV = divi.codigo and divi.cod_tabla = 'DV'
42 WHERE (a02200.CODCT >= @RentasD AND a02200.CODCT <= @OGastosH)
43 and a02200.TIT = 'T'
44 ORDER BY CODCT;
45 DROP TABLE p02360_Temp1;
46 DROP TABLE p02360_Temp2;
    
```

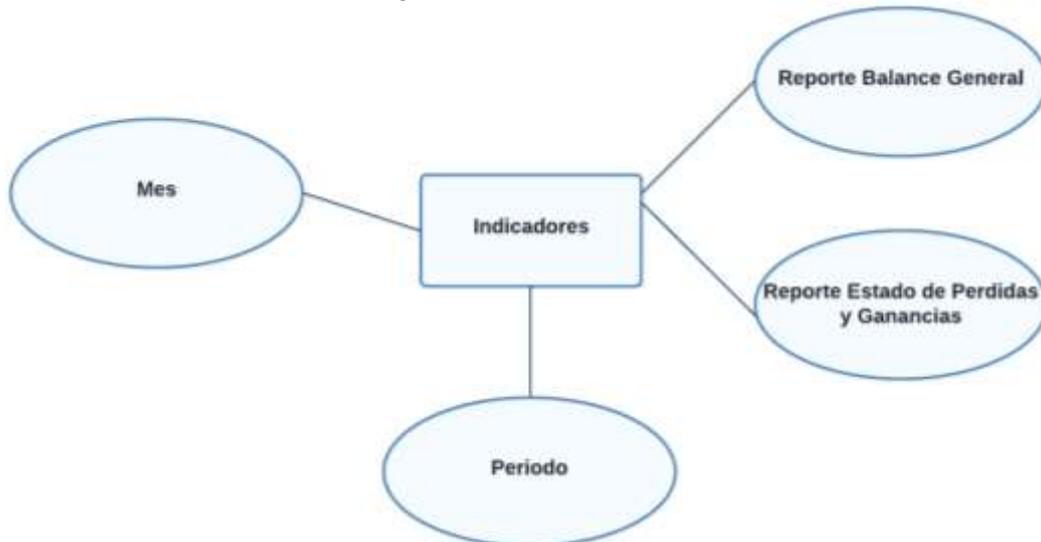
The results table is as follows:

Message	Result 1	Result 2	Profile	Status				
	NIVCT	formulario_101	CC	SC	CODCT	DESCT	TIT	SACT
▶	1	(Null)	(Null)	(Null)	4	INGRESOS	Titulo	1999214,00
	2	(Null)	(Null)	(Null)	40	INGRESOS POR VENTAS	Titulo	1999220,44
	3	(Null)	(Null)	(Null)	401	EXPORTACIONES	Titulo	-779,56
	4	(Null)	(Null)	(Null)	40101	VENTAS AL EXTERIOR	Titulo	(Null)
	6	6009	ECA	PC	40101001	VTAS.CAMARON EXPORTAC	Movimiento	(Null)
	5	6009	EPE	PP	40101002	VTAS.PESCADO EXPORTACI	Movimiento	(Null)
	5	(Null)	ECA	PC	40101003	VTAS.MANGO/LIMON	Movimiento	(Null)
	5	(Null)	ECA	PC	40101004	VTAS LIMON/EXTERIOR	Movimiento	(Null)
	5	(Null)	ECA	PC	40101005	VTAS.CHULETAS PREC.EXPC	Movimiento	(Null)
	5	6009	VAG	VA	40101006	VTAS.HAMBURGUESAS EXP	Movimiento	(Null)
	5	(Null)	ECA	PC	40101007	VTAS.CHULETAS APANADA	Movimiento	(Null)
	1	6009	ECA	TC	40101008	VTAS. EXPORTACION TROP	Movimiento	(Null)
	1	6009	ECA	PA	40101009	VTAS. EXPORTACION PT4 L	Movimiento	(Null)
	5	603	VAG	VA	40101010	VENTA BOCADITO APANAD	Movimiento	(Null)

Nota: Query extraído de la generación del reporte de estados de pérdidas y ganancias en el sistema actual de la camaronera

Después de revisar la información que se obtiene de los datos transaccionales con la ejecución de los queries se procedió a la elaboración del modelo conceptual, el cual se detalla a continuación.

Ilustración 20 Modelo Conceptual



Nota: Se muestra el modelo conceptual que se espera estructurar lógicamente en el DW

Fase 2: Análisis de Data Sources (OLPT)

Esta Fase nos ayudará a establecer los cálculos de los indicadores para poder determinar el mapeo del Modelo conceptual para generar uno ampliado.

2.1. Hechos e Indicador

Indicadores de Liquidez:

Estos indicadores miden la capacidad que tiene la camaronera para pagar sus deudas a corto plazo, es decir convertir sus activos corrientes en efectivo.

Indicadores de Endeudamiento:

Permiten determinar el nivel de endeudamiento que puede asumir la camaronera, teniendo como base su capital propio o patrimonio.

Indicadores de Rentabilidad:

Los indicadores de rentabilidad determinan la sostenibilidad en el tiempo de la camaronera.

2.2. Mapeo

En este paso las tablas transaccionales son los reportes generados almacenado en dos DataFrame, uno contiene los reportes de los balances generales de los dos años históricos y el año vigente (2023), el segundo contiene los reportes de estados de pérdidas y ganancias de los mismos periodos.

Se procedió a realizar el mapeo de la información extraída en dos Dataframe con ayuda del lenguaje Python, para de esta manera facilitar la normalización de los datos; el contenido de estos se detalla a continuación:

Ilustración 21 Dataframe de Balance General

```
df_balance_general.head(20)
```

PERIODO	MES	NIVCT	CC	SC	FORM101	CODCTA	DESCT	TIT	SACT
0	2020	1	1			1	ACTIVO	Titulo	5.038000e+08
1	2020	1	2			11	ACTIVO CORRIENTE	Titulo	1.853460e+08
2	2020	1	0			110601037	INV. MULTIPROYECTOS	Titulo	NaN
3	2020	1	3			111	CAJA	Titulo	6.890030e+04
4	2020	1	4			11101	CAJA GENERAL	Titulo	4.015030e+04
5	2020	1	2	ECA	PC	311 11101001	CAJA GENERAL IPSP	Movimiento	1.503000e+02
6	2020	1	1	MAT	OC	311 11101002	CAJA GENERAL DOLAR (SRA INES)	Movimiento	4.000000e+04
7	2020	1	5	ECA	PC	11101003	CAJA GENERAL DOLAR (PLANTA)	Movimiento	NaN
8	2020	1	5			11101004	CAJA GENERAL DOLAR (D-PLANTE)	Movimiento	NaN
9	2020	1	4			11102	CAJA CHICA	Titulo	2.875000e+04
10	2020	1	5	ECA	PC	311 11102001	CAJA CHICA ADMINIST IPSP	Movimiento	2.100000e+03
11	2020	1	5	BAL	DV	11102002	CAJA CHICA ADMINIST.DAVIPA	Movimiento	NaN
12	2020	1	5	LAB	LI	311 11102003	CAJA CHICA ADMINIST BORMAN	Movimiento	0.000000e+00
13	2020	1	5	ECA	PC	11102004	CAJA CHICA IPSP	Movimiento	NaN
14	2020	1	5	ECA	PC	11102005	CAJA CHICA MANT.MECANICO DAVIP	Movimiento	NaN
15	2020	1	6	ECA	PC	11102006	CAJA CHICA MANT.1 PLANTA IPSP	Movimiento	NaN
16	2020	1	5	BAL	DV	11102007	CAJA CHICA MANT. PLANTA DAVIPA	Movimiento	NaN
17	2020	1	5	ECA	PC	11102008	CAJA CHICA CHÓFERES IPSP	Movimiento	NaN
18	2020	1	5	ECA	PC	311 11102009	CAJA CHICA PERSONAL IPSP	Movimiento	NaN
19	2020	1	5	EPE	PP	11102010	CAJA CHICA PESCADO	Movimiento	NaN

Nota: Almacena la información del balance general de manera mensual de los años 2022-2023

Ilustración 22 Dataframe de Estado de Pérdidas y Ganancias

```
df_estado_ggg.head(20)
```

PERIODO	MES	NIVCT	CC	SC	FORM101	CODCTA	DESCT	TIT	SACT	
0	2020	1	1.0	None	None	None	4	INGRESOS	Titulo	50353076.87
1	2020	1	2.0	None	None	None	40	INGRESOS POR VENTAS	Titulo	50364274.82
2	2020	1	3.0	None	None	None	401	EXPORTACIONES	Titulo	47615707.59
3	2020	1	4.0	None	None	None	40101	VENTAS AL EXTERIOR	Titulo	47623424.23
4	2020	1	6.0	6009	ECA	PC	40101001	VTAS.CAMARON EXPORTACION PTA - 1	Movimiento	22443554.70
5	2020	1	5.0	6009	EPE	PP	40101002	VTAS.PESCADO EXPORTACIONES	Movimiento	351677.88
6	2020	1	5.0	None	ECA	PC	40101003	VTAS.MANGO/LIMON	Movimiento	NaN
7	2020	1	5.0	None	ECA	PC	40101004	VTAS LIMON/EXTERIOR	Movimiento	NaN
8	2020	1	5.0	None	ECA	PC	40101005	VTAS.CHULETAS PREC.EXPORT.	Movimiento	NaN
9	2020	1	5.0	6009	VAG	VA	40101006	VTAS.HAMBURGUESAS EXPORT.	Movimiento	NaN
10	2020	1	5.0	None	ECA	PC	40101007	VTAS.CHULETAS APANADAS EXPORT.	Movimiento	NaN
11	2020	1	1.0	6009	ECA	TC	40101008	VTAS. EXPORTACION TROPACK CAMARON	Movimiento	21504549.58
12	2020	1	1.0	6009	ECA	PA	40101009	VTAS. EXPORTACION PTA LA ALDEA	Movimiento	2174168.05
13	2020	1	5.0	603	VAG	VA	40101010	VENTA BOCADITO APANADO	Movimiento	NaN
14	2020	1	5.0	603	VAG	VA	40101011	VENTAS EXPORTACION PIEL PESC.	Movimiento	NaN
15	2020	1	1.0	6011	ECA	PA	40101012	VTAS. FLETES EXPORTACION PTA LA ALDEA	Movimiento	52346.01
16	2020	1	5.0	6009	VAG	VA	40101013	VTAS. EXP. FISH BLOCK LOMIT TR	Movimiento	NaN
17	2020	1	5.0	603	VAG	VA	40101014	ESCAMAS PESCADO	Movimiento	NaN
18	2020	1	0.0	6009	ECA	PT	40101015	VTAS. CAMARON EXPORTACION PTA - 2	Movimiento	NaN
19	2020	1	5.0	6011	ECA	PT	40101016	VTAS. FLETES EXPORTACION PTA 2	Movimiento	NaN

Nota: Almacena la información del estado de pérdidas y ganancias de manera mensual de los años 2022-2023

Tabla 8 Campos del Dataframe

Campo	Descripción
Período	Período contable
Mes	Mes para procesar
NIVCT	Nivel de la Cuenta Contable, en termino contable
CC	Centro de Costo
SC	Subcentro de Costo
FORM101	Código Formulario 101
CODCTA	Código de la Cuenta Contable
DESCT	Descripción de la Cuenta
TIT	Indica si la cuenta contable es título o movimiento
SACT	Saldo de la cuenta contable

Nota: Contenido del dataframe.

2.3. Granularidad

Con respecto al Mapeo realizado anteriormente de granular e identificar los campos necesarios para la depuración y procesamiento de datos y proceder con el cálculo de los indicadores solicitados. A continuación, se muestra los datos ya depurados.

Ilustración 23 Mapeo de Datos

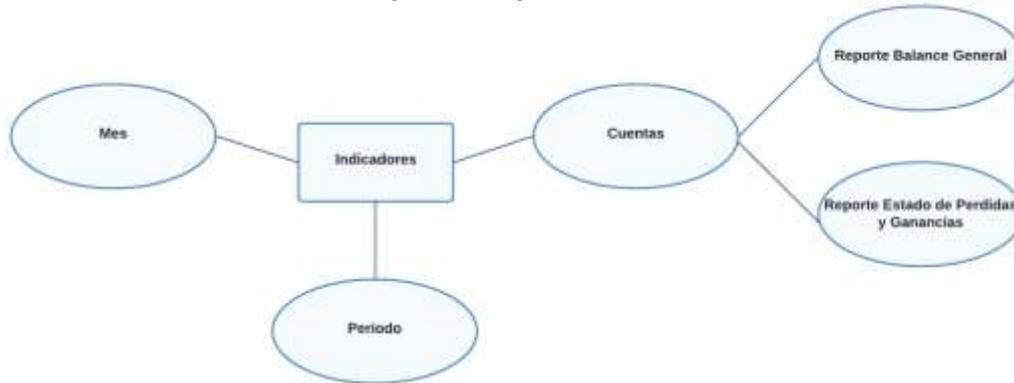
PERIODO	MES	ACTIVO CORRIENTE	PASIVO CORRIENTE	INVENTARIOS	PASIVO	ACTIVO	PATRIMONIO	PASIVOS NO CORRIENTE	UTILIDADES BRUTA	CUENTAS CORRIENTES	VENTAS	
0	2000	1	1.85340e+08	1.30535e+08	8.68254e+07	1.82279e+08	3.36800e+08	3.211862e+08	5.169134e+07	1.731344e+08	48878834.99	5.036427e+07
1	2000	2	2.101899e+08	1.392634e+08	8.526723e+07	2.047586e+08	5.304884e+08	3.211862e+08	6.549314e+07	1.731344e+08	71632402.65	1.064845e+08
2	2000	3	2.041490e+08	1.323885e+08	8.165450e+07	1.966623e+08	5.278540e+08	3.211862e+08	6.407362e+07	1.731344e+08	62841191.42	1.095007e+08
3	2000	4	2.019359e+08	1.346326e+08	8.150997e+07	1.878713e+08	5.256972e+08	3.211862e+08	6.333872e+07	1.731344e+08	60152395.66	2.197150e+08
4	2000	5	2.080549e+08	1.311190e+08	8.107610e+07	1.813895e+08	5.340775e+08	3.211862e+08	6.007046e+07	1.731344e+08	69134232.98	2.789561e+08
5	2000	6	2.037691e+08	1.252271e+08	8.651005e+07	1.848625e+08	5.291818e+08	3.211862e+08	5.963138e+07	1.731344e+08	59448150.35	3.198823e+08
6	2000	7	2.063694e+08	1.30625e+08	8.076859e+07	1.876672e+08	5.312721e+08	3.211862e+08	5.460486e+07	1.731344e+08	61076462.62	3.474495e+08
7	2000	8	2.078293e+08	1.360832e+08	8.626998e+07	1.892944e+08	5.298289e+08	3.211862e+08	5.264820e+07	1.731344e+08	67743966.27	3.830004e+08
8	2000	9	2.251020e+08	1.557080e+08	8.130065e+07	2.112673e+08	5.528502e+08	3.211862e+08	5.559225e+07	1.731344e+08	81038912.28	4.220187e+08
9	2000	10	2.077720e+08	1.370872e+08	8.204812e+07	1.916212e+08	5.368330e+08	3.211862e+08	5.453491e+07	1.686344e+08	71294550.40	4.791889e+08
10	2000	11	2.181670e+08	1.482798e+08	8.416079e+07	2.000234e+08	5.492721e+08	3.211862e+08	5.254148e+07	1.686344e+08	88819079.69	5.385072e+08
11	2000	12	2.186103e+08	1.610877e+08	8.215288e+07	2.177994e+08	5.677980e+08	3.267693e+08	5.070286e+07	1.784840e+08	63061794.02	5.787769e+08
12	2001	1	2.164613e+08	1.532857e+08	8.801361e+07	2.252785e+08	5.874082e+08	3.595052e+08	7.199285e+07	2.012199e+08	50567362.96	4.741750e+07
13	2001	2	2.358118e+08	1.676574e+08	1.025833e+08	2.463579e+08	6.188479e+08	3.595052e+08	7.879454e+07	2.012199e+08	74786574.94	1.030415e+08
14	2001	3	2.595000e+08	1.807144e+08	1.085028e+08	2.757494e+08	6.480751e+08	3.595052e+08	9.287502e+07	2.012199e+08	75308057.83	1.740146e+08
15	2001	4	2.641900e+08	1.752948e+08	1.116194e+08	2.684498e+08	6.505885e+08	3.595052e+08	9.245558e+07	2.012199e+08	80231366.78	2.471909e+08
16	2001	5	2.864074e+08	1.803775e+08	1.181533e+08	2.820937e+08	6.786832e+08	3.595052e+08	1.020161e+08	2.012199e+08	95682320.46	3.385867e+08
17	2001	6	3.129012e+08	1.799402e+08	1.378825e+08	3.017652e+08	7.354917e+08	3.595052e+08	1.218230e+08	2.012199e+08	9660495.25	4.256735e+08
18	2001	7	3.171877e+08	1.882144e+08	1.400570e+08	2.999970e+08	7.285241e+08	3.595052e+08	1.191828e+08	2.012199e+08	9718185.37	5.02534e+08
19	2001	8	3.162129e+08	1.876939e+08	1.359442e+08	3.014036e+08	7.493233e+08	3.595052e+08	1.177710e+08	2.012199e+08	98381711.34	6.00472e+08

Nota: Como se aprecia en la imagen solo tenemos los datos de las cuentas contables necesarias para los cálculos de los indicadores financieros solicitados.

2.4. Modelo Conceptual Ampliado

Una vez obtenido el modelo conceptual se procede al análisis del mismo para la normalización de la información de los reportes, como se detalla en el gráfico se condensa la información de ambos reportes en una sola estructura a la que vamos a llamar cuentas la cual almacenara las cuentas contables necesarias para el cálculo de los indicadores.

Ilustración 24 Modelo Conceptual Ampliado

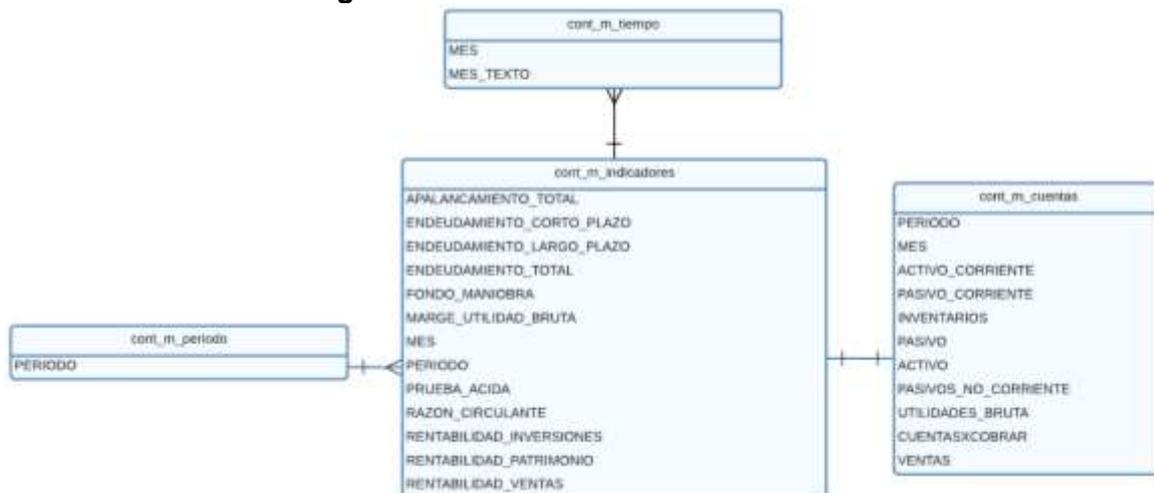


Nota: La estructura se condensa en cuentas normalizando ambos reportes para tener una sola fuente de datos

Fase 3: Modelo Lógico del Data Warehouse (DW)

Con ayuda del modelo lógico ampliado se procedió a la creación del DW que almacenara la información de los indicadores ya calculada lista para la consulta de los usuarios en cualquier momento. Así luce el modelo lógico del DW obtenido:

Ilustración 25 Modelo Lógico DW



Nota: En el modelo lógico ya podemos observar los campos procesados a partir de los datos de las cuentas contables obtenidas en la ejecución de los reportes.

Fase 4: Integración de Datos

4.1. Carga Inicial

Como se observa en las siguientes ilustraciones por medio de las técnicas ETL con el lenguaje Python se realiza la extracción de los datos gracias a las librerías de Python para el uso del motor de base MySQL como se observa en los Anexos 3 y 4 de los reportes contables conectándose a la base transaccional de la camaronera; el código indicado en las ilustraciones 26 y 27 detallan el procedimiento utilizado para extraer los balances generales

Ilustración 28 Script de Actualización

```
script_actualizacion.py > ...
1 from DBConnection import DBConnection
2 from datetime import datetime
3 meses = ['01','02','03','04','05','06','07','08','09','10','11','12']
4 resultados_balance_general = []
5 resultados_estado_pyg = []
6
7 periodo = datetime.now().year
8 tabla = 'a02300'
9
10 #Carga de Balance General
11 > for mes in meses: ...
63
64 #Carga de Estado de Perdidas y Ganancias
65 > for mes in meses: ...
113
114 import pandas as pd
115 df_balance_general = []
116 df_estado_pyg = []
117 > for resultado in resultados_balance_general: ...
123
124 > for resultado in resultados_estado_pyg: ...
130
131 df_balance_general.columns=['PERIODO','MES','NIVCT','CC','SC','FORM101','CODCTA','DESCT','TIT','SACT']
132 df_estado_pyg.columns=['PERIODO','MES','NIVCT','CC','SC','FORM101','CODCTA','DESCT','TIT','SACT']
133
134 df_balance_general['PERIODO']= df_balance_general['PERIODO'].astype("string")
135 df_balance_general['MES']= df_balance_general['MES'].astype("string")
136 df_balance_general["SACT"] = df_balance_general["SACT"].astype("string")
137 df_balance_general["SACT"] = df_balance_general["SACT"].astype(float)
138 df_balance_general['DESCT']= df_balance_general['DESCT'].str.strip()
139 df_balance_general['CODCTA']= df_balance_general['CODCTA'].str.strip()
140 df_estado_pyg['PERIODO']= df_estado_pyg['PERIODO'].astype("string")
141 df_estado_pyg['MES']= df_estado_pyg['MES'].astype("string")
142 df_estado_pyg["SACT"] = df_estado_pyg["SACT"].astype("string")
143 df_estado_pyg["SACT"] = df_estado_pyg["SACT"].astype(float)
144 df_estado_pyg['DESCT']= df_estado_pyg['DESCT'].str.strip()
145 df_estado_pyg['CODCTA']= df_estado_pyg['CODCTA'].str.strip()
146
147 datos_balance_general = df_balance_general.reindex(columns=["PERIODO","MES","CODCTA","DESCT","SACT"])
148 datos_estado_pyg = df_estado_pyg.reindex(columns=["PERIODO","MES","CODCTA","DESCT","SACT"])
149
150 > cuentas_balance_dict = { ...
161 > cuentas_estado_pyg_dict = { ...
164
165 data_cuentas = datos_balance_general[datos_balance_general['CODCTA']=='1'].reindex(columns=["PERIODO","MES"])
166 data_cuentas = data_cuentas.reset_index(drop=True)
167 > for codigo,nombre in cuentas_balance_dict.items(): ...
172 > for codigo,nombre in cuentas_estado_pyg_dict.items(): ...
177
178 lista = ['PERIODO','MES']+list(cuentas_balance_dict.values())+list(cuentas_estado_pyg_dict.values())
179 data_cuentas.columns=lista
180
181 data_indicadores = datos_balance_general[datos_balance_general['CODCTA']=='1'].reindex(columns=["PERIODO","MES"])
182 data_indicadores = data_indicadores.reset_index(drop=True)
183 #Razón Circulante
184 razon_circulante = data_cuentas['ACTIVO_CORRIENTE']/data_cuentas['PASIVO_CORRIENTE']
185 data_indicadores = pd.concat([data_indicadores, razon_circulante], axis=1)
186
187 #Prueba Acida
188 prueba_acida = (data_cuentas['ACTIVO_CORRIENTE'] - data_cuentas['INVENTARIOS])/data_cuentas['PASIVO_CORRIENTE']
189 data_indicadores = pd.concat([data_indicadores, prueba_acida], axis=1)
```

Nota: Script que al ejecutarse agrega la información actualizada de indicadores del año en curso.

Ilustración 31 Pantalla de Menú de Dashboard



Nota: Se muestra todas las opciones que se brinda en el dashboard

Ilustración 32 Pantalla de Indicadores de Liquidez



Nota: Se muestra los indicadores de liquidez los cuales de los puede filtrar por periodo y mes.

Ilustración 33 Pantalla de Indicadores de Endeudamiento



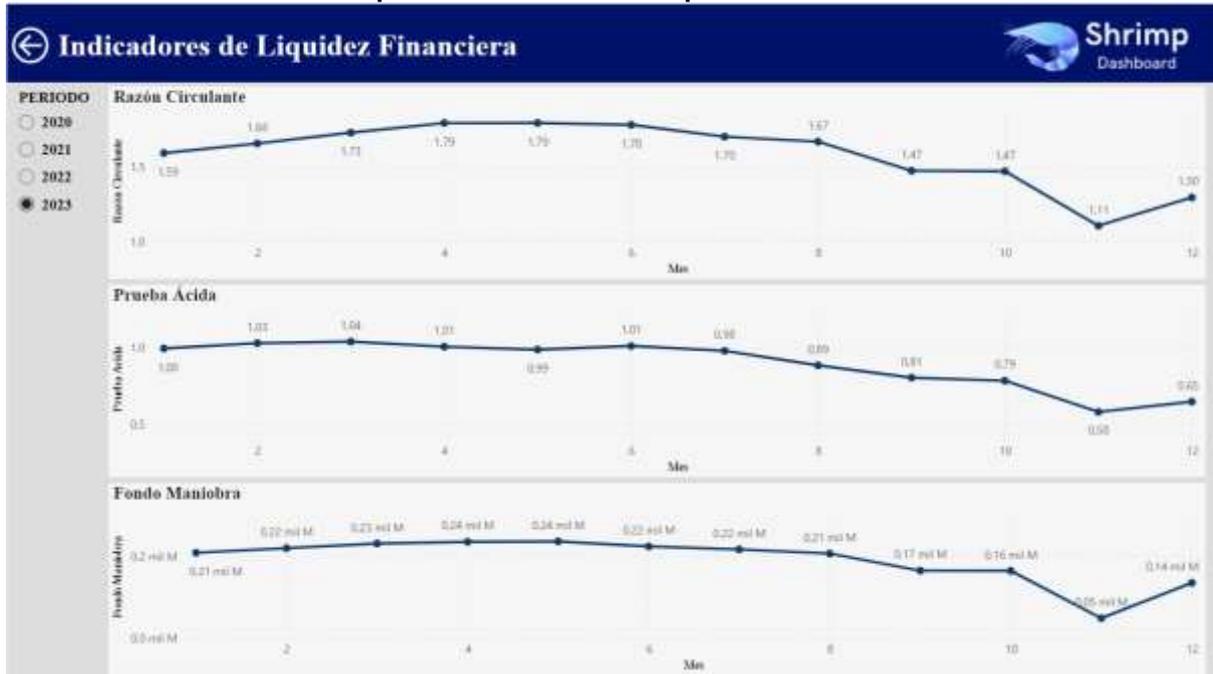
Nota: Se muestra los indicadores de endeudamiento los cuales de los puede filtrar por periodo y mes.

Ilustración 34 Pantalla de Indicadores de Rentabilidad



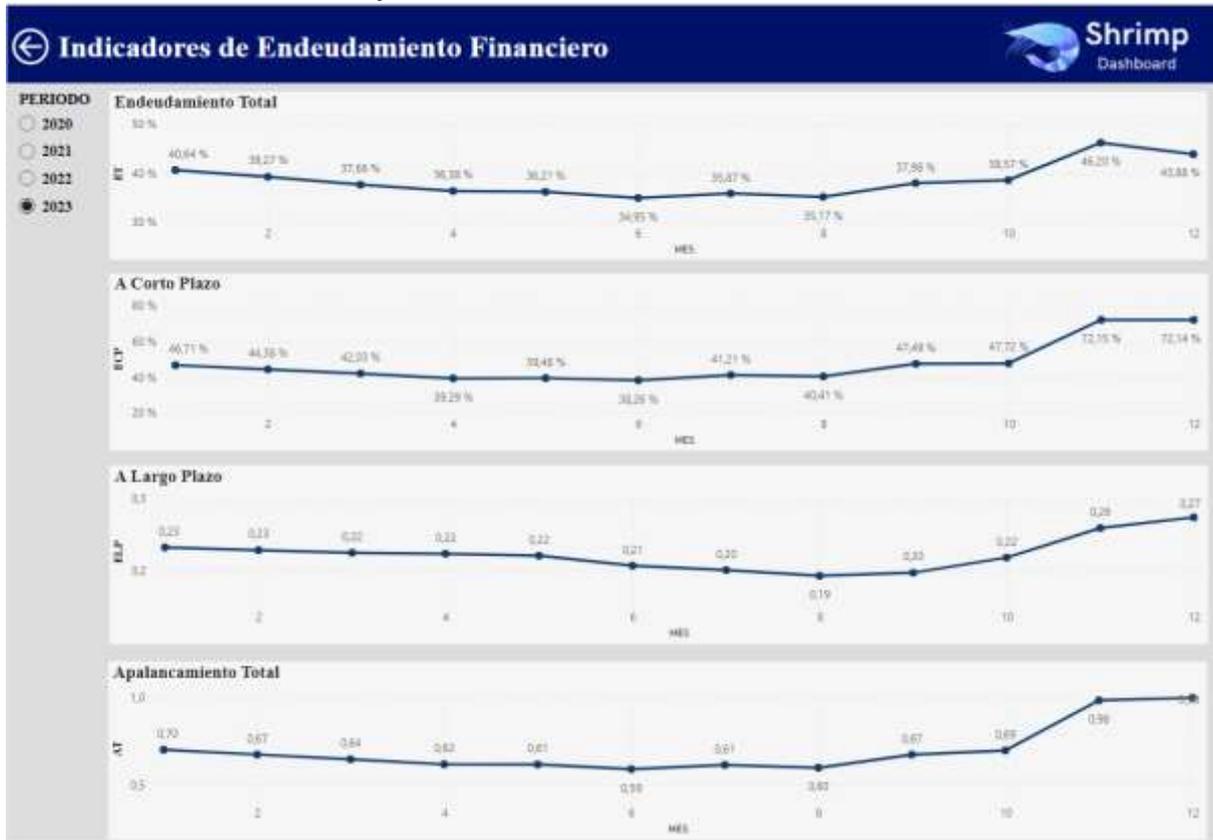
Nota: Se muestra los indicadores de rentabilidad los cuales de los puede filtrar por periodo y mes.

Ilustración 35 Series de Tiempo de Indicadores de Liquidez



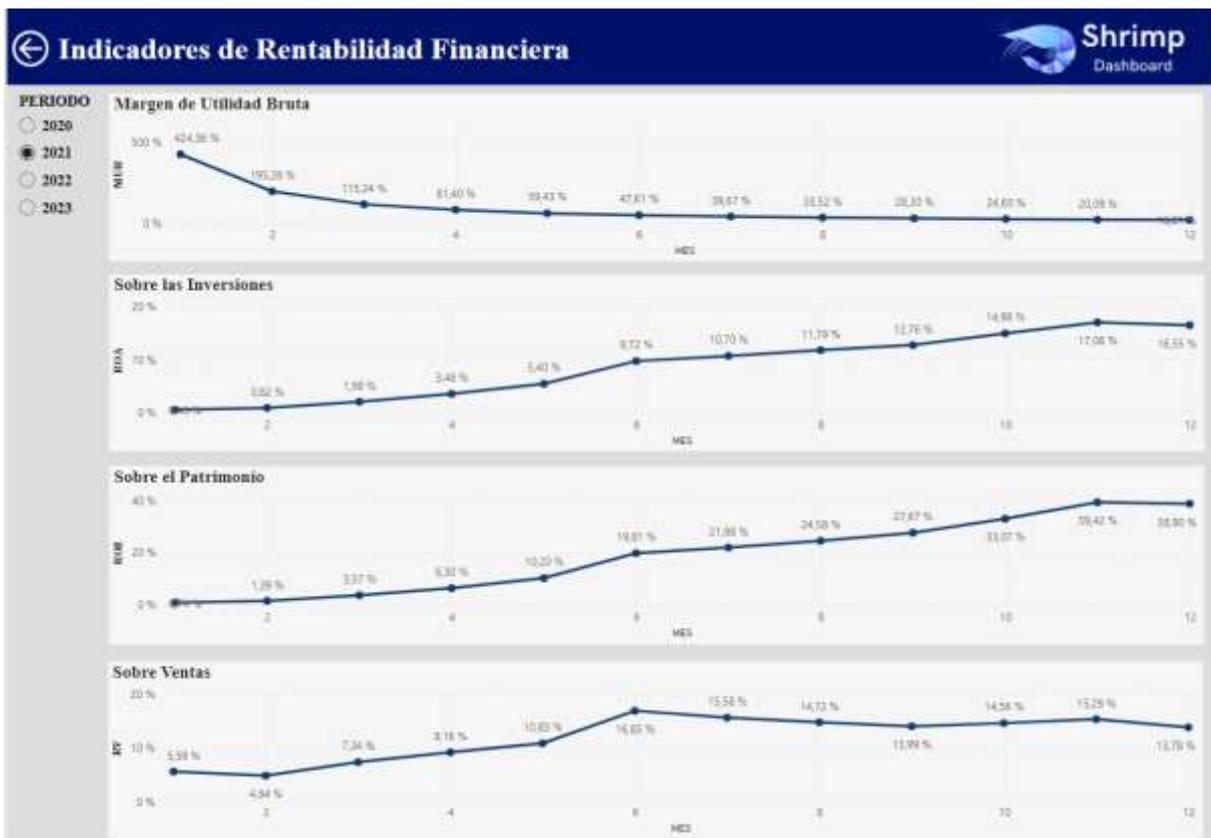
Nota: Se muestra las series de tiempo de los indicadores de liquidez los cuales se puede visualizar por periodo

Ilustración 36 Series de Tiempo de Indicadores de Endeudamiento



Nota: Se muestra las series de tiempo de los indicadores de endeudamiento los cuales se puede visualizar por periodo

Ilustración 37 Series de Tiempo de Indicadores de Rentabilidad



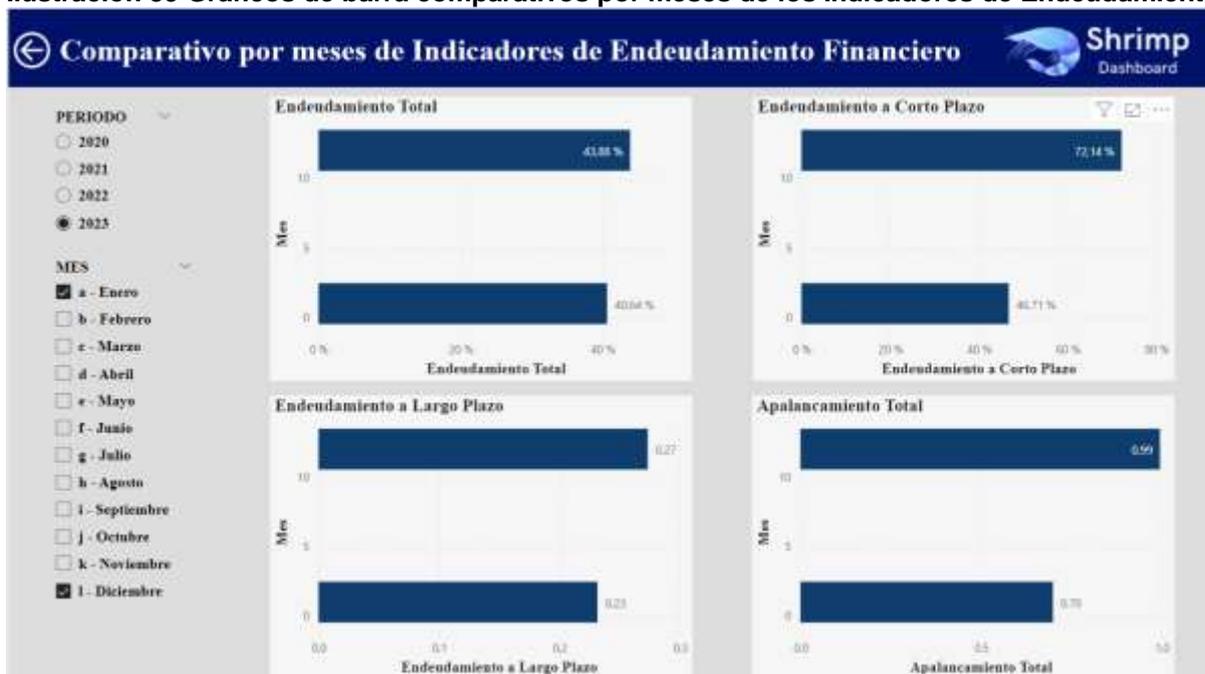
Nota: Se muestra las series de tiempo de los indicadores de rentabilidad los cuales se puede visualizar por periodo

Ilustración 38 Gráficos de barra comparativos por meses de los Indicadores de Liquidez



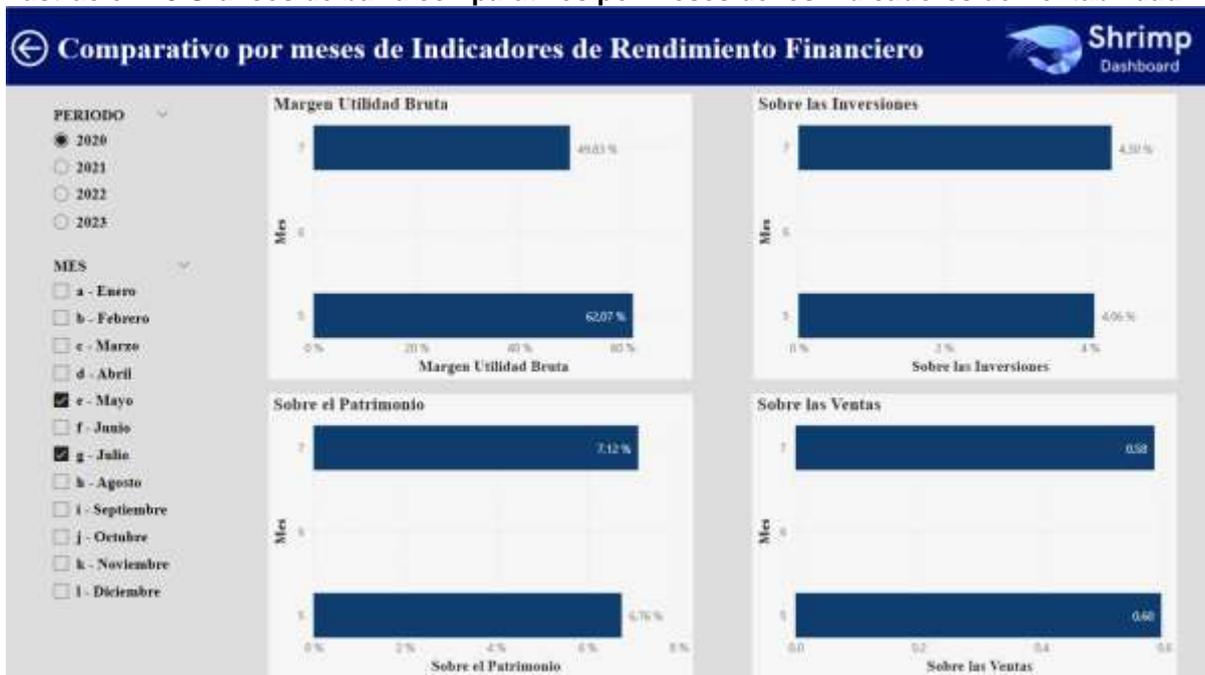
Nota: Se muestra gráficos que comparan los cambios de los indicadores de liquidez de manera mensual.

Ilustración 39 Gráficos de barra comparativos por meses de los Indicadores de Endeudamiento



Nota: Se muestra gráficos que comparan los cambios de los indicadores de endeudamiento de manera mensual

Ilustración 40 Gráficos de barra comparativos por meses de los Indicadores de Rentabilidad



Nota: Se muestra gráficos que comparan los cambios de los indicadores de rentabilidad de manera mensual

Conclusiones

- Con la implementación del prototipo el proceso previo a la toma de decisiones gerenciales paso de durar 14.75 horas a tener la información de manera inmediata desde el Power BI, logrando la eliminación de los reprocesos de la información contable requerida y la gestión manual por parte del personal contable para el cálculo de indicadores; así también, se logra tener consistencia en la información debido al filtrado de los datos disminuyendo la tasa de errores. Esta solución brinda información efectiva y oportuna, por lo que se da cumplimiento a los objetivos generales y específicos, y al alcance definido en este Proyecto.
- Para este proyecto se planteó la siguiente interrogante investigativa: ¿Permitirá el prototipo de Business Intelligence proporcionar información confiable y de manera oportuna para el proceso de toma de decisiones estratégicas de la alta gerencia?; la aplicación de los procesos ETL en el desarrollo del proyecto permite que la información obtenida pase por un proceso de filtrado y depuración lo que disminuye la tasa de errores y la información cumple con los criterios de integridad y confiabilidad, y mediante la implementación de la herramienta Power BI se tiene la información en el momento requerido es decir con una alta disponibilidad; de esta forma damos cumplimiento a la interrogante planteada.
- Se corrobora que el uso de la metodología Hefesto es de gran utilidad cuando nos queremos enfocar en un área específica de la organización y a su vez en la implementación de un Data Warehouse.
- El uso de Python y su librería pandas facilitaron el proceso ETL aplicado en los datos, gracias a esto se logró la normalización correcta y la realización de los cálculos de los indicadores.
- Power BI es una herramienta muy intuitiva, al realizar la conexión con la base DW automáticamente identificó las relaciones para el cubo de información, de esta manera los filtros son fáciles de aplicar en el Dashboard.

Recomendaciones

- Elaborar un plan de actualización y mantenimiento, con respecto a la infraestructura para el correcto funcionamiento del almacén de datos y del prototipo realizado, debido al crecimiento de la base de datos.
- Incrementar al proyecto nuevos Datamarts con otros indicadores para el control de diferentes departamentos de la empresa.
- Designar personal responsable de la validación y operación del sistema y la carga de datos.
- Agregar segmentaciones de los indicadores por Centros de Costo o por establecimientos, al menos si se agregan indicadores de productividad.

Referencias

- Arias, A. (2015). *Bases de Datos con MySQL: 2ª Edición (Spanish Edition)* (2ª Edición). CreateSpace Independent Publishing Platform; 2nd edición (23 Julio 2015).
- Bernabeu, R. D., & García Mattio, M. (2018). *Hefesto—Data Warehousing v3*. Propia.
- Coronel, C., & Morris, S. (2018). *Database Systems: Design, Implementation, & Management* (13a ed.). Cengage Learning.
- Deckler, G., Powell, B., & Gordon, L. (2022). *Mastering Microsoft Power BI: Expert techniques to create interactive insights for effective data analytics and business intelligence* (2nd Edición). Packt Publishing.
- Ellison, B. (2021). *Python for Data Analysis: A Beginners Guide to Master the Fundamentals of Data Science and Data Analysis by Using Pandas, Numpy and Ipython (Programming Languages for Beginners)* (1era Edición). Brady Ellison.
- Hueso Ibáñez, L. (2016). *Administración de Sistemas gestores de Bases de Datos* (2da Edición). Ediciones de la U.
- Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling* (3rd Edición). Wiley.
- Kishore Pandey, B., & Ro Schoof, E. (2023). *Building ETL Pipelines with Python: Create and deploy enterprise-ready ETL pipelines by employing modern methods* (Primera Edición). PACKT.
- Parmenter, D. (2015). *Key Performance Indicators: Developing, Implementing, and Using Winning KPIs* (3rd Edición). Wiley.
- Pérez Marqués, Ma. (2015). *Business Intelligence Técnicas Herramientas Y Aplicaciones* (1.ª ed.). ALFAOMEGA RC LIBROS.
- S. Kaplan, R., & P. Norton, D. (2016). *El cuadro de mando integral: The balanced scorecard* (3er Edición). Gestión 2000.
- Toro Bonilla, M. (2022). *Fundamentos de programación: PYTHON* (2.ª Edición). Universidad de Sevilla.
- Wrembel, R., & Koncilia, C. (2011). *Data Warehouses and OLAP: Concepts, Architectures and Solutions*. IRM Press.

Anexos

Anexo 1 Informe Reunión de Requerimientos con Alta Gerencia

 DEPARTAMENTO SISTEMAS Y PROYECTOS TI	
REQUERIMIENTO DE DESARROLLO	
FECHA:	6/11/2023
EMPRESA:	IPSP
UNIDAD DE NEGOCIO:	IPSP Empacadora
ÁREA:	Sistemas y Proyectos TI
RESPONSABLE:	Enrique Gilbert
CONTACTO:	
TIPO DE REQUERIMIENTO:	Proyecto
<i>Proyecto tecnológico es un plan que se ha definido para crear un producto o modificarlo atendiendo a las necesidades de los usuarios y siempre con el objetivo de mejorar la calidad de vida de los mismos.</i>	
NÚMERO DE REQUERIMIENTO:	0001
CÓDIGO DE REQUERIMIENTO:	IPSP-SIST-PROY-0001
TÍTULO DE REQUERIMIENTO (PROYECTO O CAMBIO):	
Reunion con el personal encargado del proceso manual de obtencion de indicadores financieros para la toma de decisiones gerenciales	
PROBLEMÁTICA: CUÁL ES EL PROBLEMA O NECESIDAD?	
El problema se refleja en el proceso de toma no oportuna de decisiones estratégicas, la generación de reportes, la intervención de personal del área contable para general reportes gerenciales, el tiempo perdido en este proceso, para obtener información no tan exacta para las estrategias gerenciales.	
OBJETIVOS: SMART	
Mejorar el proceso de toma de decisiones de la alta gerencia mediante el uso de herramientas de TI.	
RESULTADOS ESPERADOS: QUÉ SE DESEA OBTENER AL SOLUCIONAR EL PROBLEMA?	
La implementación del prototipo de Business Intelligence a través de un dashboard, permitirá a la alta gerencia tener una mejor perspectiva sobre los indicadores financieros para la toma de decisiones estratégicas.	
PARTICIPANTES	
Priscila Salem	
Santiago Salem	
Julio Salem	
Cristina Salem	
Jackeline Choez	
Gerente General : Roberto Coronel	
Gerente Financiero: Jackeline Cabezas	
Gerente de Sistemas: Enrique Gilbert	
Delegado Sistemas: Eileen Sanchez	
Delegado Contable 1: Maryuri Macias	
Delegado Contable 2: Cecibel Garcia	
Delegado Contable 3: Luis Leon	
RESOLUCIÓN	
Se determino las necesidades y requerimientos para poder definir que informacion necesitan obtener en tiempo real una vez realizado el desarrollo. Se identifico que reportes usan para el calculo manual de indicadores, los actores del proceso y reproceso en este caso son 3 personas del area contable. Se planteo las siguientes interrogantes a solucionar con el desarrollo del proyecto las cuales se detallan a continuacion. Adicional a esto se estableció un periodo de entrega del proyecto a analizar y con entregables para comprobar el avance del mismo.	
Preguntas a solucionar	
1.¿Desea saber acerca de la Liquidez de la empresa?	
2.¿Desea reportes de los indicadores de Liquidez de manera histórica?	
3.¿Desea saber acerca del Endeudamiento de la empresa?	
4.¿Desea reportes de los indicadores de Endeudamiento de manera histórica?	
5.¿Desea saber acerca de la Rentabilidad de la empresa?	
6.¿Desea reportes de los indicadores de Rentabilidad de manera histórica?	

Anexo 2 Reportes Contables de Diciembre

			
Estado pyg 2021.xlsx	Balance General 2022.xlsx	Balance General 2021.xlsx	Estado pyg 2022.xlsx

Anexo 3 Script de Conexión a la Base de Datos



DBConnection.py



DBConnector.py

Anexo 4 Script de Carga Histórica



script_data_historica.
py

Anexo 5 Script de Actualización



script_actualizacion.p
y

Anexo 6 Proyecto BI



INDICADORES
CAMARONERA.pbix

Glosario

- BI:** Es un conjunto de metodologías, aplicaciones, estrategias, prácticas y capacidades que se enfocan en crear y administrar la información para su utilidad en el sector financiero de una organización, permitiendo tomar mejores decisiones a los usuarios.
- OLAP:** Online Analytical Processing son las bases de datos OLAP están diseñadas para facilitar el análisis y la consulta eficiente de grandes conjuntos de datos.
- OTLP:** Online Transaction Processing son las bases de datos OLTP están diseñadas para procesar transacciones en tiempo real, lo que significa que son eficientes para insertar, actualizar y eliminar registros en una base de datos.
- DBMS:** Un Gestor de Bases de Datos es un software que facilita la creación, el mantenimiento y el uso de bases de datos.
- ETL:** Extract, Transform, Load, ETL es un proceso clave en el desarrollo de un Data Warehouse, donde se extraen datos de diversas fuentes, se transforman según las necesidades y se cargan en el almacén de datos para su análisis.
- KPI:** Son métricas cuantitativas utilizadas para evaluar el rendimiento de una organización, departamento o proceso; representan medidas críticas alineadas con los objetivos estratégicos o metas de la empresa. Sirven para la medición de rendimiento proporcionando una medida clara y cuantificable y para la toma de decisiones basada en datos y en el seguimiento del progreso hacia metas estratégicas.



DECLARACIÓN Y AUTORIZACIÓN

Yo, **Sánchez Naranjo Eileen Belén**, con C.C: # **0950467290** autora del trabajo de titulación: **Desarrollo de prototipo de sistema de Business Intelligence para la toma de decisiones estratégicas en una camaronera localizada en la provincia del Guayas**, previo a la obtención del título de **Ingeniero en Ciencias de la Computación** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, 8 de febrero del 2024

Nombre: **Sánchez Naranjo Eileen Belén**
C.C: **0950467290**

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TEMA Y SUBTEMA:	Desarrollo de prototipo de sistema de Business Intelligence para la toma de decisiones estratégicas en una camaronera localizada en la provincia del Guayas.		
AUTOR(ES)	Eileen Belen Sanchez Naranjo		
REVISOR(ES)/TUTOR(ES)	Byron Severo Yong Yong		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Ingeniería		
CARRERA:	Ingeniería en Ciencias de la Computación		
TÍTULO OBTENIDO:	Ingeniero en Ciencias de la Computación		
FECHA DE PUBLICACIÓN:	08 de febrero del 2024	No. DE PÁGINAS:	49
ÁREAS TEMÁTICAS:	Área Financiera de una Camaronera		
PALABRAS CLAVES/ KEYWORDS:	Inteligencia de Negocio, Power BI, MySql, Indicadores Financieros, Hefesto.		
RESUMEN:	<p>El presente trabajo de titulación tiene como objetivo la implementación de un prototipo de Inteligencia de Negocios para ayudar en la toma de decisiones de la alta gerencia de una camaronera localizada en la ciudad de Guayaquil, esto simplificaría la obtención de información efectiva y fundamentada en datos históricos, necesarios para el proceso. Para ellos se hará uso de tecnología que permita realizar el análisis, diseño, consolidación y extracción de los datos transaccionales de la organización. Los reportes desactualizados y los reprocesos solicitados hacen que la toma de decisiones no se realice de manera efectiva, la aplicación de BI permite obtener indicadores financieros y de gestión de manera fácil y eficiente en cualquier momento sin necesidad de un proceso manual, para la construcción de la solución informática se hará uso de: la metodología de Hefesto, la herramienta Power BI y la base de datos MySQL permitiendo la creación de un propio Data Warehouse que permitirá la consulta de información efectiva en cualquier momento.</p>		
ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	
CONTACTO CON AUTOR/ES:	Teléfono:+593-939845904	E-mail: eileen_sanchez@hotmail.es	
CONTACTO CON LA INSTITUCIÓN(COORDINADOR DEL PROCESO UTE):	Toala Quimí, Edison José		
	Teléfono: +593-990-976776		
	E-mail: edison.toala@cu.ucsg.edu.ec		
SECCIÓN PARA USO DE BIBLIOTECA			
Nº. DE REGISTRO (en base a datos):			
Nº. DE CLASIFICACIÓN:			
DIRECCIÓN URL (tesis en la web):			