



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ingeniería de Sistemas e Informática

Escuela Profesional de Ingeniería de Sistemas

**Una solución de inteligencia de negocios para apoyar el
análisis del desempeño de los proyectos en una
empresa constructora**

TESIS

Para optar el Título Profesional de Ingeniero de Sistemas

AUTOR

Grace Stefanny MORANTE PALOMINO

ASESOR

Mg. César LUZA MONTERO

Lima, Perú

2021



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Morante, G. (2021). *Una solución de inteligencia de negocios para apoyar el análisis del desempeño de los proyectos en una empresa constructora*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Escuela Profesional de Ingeniería de Sistemas]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Grace Stefanny Morante Palomino
Tipo de documento de identidad	DNI
Número de documento de identidad	72452711
URL de ORCID	-
Datos de asesor	
Nombres y apellidos	Cesar Luza Montero
Tipo de documento de identidad	DNI
Número de documento de identidad	06111988
URL de ORCID	https://orcid.org/0000-0003-3985-0232
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Luis Soto Soto
Tipo de documento	DNI
Número de documento de identidad	22409532
Miembro del jurado 1	
Nombres y apellidos	Jorge DiazMuñante
Tipo de documento	DNI
Número de documento de identidad	07216161
Datos de investigación	
Línea de investigación	C.0.3.4 Inteligencia de Negocios
Grupo de investigación	No aplica

Agencia de financiamiento	Sin Financiamiento
Ubicación geográfica de la investigación	Edificio: Universidad Nacional Mayor de San Marcos País: Perú Departamento: Lima Provincia: Lima Distrito: Cercado de Lima Calle: Av. Amezaga 375 Latitud: -12,056 Longitud: -77,0844
Año o rango de años en que se realizó la investigación	2019-2021
URL de disciplinas OCDE	Ingeniería de Sistemas y Comunicaciones https://purl.org/pe-repo/ocde/ford#2.02.04

DEDICATORIA

Todo este esfuerzo va dedicado a mis padres y a mi familia, quienes son mi principal motor y motivación. A mi ángel, mi mamita bella Olga, que desde el cielo me guía en cada paso dado.

AGRADECIMIENTOS

A Dios, por guiar mis pasos tanto en mi vida profesional como personal, por darme la fuerza de ser constante y perseguir mis sueños y metas. Agradezco a mis amigos de la base 13 por hacer de mi experiencia universitaria un momento especial y de muchos gratos recuerdos. Por la oportunidad de pertenecer a esta gran casa de estudios, mi querida Decana de América, mi universidad Nacional Mayor de San Marcos.

A mi familia por su constante apoyo incondicional y motivación a siempre seguir adelante en todo momento, con responsabilidad, respeto y perseverancia.

A mi asesor de tesis, el Magíster César Luza Montero, quien me apoyó y me motivó en cada momento durante el desarrollo y documentación de este proyecto de tesis. Agradezco el tiempo dedicado a nuestras reuniones de avances, a la predisposición de asesoría que obtuve de su parte.

A cada uno de mis profesores durante mi etapa universitaria, quienes ayudaron y guiaron mi crecimiento y desarrollo profesional.

A Edson por ser mi mentor en mis pininos, agradecerle por despertarme el interés en el mundo de la analítica de datos.

A todos, muchas gracias.

RESUMEN

El propósito de esta investigación fue implementar una solución de Inteligencia de Negocios para apoyar la gestión de los indicadores de desempeño, en el Área de Control de Proyectos, de una empresa constructora del Perú que monitoreaba solo indicadores de costos, sin considerar otros indicadores de desempeño importantes como tiempo (p.e índice de avance de lo ejecutado respecto a lo planeado) y seguridad y salud (p.e índice de accidentabilidad). La información relacionada a estos indicadores se encontraba en fuentes de información dispersas y el procesamiento era realizado de manera tradicional apoyados en hojas de cálculos, en consecuencia el tiempo de procesamiento para la generación de la información excedía lo esperado generando dificultad para la toma decisiones de manera oportuna y eficiente.

Se utilizó la metodología de Ralph Kimball para la construcción del modelo Datawarehouse y la herramienta Tableau para la explotación de data, considerándose indicadores de Costo, Tiempo, y Seguridad y Salud. La información obtenida se almacenó en el Datawarehouse compuesto por 3 modelos datamarts. Para el entregable final se creó 3 tableros: Resultado Operativo, Avance de Proyectos y Accidentabilidad.

Para la validación de la propuesta la empresa constructora proporcionó información relacionada con los indicadores Costo, Tiempo, y Seguridad y Salud, tanto de los proyectos que tienen adjudicados como los ejecutados. Se compararon los tiempos de procesamiento antes de la solución, que consideraba solo indicadores de costos, con el tiempo de procesamiento usando la solución propuesta, que considera además de los indicadores de costos, los indicadores de tiempo, y seguridad y salud. Los resultados obtenidos con la solución de inteligencia de negocios evidencian una reducción significativa del tiempo del procesamiento debido a la centralización de información y automatizaciones de cálculos de estos indicadores.

Palabras Claves: *Inteligencia de negocios, Kimball, Indicadores, Análisis de desempeño.*

ABSTRACT

The purpose of this research was to implement a Business Intelligence solution to support the management of performance indicators, in the Project Control Area, of a construction company in Peru that only monitored cost indicators, without considering other important performance indicators such as time (rate of progress of what was executed compared to what was planned) and health and safety (rate of accident rate). The information related to these indicators was found in dispersed information sources and the processing was carried out in a traditional way supported by spreadsheets, consequently the processing time for the generation of the information exceeded what was expected, generating difficulty in making decisions timely and efficient.

Ralph Kimball's methodology was used for the construction of the Datawarehouse model and the Tableau tool for data exploitation, considering Cost, Time, and Health and Safety indicators. The information obtained was stored in the Datawarehouse composed of 3 datamarts models. For the final deliverable, 3 dashboards were created: Operating Result, Project Progress and Accident Rate.

For the validation of the proposal, the construction company provided information related to the indicators Cost, Time, and Safety and Health, both of the projects that have been awarded and those that have been executed. The processing times before the solution, which considered only cost indicators, were compared with the processing time using the proposed solution, which considers in addition to the cost indicators, the time indicators, and health and safety. The results obtained with the business intelligence solution show a significant reduction in processing time.

Keywords: *Business Intelligence, Kimball, KPI, Performance Analysis.*

INDICE

AGRADECIMIENTOS.....	III
RESUMEN	IV
ABSTRACT	V
LISTA DE FIGURAS.....	VIII
LISTA DE TABLAS.....	X
INTRODUCCIÓN	1
CAPÍTULO 1: PLANTEAMIENTO DEL PROBLEMA	3
Situación problemática.....	3
Definición del Problema	5
Objetivos	5
Objetivo General	5
Objetivos Específicos	5
Justificación	6
CAPÍTULO 2: MARCO TEÓRICO.....	9
El análisis de desempeño en los proyectos de construcción.....	9
Inteligencia de Negocios	11
Información y afines	11
Sistemas de Información	15
Gestión de información y conocimiento	17
Definición de Inteligencia de Negocios	18
Proceso de Business Intelligence	21
Componentes de una Solución BI.....	25
Marco Conceptual	32
CAPÍTULO 3: REVISIÓN DE LA LITERATURA.....	34
Artículos Revisados	34
Resumen de Revisión de la Literatura	40
Indicadores de proyectos de Construcción	42
Herramientas de Inteligencia de Negocios	45
Metodologías de Inteligencia de Negocios	52
Benchmarking de Metodologías de Inteligencia de Negocios.....	56

CAPÍTULO 4: APLICACIÓN DE LA METODOLOGÍA.....	60
Descripción de la metodología.....	60
Fase Planificación.....	60
Fase de Requerimientos.....	61
Fase de Análisis.....	61
Fase de Diseño.....	63
Fase de Construcción.....	64
Fase de Despliegue.....	65
Aplicación de las fases I y II de la Metodología.....	66
Aplicación de Fase I: Planificación.....	66
Aplicación de Fase II: Requerimientos del Proyecto.....	69
CAPÍTULO 5: SOLUCIÓN TECNOLÓGICA	71
Aplicación de Fase III: Análisis.....	71
Aplicación de Fase IV: Diseño.....	83
Aplicación de Fase V: Construcción.....	83
Aplicación de Fase VI: Despliegue	101
CAPÍTULO 6: ANÁLISIS DE RESULTADOS.....	105
Resultado de los indicadores por categoría	105
Categoría Costo	105
Categoría Tiempo	106
Categoría Seguridad y Salud	107
Análisis comparativo	107
CONCLUSIONES, RECOMENDACIONES Y TRABAJOS FUTUROS.....	110
Conclusiones.....	110
Recomendaciones.....	111
Trabajos Futuros.....	111

LISTA DE FIGURAS

<i>Figura 1 Producto Bruto Interno por Actividad Económica 2021</i>	<i>7</i>
<i>Figura 2 Pirámide del conocimiento</i>	<i>12</i>
<i>Figura 3 Proceso de Gestión de Información.....</i>	<i>17</i>
<i>Figura 4 Cadena de Valor de la Inteligencia</i>	<i>18</i>
<i>Figura 5 Pirámide de Niveles Organizacionales BI.....</i>	<i>20</i>
<i>Figura 6 Fases del Proceso BI</i>	<i>21</i>
<i>Figura 7 Esquema típico del proceso ETL</i>	<i>22</i>
<i>Figura 8 Datos importantes en el Proceso de Extracción.....</i>	<i>23</i>
<i>Figura 9 Actividades en la Transformación de Datos.....</i>	<i>24</i>
<i>Figura 10 Características de un Data Warehouse</i>	<i>26</i>
<i>Figura 11 Ejemplo de Cubo OLAP.....</i>	<i>28</i>
<i>Figura 12 OLAP según su arquitectura de almacenamiento</i>	<i>30</i>
<i>Figura 13 Ejemplo de Modelo Estrella.....</i>	<i>31</i>
<i>Figura 14 Cuadrante Mágico de Gartner para Analytics y Plataformas de Inteligencia de Negocios.....</i>	<i>46</i>
<i>Figura 15 Tablero SINADEF- MINSA construido en Tableau.....</i>	<i>48</i>
<i>Figura 16 Flujo de trabajo de la Plataforma Power BI.....</i>	<i>50</i>
<i>Figura 17 Arquitectura Datawarehouse propuesta por Bill Inmon</i>	<i>53</i>
<i>Figura 18 Arquitectura Data Warehouse propuesta por Ralph Kimball</i>	<i>55</i>
<i>Figura 19 Tareas de la Metodología de Ralph Kimball.....</i>	<i>61</i>
<i>Figura 20 Gantt del Proyecto de Inteligencia de Negocios.....</i>	<i>68</i>
<i>Figura 21 Arquitectura de la Solución propuesta</i>	<i>71</i>
<i>Figura 22 Modelo Constructora CP Avance</i>	<i>74</i>
<i>Figura 23 Fact Table del modelo CP Avance.....</i>	<i>78</i>
<i>Figura 24 Modelo CP Margen</i>	<i>80</i>
<i>Figura 25 Modelo CP Requerimiento Equipo</i>	<i>81</i>
<i>Figura 26 Modelo SSOMA - Accidentabilidad.....</i>	<i>82</i>
<i>Figura 27 Esquema de Transformación de Datos</i>	<i>84</i>
<i>Figura 28 Carga Staging Area – Lista de Procesos ETL.....</i>	<i>85</i>

<i>Figura 29 Esquema de Flujo de Control de Fuentes PeopleSoft</i>	<i>86</i>
<i>Figura 30 Esquema de Flujo de Datos de Fuentes PeopleSoft</i>	<i>87</i>
<i>Figura 31 Flujo de Datos de la Carga CP Avance</i>	<i>89</i>
<i>Figura 32 Conversor de Datos de XLS_CP_Avance</i>	<i>89</i>
<i>Figura 33 Flujo de Datos PS_CO_CR_VW y PS_CO_PROJECT</i>	<i>90</i>
<i>Figura 34 Carga Datawarehousing - lista de procedimientos almacenados.....</i>	<i>91</i>
<i>Figura 35 Procedimiento de carga a fuente BI_F_Control_Proyecto_Avance</i>	<i>92</i>
<i>Figura 36 Tableau – Conexión de Orígenes de Datos</i>	<i>94</i>
<i>Figura 37 Tableau – Esquema de Trabajo Data Source</i>	<i>95</i>
<i>Figura 38 Tableau- Creación de nueva hoja de trabajo Worksheet</i>	<i>95</i>
<i>Figura 39 Tableau – Hoja de trabajo Avance Proyecto.....</i>	<i>96</i>
<i>Figura 40 Tableau – Creación de indicador Avance Cumplimiento.....</i>	<i>97</i>
<i>Figura 41 Tableau – Campos en filas y columnas</i>	<i>97</i>
<i>Figura 42 Tableau – tipos de visualizaciones disponibles</i>	<i>98</i>
<i>Figura 43 Ventas por Proyecto - Gráfico de barras evolutivo por tipo de venta.....</i>	<i>99</i>
<i>Figura 44 Tablero Avance Proyectos</i>	<i>100</i>
<i>Figura 45 Tablero de Accidentabilidad en Proyectos.....</i>	<i>101</i>
<i>Figura 46 Plan de Pruebas - Carga de datos al Stage mediante ETL</i>	<i>103</i>
<i>Figura 47 Plan de Pruebas - Carga de datos al DW</i>	<i>104</i>
<i>Figura 48 Resultados del índice de margen de ganancia.....</i>	<i>105</i>
<i>Figura 49 Resultados del índice de avance de proyecto</i>	<i>106</i>

LISTA DE TABLAS

<i>Tabla 1 Paradigma de la información.....</i>	<i>13</i>
<i>Tabla 2 Resumen de Revisión de Literatura</i>	<i>41</i>
<i>Tabla 3 Categorías e indicadores de rendimiento en Proyectos de Construcción.....</i>	<i>45</i>
<i>Tabla 4 Cuadro comparativo de herramientas de visualización BI.....</i>	<i>52</i>
<i>Tabla 5 Comparación de los enfoques acorde a la metodología y la arquitectura.....</i>	<i>56</i>
<i>Tabla 6 Ponderación de factores - Benchmarking de metodologías</i>	<i>57</i>
<i>Tabla 7 Matriz de puntuación por afinidad de criterio- Benchmarking Metodologías.....</i>	<i>58</i>
<i>Tabla 8 Matriz de criterios – Benchmarking de Metodologías</i>	<i>58</i>
<i>Tabla 9 Benchmarking de metodologías de inteligencia de negocios</i>	<i>59</i>
<i>Tabla 10 Tabla de indicadores a monitorear en Gerencia de Control de Proyectos.....</i>	<i>70</i>
<i>Tabla 11 Dimensiones relacionadas a la Perspectiva Gerencia</i>	<i>75</i>
<i>Tabla 12 Dimensiones relacionadas a la Perspectiva Centro de Responsabilidad</i>	<i>76</i>
<i>Tabla 13 Dimensiones relacionadas a la perspectiva Cliente y Tiempo</i>	<i>77</i>
<i>Tabla 14 Fuentes en construcción de CP Avance Staggering Area</i>	<i>88</i>
<i>Tabla 15 Listado de tableros contruidos en la herramienta Tableau.....</i>	<i>93</i>
<i>Tabla 16 Plan de Pruebas - Criterios de aceptación.....</i>	<i>102</i>
<i>Tabla 17 Plan de Pruebas – Requerimiento Equipo.....</i>	<i>102</i>
<i>Tabla 18 Resultados del Índice ATP.....</i>	<i>107</i>
<i>Tabla 19 Tiempo usado para obtener los indicadores, antes y después de la solución BI</i>	<i>108</i>

INTRODUCCIÓN

En la actualidad, las empresas dedican gran parte de su tiempo y recursos a la obtención, procesamiento, aplicación y proyección de información para la toma de decisiones. En muchos casos esta información debe ser generada a partir de diversas fuentes de datos. En este contexto, muchas empresas están adoptando la tecnología de Inteligencia de Negocio para reducir el tiempo de los procesos para la generación de la información y optimizar el uso de recursos, en consecuencia, lograr ventaja competitiva.

La Inteligencia de Negocio hace referencia a los diversos procesos iterativos y automatizados que la empresa realiza para organizar y explotar la información a partir de distintas fuentes de datos con el fin de facilitar su interpretación y análisis permitiendo que la organización logre ventaja competitiva a través de la mejora en la toma de decisiones (Muñoz Hernandez , Osorio-Mass, & Zuñiga Pérez, 2016)

En el Perú, la mayoría de las empresas constructoras no cuentan con una solución de Inteligencia de Negocios que les permita explotar su inmensa cantidad de información almacenada a lo largo de los años; tales como información relacionada a costos y márgenes de ganancias obtenidas en cada uno de los proyectos, o si lo obtienen es mediante procesos tradicionalmente manuales que generan retrasos en el Análisis de Desempeño de los proyectos impidiendo la toma de decisiones efectivas y eficientes por parte de los gerentes.

La empresa de construcción objeto de estudio en esta investigación no es ajena a esta situación, por lo que ha requerido de la suscrita la implementación de una solución de Inteligencia de Negocios con el fin obtener la información pertinente, relacionada a sus proyectos, a partir de sus diversas fuentes de datos y almacenarla en un Datawarehouse que permita generar indicadores de desempeño relacionados con sus proyectos.

Este documento presenta los resultados de la investigación, realizada por la autora, sobre la implementación de una solución de Inteligencia de Negocios en una empresa constructora. Para facilitar su lectura, este documento se ha organizado en los siguientes capítulos:

En el Capítulo 1 se describe el Planteamiento del Problema detallándose la Situación Problemática, la Definición del Problema, los Objetivos y la Justificación de la investigación.

En el Capítulo 2 se describe el Marco Teórico de la investigación, el cual está conformada por: El análisis de desempeño en los proyectos de construcción, La inteligencia de negocios y su implementación y Marco conceptual.

En el Capítulo 3 se presenta la Revisión de la Literatura, el cual está conformado por los Artículos Revisados, Resumen de Revisión de la Literatura y el Benchmarking de Metodologías de Inteligencia de Negocios.

En el Capítulo 4 se describe la Aplicación de la Metodología, primero se realiza una Descripción de la Metodología de Ralph Kimball y finalmente la Aplicación de las fases I y II de la Metodología.

En el Capítulo 5 se detalla la construcción de la solución tecnológica, se presentan la aplicación de las fases III a la VI de la Metodología, desde el Análisis hasta el Despliegue.

En el Capítulo 6 se detalla los resultados obtenidos, tanto los Resultados de Categorías de Indicadores que conforman la solución; así como un Análisis Comparativo del tiempo de procesamiento de la forma tradicional y la de Inteligencia de Negocios.

Finalmente, se presentan las Conclusiones, Recomendaciones y los Trabajos Futuros para la presente tesis.

CAPÍTULO 1: PLANTEAMIENTO DEL PROBLEMA

Situación problemática

El sector de la construcción es considerado clave en el desarrollo económico de un país, pues dinamiza las actividades de otros sectores que brindan insumos como las de cemento, fierro, mayólicas, entre otras. Esta relación de dependencia de las actividades de las empresas constructoras con otras empresas proveedoras de insumos, más la diversidad de localizaciones en las que puede realizarse una obra y el cumplimiento de normativas técnicas son características distintivas de los proyectos de construcción.

Los proyectos de construcción involucran tareas de alto grado de complejidad y riesgo, por lo cual, la gestión de este tipo de proyectos requiere del análisis de diversos aspectos involucrados en la responsabilidad del gerente, quien busca lograr proyectos exitosos. Uno de los aspectos importantes en la gestión de proyectos de construcción es el Análisis del Desempeño, que se refiere a la medición y evaluación del rendimiento de los indicadores que influyen en la ejecución dentro de los plazos y costos definidos en la planificación de los proyectos (Diez-Silva, Pérez-Ezcurdia, Gimena Ramos, & Montes-Guerra, 2019).

El Análisis de Desempeño de proyectos de construcción está inmerso dentro de la función de Control de Proyectos, cuyo objetivo es el adecuado monitoreo de la ejecución de la obra en concordancia con los planos del proyecto, con las normas técnicas de construcción vigentes y con la planificación de recursos. Para tal efecto, se consideran factores de desempeño como: alcance, tiempo, costos, calidad, personal, riesgos, adquisiciones entre otros; los cuáles son conocidos como indicadores claves de desempeño o indicadores de gestión (Key Performance Indicators - KPIs). Estos indicadores ayudan a identificar el rendimiento de una determinada acción o estrategia y conocer el nivel de desempeño en base a los objetivos fijados (Bohórquez Castellanos, 2018).

En general, las empresas constructoras peruanas evidencian dificultades en el Control y Gestión de los Proyectos. Según el estudio de (Gordillo-Otárola, 2014), en el cual evidencia que si bien un 80% de empresas constructoras posee un cronograma de actividades, solamente un 49% realiza un control adecuado comparando el tiempo de ejecución real contra lo programado. Es decir, un 51% de empresas constructoras no controlan el índice de desempeño de sus proyectos tanto de sus

cronogramas como de costos. Otro problema importante es la frecuencia de monitoreo, donde se evidencia que un 37% lo realiza de manera mensual, un 32% de manera semanal y un 14% de manera quincenal, principalmente por las fuentes de información dispersas, lo que ocasiona un gran esfuerzo para su recopilación y su posterior análisis, impidiendo generar reportes de manera oportuna, confiable y con integridad para tomar decisiones efectivas y eficientes sobre el análisis de desempeño de los proyectos de construcción.

Actualmente, en muchas empresas de construcción, la información a recopilar se obtiene de diversas fuentes de datos no integradas, tales como: hojas Excel y tablas de bases de datos transaccionales, entre otras; las cuáles se consolidan de acuerdo a ciertos parámetros implementados en complementos de Excel (macros y power query), y posteriormente son almacenados en archivos Excel, para finalmente ponerlos a disposición de los gerentes responsables de dar seguimiento a los proyectos de la empresa. Esta herramienta es de gran ayuda, pero la manipulación constante de la información por parte de los usuarios genera cambios en los datos y en su parametrización, por lo cual se pierde confiabilidad e integridad de la información.

Este procedimiento de trabajo tradicional genera pérdidas económicas debido a que la información sobre los indicadores de desempeño al no estar integrada causa un gran trabajo de recopilación, impactando que la información relacionada a estos indicadores no se encuentre disponible oportunamente. Por ejemplo, el no disponer de información de las pérdidas de horas hombre por accidentes de trabajo durante la ejecución del proyecto influye en los costos estimados y el cumplimiento del plan de trabajo impactando, en consecuencia, en la competitividad de la empresa constructora.

Por lo expuesto se considera fundamental que las empresas constructoras dispongan de un mecanismo más adecuado para la recuperación, análisis, consolidación y generación de la información que permita apoyar el análisis de desempeño de los proyectos de construcción y por lo tanto mejorar la toma de decisiones.

Definición del Problema

El problema de investigación consiste en integrar las diversas fuentes de datos en un repositorio centralizado a partir del cual se parametrizan las categorías de indicadores de desempeño, tales como: costo, tiempo, y seguridad y salud. De esta manera se podrían generar reportes oportunos que permitan tomar decisiones efectivas y eficientes sobre el análisis de desempeño de los proyectos de construcción. A pesar de informar los indicadores de costo y tiempo exigidos por los órganos de supervisión gubernamental, el Area de Control de proyectos de la empresa constructora monitorea solo la categoría de indicadores de costos, sin considerar las otras categorías que también son importantes para evaluar el desempeño de los proyectos en todo su ciclo de vida.

Objetivos

Objetivo General

Implementar una Solución de Inteligencia de Negocios para apoyar el Análisis del desempeño en los proyectos de construcción que permita integrar las diversas fuentes de datos en un repositorio centralizado a partir del cual se parametrizan los indicadores de desempeño, tales como: costo, tiempo, y seguridad y salud para tomar decisiones eficientes y efectivas.

Objetivos Específicos

- a) Seleccionar una metodología para implementar soluciones de inteligencia de negocios.
- b) Seleccionar una herramienta para implementar la solución de inteligencia de negocios.
- c) Identificar indicadores de desempeño propuestos en la literatura científica para proyectos de construcción.
- d) Implementar un modelo dimensional de información relacionada al monitoreo de proyectos de construcción.
- e) Validar la propuesta en la empresa de estudio para demostrar la mejora en el tiempo de procesamiento para la obtención de los indicadores de las categorías: costo, tiempo y seguridad y salud.

Justificación

El presente trabajo de investigación se justifica como una alternativa de solución al Análisis de Desempeño de Proyectos de Construcción basado en indicadores claves. Tradicionalmente, la recopilación de la información de estos indicadores de desempeño utiliza fuentes de diversas características y diversas localizaciones las cuales son procesadas y consolidadas de manera informal. Esta propuesta considera la sistematización de la recopilación de la información haciendo uso de técnicas y herramientas de la Inteligencia de Negocios que permita disponer de información oportuna y fiable, en consecuencia, permitir a la empresa tomar decisiones de forma eficiente y efectiva.

Desde el punto de vista académico, esta investigación realiza un benchmarking de las metodologías más representativas e influyentes en el mercado de soluciones de inteligencia de negocios; a su vez, se presenta una comparativa de las principales herramientas de explotación de data. Este aporte, permitirá elegir las herramientas más adecuadas según las necesidades específicas de implementación.

Desde el punto de vista práctico, la solución de inteligencia de negocios apoyará el proceso de toma de decisiones permitiendo integrar en un repositorio (datawarehouse) único las diversas fuentes de datos. La información contenida en el datawarehouse podrá ser accedida en el momento requerido, permitiendo crear informes y tableros requeridos por el área de control de proyectos.

Además, esta propuesta podrá ayudar a las empresas constructoras mejorando su competitividad en general al sector de la industria de la construcción. Precisamente, en el informe situacional del Sector construcción del PMI mayo 2020, se señala que “En el sector construcción se conoce que existen alrededor de 23,700 empresas constructoras y 62,200 empresas inmobiliarias” ... “De acuerdo a la data registrada, las dos terceras partes de las obras son ejecutadas en los gobiernos regionales y locales, quienes, por las estadísticas de años anteriores, no han demostrado tener la mejor capacidad de ejecución” ... “Esta puede ser también la oportunidad para promover el establecimiento de las oficinas de gestión de proyectos, lo cual se encuentra expresado en el Plan Nacional de Competitividad y Productividad 2019 – 2030” (Ormeño, 2020)

Por otro lado, el sector Construcción, aparte de la Minería y Agroindustria, es uno de los más importantes para dinamizar la demanda interna y el empleo en el país. A pesar de haber sido impactado por las medidas restrictivas del gobierno peruano producto del Covid-19; tal como se observa en la Figura 1, el sector construcción registró un crecimiento de 257,0 por ciento en Mayo (20,5 por ciento con relación a mayo de 2019), por la mayor actividad de autoconstrucción, así como por la reanudación de proyectos privados y públicos. Como se sabe los presupuestos asignados a este tipo de proyectos son de grandes cantidades pudiendo hablar de millones de soles es por ello que requieren de un monitoreo constante para cumplir con lo pactado.

Figura 1

Producto Bruto Interno Mayo 2021

Producto Bruto Interno					
(Var. % anual)					
	Peso	Mayo		Enero-Mayo	
	2020 ^{1/}	2021/2019	2021/2020	2021/2019	2021/2020
PBI Primario	23,1	-3,6	34,7	-2,4	13,5
Agropecuaria	6,4	0,7	-4,0	1,5	-2,2
Pesca	0,5	8,6	102,4	11,6	67,3
Minería metálica	8,8	-8,5	82,5	-5,6	24,6
Hidrocarburos	1,5	-15,4	6,7	-16,5	-10,6
Manufactura primaria	3,6	4,9	50,0	7,7	25,2
PBI No Primario	76,9	0,7	52,7	-0,2	21,6
Manufactura no primaria	8,4	1,5	108,2	1,9	45,5
Electricidad, agua y gas	2,0	0,4	28,2	0,4	12,4
Construcción	5,6	20,5	257,0	19,9	102,2
Comercio	10,2	2,9	104,1	-5,1	31,0
Servicios	50,7	-1,9	30,3	-1,8	11,3
PBI Global	100,0	-0,4	47,8	-0,7	19,7

1/ Ponderación implícita del año 2020 a precios de 2007.
Fuente: INEI y BCRP.

Nota. Fuente: Recuperado de (BCRP, 2021)

En general la industria de la construcción suele ser inestable y las demoras son consideradas como uno de los mayores problemas que atañan a los proyectos de construcción. Los retrasos en la construcción pueden originar sobre costos, disputas entre partes, abandono del proyecto e incluso, litigios legales, por lo cual es muy importante que estas empresas de una solución de Inteligencia de Negocios como la que se plantea en este estudio.

Por otro lado, según el estudio de la Universidad de Montevideo sobre Causas de Retrasos en Proyectos de Construcción, el 80% de los retrasos se deben a problemas durante la ejecución; relacionados a la mano de obra, diseño y aspectos administrativos (Rudeli, Vilez, Gonzales, & Santilli, 2018), por lo tanto una solución de inteligencia de negocios proveerá conocimientos para apoyar los aspectos administrativos en la ejecución de un proyecto.

CAPÍTULO 2: MARCO TEÓRICO

En este capítulo se desarrolla las teorías y conceptos asociados al análisis de desempeño en los proyectos de construcción. Así mismo, se describen las teorías y conceptos asociados a la inteligencia de negocios y su implementación. Para ello se ha realizado una revisión exploratoria de literatura, las cuales han sido consolidadas en los siguientes sub capítulos: El análisis de desempeño en los proyectos de construcción; La inteligencia de negocios y su implementación; Marco conceptual de los términos relevantes.

El análisis de desempeño en los proyectos de construcción

Un proyecto de construcción o proyecto de obra civil, al igual que cualquier proyecto, se puede definir como un emprendimiento temporal que se lleva a cabo para crear un producto o servicio, y se puede considerar como un proceso, con una duración determinada y un fin concreto, compuesto por actividades y tareas diferentes, que puede ser elaborado de manera gradual (PMI, 2021).

Específicamente, un proyecto de construcción busca generar un beneficio concreto para la organización de un territorio y mejorar la calidad de vida de la comunidad y se caracteriza el alto grado de complejidad de sus actividades y tareas, las cuales se realizan de forma simultánea y con muchas dependencias entre ellas. Sin embargo, es necesario aclarar que las acciones que constituyen un proyecto no pueden ser repetitivas, debiendo tener una duración específico y estar formalmente organizadas (COPECO, 2018).

Con la finalidad de organizar los trabajos que se deben llevar a cabo en los proyectos de construcción, es necesario planificar cada actividad definida en el ciclo de vida del proyecto. Debido a la diversidad de los proyectos de construcción, muchas de sus actividades dependen de la naturaleza, tamaño y área de aplicación (Gordillo-Otárola, 2014). En ese contexto la literatura agrupa las actividades en dos grandes fases: Planeamiento y Ejecución. En el Planeamiento se llevan a cabo actividades relacionadas al estudio de campo y al desarrollo del expediente técnico. En la Ejecución se realizan actividades referidas a la construcción y supervisión de la obra (Gordillo-Otárola, 2014).

Durante la fase de ejecución intervienen diversos elementos que hacen complejo su control; estos elementos pueden incluir recursos, actividades, información, entre otras. Por ello es de vital

importancia desarrollar de manera sistemática un proceso de análisis de desempeño que utilicen indicadores de desempeño establecidos en la literatura académica y profesional, conocidos como KPI (Key Performance Indicator). En ese sentido, el análisis de desempeño se refiere a la medición y evaluación del rendimiento de los indicadores que influyen en la ejecución dentro de los plazos y costos definidos en la Planificación de los proyectos, es decir se enfoca en los aspectos relacionados con la evaluación cuantitativa utilizando medidas de progreso que permiten tomar acciones correctivas cuando es necesario. (Diez-Silva, Pérez-Ezcurdia, Gimena Ramos, & Montes-Guerra, 2019).

Tradicionalmente para el análisis de desempeño se consideraban las variables definidas por el triángulo de hierro propuesto por el doctor Martín Barnes en 1969, estas variables son: Tiempo, Coste y Alcance; sin embargo, debido a la multidimensionalidad que se requiere en un proyecto, surge la necesidad de nuevos modelos de medición de rendimientos (Diez-Silva, Pérez-Ezcurdia, Gimena Ramos, & Montes-Guerra, 2019).

Utilizando los resultados teóricos de revisión sobre medición del desempeño y éxito en la dirección de proyectos, se sintetiza a continuación los modelos propuestos por diversos autores sobre indicadores a considerar en el diseño de un sistema de evaluación del rendimiento de proyectos de construcción. (Diez-Silva, Pérez-Ezcurdia, Gimena Ramos, & Montes-Guerra, 2019).

1. (Coste, Tiempo, Calidad, Seguridad, Salud, Medioambiente, Satisfacción del cliente, Comunicación) propuesto por Cheung y Seung.
2. (Coste, Tiempo, Calidad, Propietarios) propuesto por Ling.
3. (Coste, Programación) propuesta por Cho & Hyun.
4. (Equipo de trabajo, Mejora continua, Tiempo, Presupuesto, Especificaciones, Recursos/eficiencia, Cosas correctas/efectividad, Seguridad, Defectos, Involucrados, Conflictos) propuesta por Toor & Ogunlana.
5. (Programación, Coste, Calidad, Disputas) propuesta por Jha & Iyer.
6. (Interesados, Tiempo, Coste, Aplicación de estándares, Implementación, Formación) propuesto por Bernroider & Ivanov.

7. (Coste, Tiempo, Satisfacción del cliente, Implementación del SGC, Equipo de proyecto, Gestión de cambios, Gestión de materiales, Gestión de seguridad) propuesta por Luu, Kim, & Huynh.

La elección de uno de los modelos mencionados puede seguir diversos criterios dependiendo del tamaño, tipo, alcance, desafíos del proyecto de construcción, por ello se hace imposible generar una lista universal de criterios fijos que deben considerarse en cualquier proyecto. Por ello, en base a los modelos presentados, se indican las cuatro categorías de mayor especificidad y representatividad de acuerdo con la frecuencia con la que son mencionados, estos : Coste, Tiempo, Calidad, Seguridad y Salud

Es esencial resaltar la importancia del análisis de desempeño en el proceso de control de proyectos, y es considerado un factor clave que determina su éxito, ya que permite vigilar el desempeño de los objetivos planificados, y a su vez evaluar el desempeño real de dichas actividades, con ello la organización podrá tomar las acciones correctivas en el momento indicado, con el objetivo de cumplir su plan de trabajo planificado. (Diez-Silva, Pérez-Ezcurdia, Gimena Ramos, & Montes-Guerra, 2019).

Inteligencia de Negocios

La literatura revisada muestra que una de las alternativas para realizar el análisis de desempeño basado en indicadores de rendimiento en el área de Control de Proyectos es la Inteligencia de Negocios (Business Intelligence BI), la cual apoyará la gestión organizacional, la toma de decisiones efectivas por parte de la gerencia y así contribuir en la consecución de los objetivos estratégicos (Garcia & Macias, 2010).

Con el fin de organizar este subcapítulo, se han considerado los siguientes apartados: Información y afines, Sistemas de Información, Gestión de Información y de Conocimiento, Definición de Inteligencia de Negocios, Proceso de Inteligencia de Negocios y finalmente Componentes de una Solución de Inteligencia de Negocios.

Información y afines

Actualmente las organizaciones afrontan retos desafiantes en el mercado competitivo, por ello es fundamental que se desarrollen procesos adecuados de planificación estratégica con la finalidad

de lograr una gestión sostenida y escalable frente al crecimiento continuo de la organización, sin perder las características fundamentales: eficiencia, eficacia y calidad de la toma de decisiones (Cordero-Guzmán & Rodríguez-López, 2017).

El factor que permitirá alcanzar este propósito es la información que se pueda recopilar de la organización, contenida en diversas fuentes de datos y disponible de diversas formas, pudiendo ser: en tiempo real o histórica. En este sentido, la información constituye un recurso decisivo que puede determinar el éxito de una organización; por ello es necesario definir y distinguir los conceptos relacionados a ella, como lo son: dato, información y conocimiento (Cordero-Guzmán & Rodríguez-López, 2017).

El dato es la mínima unidad semántica y se relaciona con elementos primarios de información que por sí solos son irrelevantes y no orientativos como apoyo a la toma de decisiones. La información se puede definir como un conjunto de datos procesados que tienen un significado (relevancia, propósito y contexto) y que, por lo tanto, son de utilidad para quien debe tomar decisiones. Con respecto al conocimiento, está dado en el análisis, síntesis, determinación de experiencias de la información; se origina y aplica en la mente de los conocedores permitiendo incorporar nuevas experiencias de información (Novoa Torres, Bermudez Huérfino, & Zamora Carrillo, Diciembre 2019). La Figura 2 muestra a detalle los niveles de la pirámide de generación del conocimiento.

Figura 2

Pirámide del conocimiento



En el contexto empresarial o desde el punto de vista del Control de Proyectos, la información constituye un recurso decisivo que puede determinar el éxito de una organización siempre y cuando se gestione de manera eficiente (Suarez Alfonzo, Cruz Rodriguez, & Perez Macias, 2015). En la actualidad, la problemática del Data Quality (DQ) o calidad de datos en español es cada vez más evidente, ello debido a la multidiversidad de fuentes en el proceso de recopilación de información. En tal sentido es necesario asegurar la calidad evaluando la información en base a las siguientes categorías y dimensiones mostradas en la Tabla 1 (M.Strong, W.Lee, & Y.Wang, August 2012).

Tabla 1

Paradigma de la información

Categoría Calidad de Datos	Dimensiones
Intrínseca	Precisión, Objetividad, Credibilidad
Accesibilidad	Accesibilidad , Acceso seguro
Contextual	Relevante, Valor añadido, Completa
Representación	Fácil de entender, Concisa, Consistente

Nota. Fuente: Adaptado de (M.Strong, W.Lee, & Y.Wang, August 2012)

1. Intrínsecas:

Esta categoría considera dimensiones tradicionales: precisión de los datos y confiabilidad en la fuente. La información que supere el análisis desde todas las dimensiones mencionadas puede ser considerada como “Lista para usar”; es decir, de buena calidad y apta para iniciar un proceso decisorio o un análisis de Inteligencia de Negocios.

2. Accesibilidad:

Esta categoría es la encargada de gestionar accesos a usuarios para que puedan revisar la información deseada. La accesibilidad está directamente relacionada con la seguridad, ya que las organizaciones deben garantizar los accesos seguros a la información, apoyados en

la implementación de medidas de seguridad (Ej.: criptografía, claves de acceso, controles, etc.).

3. Contextual:

La categoría contextual está relacionada al contexto y a la completitud de la información, en tal sentido, los datos deben tener relación con el tema en cuestión y ser actuales. El problema de datos incompletos, datos medidos o definidos inadecuadamente impide un correcto análisis. Para resolver estos problemas de DQ contextuales, la organización debe iniciar proyectos para proporcionar datos relevantes que agreguen valor a las tareas de los consumidores de datos.

4. Representación:

La categoría representación considerar atributos como: el idioma, el soporte, las unidades de medida y la codificación con las que el consumidor está familiarizado. Es importante que se tenga consistencia metodológica de los datos, en nuestro contexto del Control de Proyectos, imaginemos que el reporte gerencial muestra solo los códigos de los proyectos en ejecución, más no la descripción del proyecto; esta casuística generaría un problema de fácil entendimiento, al no representar al proyecto de forma concisa.

Cabe precisar que el énfasis que se hará de cada atributo dependerá de la función que se ocupe con respecto a la información. En nuestro contexto, las gerencias involucradas en el Control de Proyectos pondrán el énfasis en las dimensiones contextuales e intrínsecas de la información que apoye al proceso decisorio de la toma de decisiones.

De lo anteriormente detallado, se manifiesta claramente, que la información es un recurso importante para la consecución de los objetivos de cualquier organización y que, por tanto, debe administrarse de manera eficiente y eficaz. En tal sentido, la organización debe apoyarse de instrumentos necesarios para procesar y gestionar la información, siendo estos: los sistemas y redes de información (García Orozco, Javier Francisco, 2010).

Sistemas de Información

La literatura actual muestra un gran repertorio de definiciones para el término Sistemas de Información, tal vez una definición más precisa define al SI como : “conjunto formal de procesos que, operando sobre una colección de datos estructurada de acuerdo a las necesidades de la empresa, recopila, elabora y distribuyen selectivamente la información necesaria para la operación de dicha empresa y para las actividades de dirección y control correspondientes, apoyando, al menos en parte, los procesos de toma de decisiones necesarios para desempeñar funciones de negocio de la empresa de acuerdo con su estrategia” (Andreau, Ricart, & Valor, 1991).

En general, todo sistema de información cumple las siguientes funciones: utiliza como materia prima los datos, los cuales almacena, procesa y transforma para obtener como resultado final información, la cual será suministrada a los diferentes usuarios del sistema, existiendo además un proceso de retroalimentación, en la cual se ha de valorar si la información obtenida se adecua a lo esperado. Este proceso ayuda a generar información de salida útil, ayudando en el proceso de toma de decisiones(Ver Figura 3) (Hernandez Trasobares, 2013).

La literatura actual muestra muchas clasificaciones de los Sistemas de Información, sin embargo, la más útil es la propuesta por K y J Laudon (1996), en la cual los SI se agrupan según su utilidad en los cuatro niveles de la organización empresarial: nivel operativo (referido a las operaciones diarias de la organización), nivel del conocimiento (involucra al personal que gestiona la información), nivel administrativo (orientado a gerencias intermedias de la organización) y un nivel estratégico (orientado a la alta dirección de la empresa). Estos niveles clasifican a los SI en seis tipos específicos:

Sistemas de procesamiento de Operaciones (SPO): Son utilizados para la administración de aquellas operaciones diarias de rutina necesarias en la gestión empresarial (aplicaciones de nóminas, auditoría, registro y datos de empleados). Estos sistemas generan información que será utilizada por el resto de la organización siendo utilizados por el Nivel Operativo.

Sistemas de Trabajo del Conocimiento (STC): Son los encargados de apoyar a los agentes que manejan información en la creación e integración de nuevos conocimientos para la empresa (estaciones de trabajo para la administración); forman parte del nivel de conocimiento.

Sistemas de automatización en la oficina (SAO): Buscan incrementar la productividad de los empleados que manejan la información en los niveles inferiores de la organización, para ello utilizan herramientas como (procesador de textos, agendas electrónicas, programas Office, correo electrónico, etc); se encuentran encuadrados en el nivel de conocimiento al igual que los STC.

Sistemas de apoyo a ejecutivos (ESS): Apoyan el nivel administrativo en la organización, empleados en el proceso de planificación, control y toma de decisiones proporcionando informes sobre las actividades ordinarias (control de inventarios, presupuesto anual, análisis de las decisiones de inversión y financiación).

Sistemas de Información Gerencial (MIS): Son esenciales para la toma de decisiones estratégicas las cuales requieren el entendimiento de situaciones complejas, se apoyan en el empleo de gráficos y comunicaciones avanzadas. Son utilizados por la alta dirección de la organización con el fin de elaborar la estrategia general de la empresa (planificación de ventas para 4 años, plan de operaciones, planificación de la mano de obra).

Sistemas de Apoyo a las Decisiones (DSS): Empleados por la gerencia intermedia de la organización apoyando la toma de decisiones de situaciones rápidamente cambiantes y no especificadas fácilmente con anticipación. A su vez ayudan a los distintos usuarios en el proceso de toma de decisiones, a la hora de utilizar diferentes datos y modelos para la resolución de problemas no estructurados (análisis de costes, análisis de precios y beneficios, análisis de ventas por zona geográfica, etc.).

Todos estos sistemas de información a su vez podrían analizarse según las diferentes áreas de la empresa: ventas y mercadotecnia, manufactura y producción, finanzas, contabilidad y recursos humanos. Para cada una de estas áreas existe un conjunto específico de aplicaciones informáticas y equipos, los cuales han de estar coordinados entre sí. Por ello resulta clave la correcta planificación y desarrollo de los sistemas de información, en tal sentido la gestión

la información y el conocimiento constituye para muchos la respuesta práctica y sustancial a las necesidades actuales de la sociedad y de las empresas en general, al hacerlas más competitivas en los tiempos de cambio (García Orozco, Javier Francisco, 2010).

Gestión de información y conocimiento

El concepto de gestión de información (GI), de acuerdo a Woodman (1985), es considerado como el proceso para la obtención de la información adecuada, en la forma correcta, para la persona u organización indicada, al precio adecuado, en el tiempo oportuno y lugar apropiado, para tomar la decisión adecuada. La gestión del conocimiento (GC), se refiere a un proceso sistémico y específico el cual hace uso de la información útil de una organización, enfocado en el correcto manejo de los recursos humanos (capital humano), formados y preparados para obtener el máximo provecho y emplear la inteligencia colectiva para lograr objetivos estratégicos (Gil-Montelongo, López-Orozco, Molina-García, & Bolio-Yris, 2011).

En este contexto y basado en los enfoques de GC y de GI, así como la interrelación dialéctica entre ambos términos, las cuales pretenden extraer el conocimiento de las personas y almacenarlo con ayuda de los sistemas tecnológicos para su reutilización; en la Figura 4 se representan las fases del proceso de gestión de información, el cual tiene como base la Tecnología de Información (TI), las herramientas y aplicaciones utilizadas en las TI dan soporte a las fases de adquisición, distribución, interpretación; de tal modo que estas se lleven a cabo de manera óptima con el fin de disponibilizar una base de conocimiento en la organización (Gil-Montelongo, López-Orozco, Molina-García, & Bolio-Yris, 2011)

Figura 3

Proceso de Gestión de Información

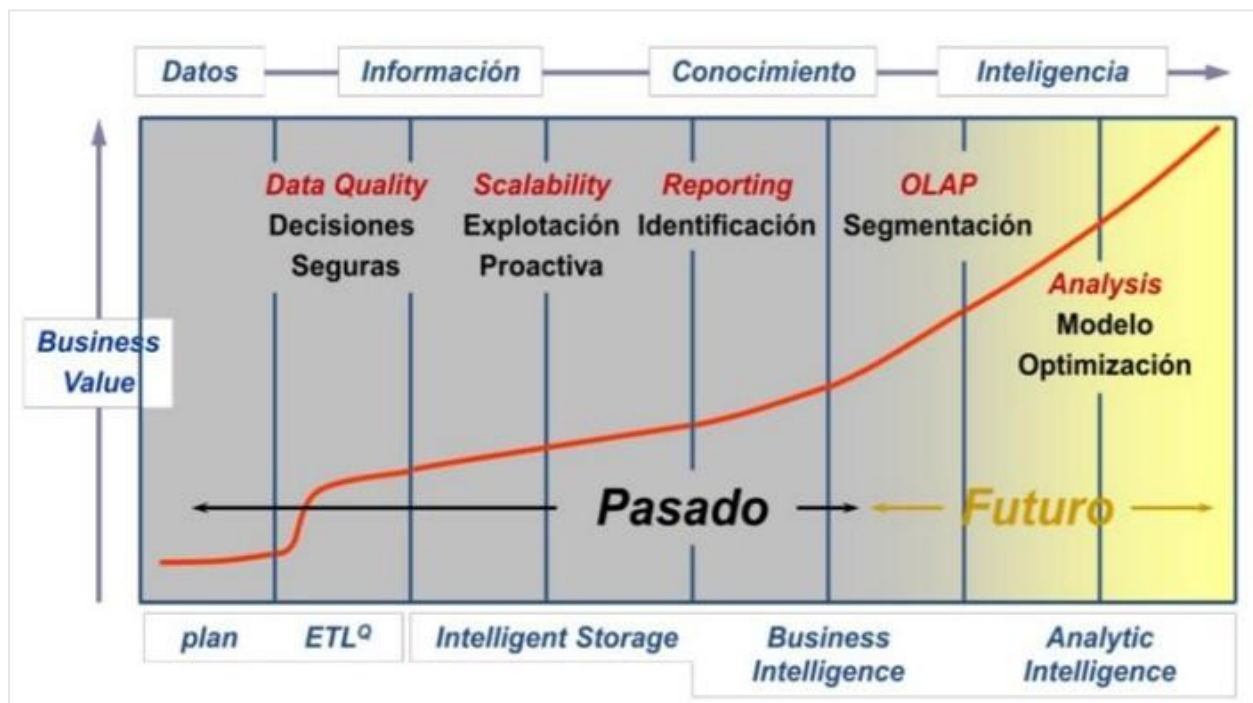


Nota. Fuente: Adaptado de (Gil-Montelongo, López-Orozco, Molina-García, & Bolio-Yris, 2011)

Los rápidos avances en las tecnologías de la información han contribuido en el aumento del volumen y variedad de información contenidas en múltiples fuentes de información generando que el análisis se torne complejo, lo que afirma que el almacenamiento de información ha crecido más rápidamente que la capacidad de analizarla. Ello ha detonado que las empresas busquen generar inteligencia del conocimiento, dando lugar a la Inteligencia de Negocios, la cual será detallada posteriormente ;y a la Inteligencia Analítica (Albarran & Salgado Gallegos, 2013). Para concluir, la Figura 4 muestra la cadena de valor de la Inteligencia detallada a lo largo de este subcapítulo.

Figura 4

Cadena de Valor de la Inteligencia



Nota. Fuente: Adaptado de (Albarran & Salgado Gallegos, 2013)

Definición de Inteligencia de Negocios

Partiendo del concepto de gestión del conocimiento, se acuña el de Business Intelligence (BI) (inteligencia de negocios, en su traducción al español), el cual, en primer lugar, ha sido abordado en el área tecnológica con el fin de implementar sistemas de información que refuercen las estrategias a aplicar en áreas directivas y administrativas. La evolución del concepto ha sido

compleja debido a que incluye otros conceptos como la minería de datos, el manejo de datos a gran escala y su analítica, tanto que a BI se le considera una “sombrilla” que incluye metodologías, prácticas, sistemas y aplicaciones que permitirán un mejor entendimiento del mercado basado en los datos, para posteriormente tomar decisiones (Mora, 2018).

Desde su primera mención en 1958 por el pionero de la ciencia de la información H. P. Luhn, la inteligencia empresarial (BI) se ha definido de muchas formas y el concepto aún está evolucionando. Los investigadores del Instituto Gartner, encabezados por Howard Dresner, popularizaron la siguiente definición del término BI: “conceptos y métodos para mejorar la toma de decisiones comerciales mediante el uso de sistemas de apoyo basados en hechos”. Forrester, utiliza la siguiente definición: “es un conjunto de metodologías, procesos, arquitecturas y tecnologías que transforman los datos en información significativa y útil para la toma de decisiones estratégicas, tácticas y operativas” (Rimvydas Skyrius, 2012).

El detalle de estas pocas definiciones seleccionadas permite comprender la estructura subyacente de BI. En primer lugar, comprender que BI reúne un conjunto de tecnologías, procesos o enfoques que funcionan juntos como un sistema complejo. En pocas palabras, BI brinda la capacidad de obtener conocimientos y desempeñar una mejor toma de decisiones mediante el uso de tecnologías y herramientas, sin olvidar el factor humano, el cual debe poseer las siguientes competencias: capacidad de liderazgo, perfil analítico, manejo de tecnología, cultura empresarial y toma de decisiones (Mora, 2018).

Existen muchas particularidades de los usuarios a los cuáles está dirigido BI, sin embargo, los sistemas de Business Intelligence tienen en común ciertas características mencionadas a continuación (Chalco Valencia & Galo Andrés, 2018):

Accesibilidad permanente a la información: Se debe garantizar el acceso de los usuarios a los datos y la información que contienen, independientemente de su procedencia, esta información debe estar disponible y ser confiable.

Orientación al usuario final: El propósito principal es que el usuario pueda utilizar herramientas de manera intuitiva y así poder consumir la información y auto gestionar sus necesidades, sin depender de un conocimiento técnico.

Apoyo en la toma de decisiones: No se limita a realizar la presentación de la información, de hecho, provee una herramienta que permita seleccionar y manipular las variables de interés, de modo que se apoye la toma de decisiones de una manera informada.

En toda organización existen tres tipos de necesidades de información: Operacional, táctica y estratégica. Los sistemas de Business Intelligence recopilan los datos registrados en la operación del día a día, de forma integrada, ordenada y consistente, con el propósito de convertirla en información valiosa para los niveles tácticos y estratégicos. En la Figura 5, se muestra la estructura organizacional comentada (Rimarachin, 2015).

Figura 5

Pirámide de Niveles Organizacionales BI



Nota. Fuente: Adaptado de (Rimarachin, 2015)

A continuación se detallan los niveles organizacionales mostrados (Muñoz Hernandez , Osorio-Mass, & Zuñiga Pérez, 2016):

Nivel operacional: Corresponde a todos procesos vitales que respaldan la operación diaria de la organización asegurando su correcto funcionamiento. En este nivel se tiene toda la información cruda que servirá como input para los siguientes niveles operacionales.

Nivel táctico: Conformado por la integración funcional de la comunicación y la coordinación de las áreas de la organización. En este nivel, los gerentes necesitan que se les presente soluciones que les apoyen a tomar decisiones en el mediano y corto plazo.

Nivel Estratégico: Los procesos que conforman el nivel estratégico, son los que se encargan de gestionar información de forma más agregada, la que nos permite un adecuado macro análisis con la finalidad de interpretar y decidir a partir de los resultados obtenidos de las diversas operaciones claves de la empresa como las ventas, el gasto, los impuestos etc.

Proceso de Business Intelligence

Para comprender cómo una organización puede crear inteligencia a partir de sus datos y así proveer a los usuarios finales el acceso a esta información de manera oportuna. El proceso de BI está dividido en cinco fases, las cuales serán explicadas teniendo como referencia la Figura 6, que sintetiza todo el proceso (Bernabeu, 2009):

Figura 6

Fases del Proceso BI



Nota. Fuente: Adaptado de (Bernabeu, 2009)

Para ello se definen las 5 fases de la siguiente forma:

FASE 1. Dirigir y planear: En esta fase inicial se deberán recolectar los requerimientos de información específicos de los diferentes usuarios, así como entender sus diversas necesidades, para que luego en conjunto con ellos se generen las preguntas que les ayudarán a alcanzar sus objetivos.

FASE 2. Recolección de información: Se realiza el proceso de extraer desde las diferentes fuentes de información de la empresa, tanto internas como externas. Se recopilarán todos los orígenes de datos que nos ayuden a responder las interrogantes formuladas en la fase 1.

FASE 3. Procesamiento de datos: En esta fase se integran y cargan los datos en crudo en un formato utilizable para el análisis. Esta actividad puede realizarse mediante la creación de una nueva base de datos, agregando datos a una base de datos ya existente o bien consolidando la información.

FASE 4. Análisis y producción. Ahora se procede a trabajar sobre los datos extraídos e integrados, utilizando herramientas y técnicas propias de la tecnología BI, para crear inteligencia. Como resultado final de esta fase se obtendrán las respuestas a las preguntas, mediante la creación de reportes, indicadores de rendimiento, cuadros de mando, gráficos estadísticos, etc.

FASE 5. Difusión. Finalmente, se les entregará a los usuarios que lo requieran las herramientas necesarias, que les permitirán explorar los datos de manera sencilla e intuitiva.

Una de las fases más complejas del Proceso de BI es la del procesamiento de datos, en esta, se lleva a cabo el proceso de extracción, transformación y carga de datos (ETL), la importancia de este radica en el hecho de clasificar información para poder obtener datos relevantes contenidos en cada base de conocimiento con ello se asegura la asertividad y confiabilidad en la información, la cual posteriormente será utilizada en análisis que apoyen la toma de decisiones estratégicas (Contreras de La Fuente, Agosto , 2018).

Figura 7

Esquema típico del proceso ETL



Nota. Fuente: Adaptado de (Espinoza, 2010)

ETL son las siglas en inglés de Extract, Transform and Load; en español: Extraer, Transformar y Cargar. Se define como el proceso que permite a las organizaciones recopilar datos de múltiples fuentes y disponibilidad la información en un repositorio unificado, pudiendo ser: otra base de datos, un datamart o un datawarehouse. El objetivo de este proceso es tener información formateada y limpia, la cual podrá apoyar el análisis de un proceso de negocio (ver Figura 7) (Espinoza, 2010).

Para poder alimentar nuestras fuentes de datos dimensionales: datawarehouse o datamart, se torna crucial una correcta implementación del proceso ETL para poder consolidar datos que puedan ser procesados de forma eficiente y rápida. El ETL está conformado por tres etapas principales; extracción, transformación y carga, los cuales se presentan a continuación de forma separada, en cada uno de ellas se mencionan las actividades a considerar para obtener un mejor rendimiento en cada etapa (Contreras de La Fuente, Agosto , 2018):

1. Extracción:

Es el primer paso por lo tanto es responsable de recopilar la información que se tiene en las diferentes fuentes de datos y asegurar el recaudamiento diario de los datos. Generalmente son datos de texto plano como archivos de texto, hojas de cálculo, bases de datos transaccionales entre otros. En el proceso de extracción se pueden encontrar tres tipologías de datos mostrados en la Figura 8, las cuales debemos considerar para obtener mayor eficiencia del proceso:

Figura 8

Datos importantes en el Proceso de Extracción



Nota. Fuente: Elaboración propia

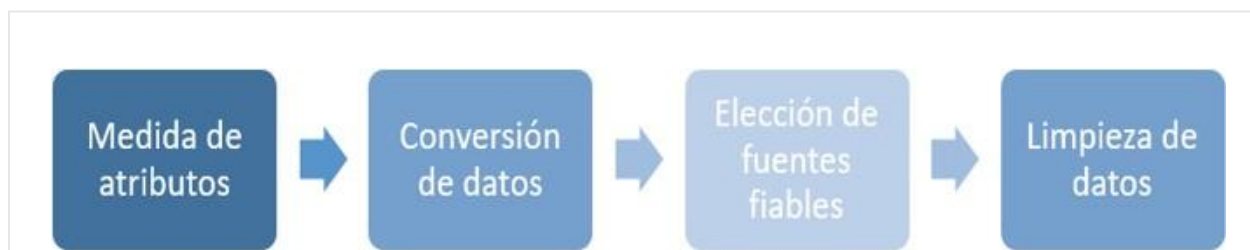
Los datos históricos, son aquellos datos almacenados a través del tiempo, estos nos permiten conocer la evolución de la organización, por lo tanto, es de suma importancia disponibilizar la historia de las fuentes consideradas. Por otro lado, un dato operacional, son los datos actuales producto de la operación diaria, generalmente almacenados en los sistemas transaccionales(OLTP). Finalmente, se deben considerar los datos relevantes encontrados en diferentes formatos, para ello se debe inspeccionar la variedad de fuentes de almacenamiento utilizadas por la organización.

2. Transformación:

Después de haber identificado las fuentes requeridas de información, se deben estandarizar los datos con diferentes formatos. Para ello, se requiere de la limpieza absoluta de la información y así poder unificar todos los datos a un mismo formato específico. Los datos deben ser congruentes, así como compatibles con el almacén y es debido a ello que es necesario definir estándares a la información que deban ser integrados para normalizar de manera correcta (ver Figura 9).

Figura 9

Actividades en la Transformación de Datos



Nota. Fuente: Adaptado de (Contreras de La Fuente, Agosto , 2018)

En muchos escenarios, un dato se encuentra contenido en diversas fuentes de información, es por ello que se debe revisar la documentación de fuentes para así poder identificar el origen más confiable. Luego de ello, se deben establecer medidas de los atributos (valores mínimo y máximo de datos), así delimitamos los caracteres y prevenimos una sobrecarga; y, por otro lado, la actividad de conversión de datos nos ayudará a nombrar los campos en función a la relación que exista entre ellos. Para culminar la etapa de transformación, se requiere de una limpieza de datos adecuada,

para ello se excluye información incorrecta o inconsistente con el fin de entregar solo información valiosa.

3. Carga:

En esta etapa, los datos procedentes de la fase anterior (fase de transformación) son cargados en el sistema de destino. Dependiendo de los requerimientos de la organización, este proceso puede abarcar una amplia variedad de acciones diferentes. Existen dos formas básicas de desarrollar el proceso de carga: por acumulación simple y Rolling definidas a continuación: (Power Data, 2017).

- *Acumulación simple:* Consiste en realizar un resumen de todas las transacciones comprendidas en un período de tiempo seleccionado y transportar el resultado como una única transacción hacia el datawarehouse, almacenando un valor calculado que consistirá típicamente en un sumatorio o un promedio de la magnitud considerada. Es la forma más sencilla y común de llevar a cabo el proceso de carga.
- *Rolling:* Este proceso sería el más recomendable en los casos en que se busque mantener varios niveles de granularidad. Para ello se almacena información resumida a distintos niveles, correspondientes a distintas agrupaciones de la unidad de tiempo o diferentes niveles jerárquicos en alguna o varias de las dimensiones de la magnitud almacenada (por ejemplo, totales diarios, totales semanales, totales mensuales, etc.).

Componentes de una Solución BI

Al pensar en Business Intelligence, inevitablemente nos vienen a la mente conceptos como datamart, datawarehouse, OLAP (On Line Analytical Processing), indicadores de rendimiento y Reporting. Alguno de estos son esenciales y otros solo se presentan de acuerdo con el nivel de madurez en que se encuentre el proyecto de Inteligencia de Negocios (Acosta, 2019).

Un datamart es un fracción o subconjunto de un almacén de datos orientado a una línea de específica (por ejemplo, el departamento de ventas), el cual busca satisfacer necesidades específicas de información. Por ello, obtiene datos de pocas fuentes de información, acorde a ello

presenta datos resumidos y específicos para el análisis haciendo fácil su uso, diseño e implementación (Acosta, 2019).

Un datawarehouse es un repositorio unificado, que puede ser definido como una colección de datamarts. La información que contiene se encuentra integrada permitiendo que la organización o empresa ejecute análisis potentes en grandes volúmenes (petabytes) de datos históricos de formas que una base de datos estándar no podría procesar. (Mazon, Pardillo, & Trujillo, 2011)

Una de las definiciones más famosas sobre DW, es la de William Harvey Inmon, quien define: “Un Data Warehouse es una colección de datos orientada al negocio, integrada, variante en el tiempo y no volátil para el soporte del proceso de toma de decisiones de la gerencia”. Debido a que W. H. Inmon, es reconocido mundialmente como el padre del DW, la explicación de las características más sobresalientes de esta herramienta mostrada en la Figura 10 se basó en su definición (Bernabeu, 2009).

Figura 10

Características de un Data Warehouse



Nota. Fuente: Adaptado de Inmon

- **Orientado al negocio:** Los datos están organizados por temas para facilitar el entendimiento por parte de los usuarios, de forma que todos los datos relativos a un mismo tema queden integrados en un solo lugar. Por ejemplo, todos los datos de un cliente pueden estar consolidados en una misma tabla, todos los datos de los productos en otra, y así sucesivamente.

- **Integrada:** Los datos se deben integrar en una estructura consistente que posea diversos niveles de detalle para así responder las necesidades de información por parte de los usuarios. Es vital que, al integrar la información, se depuren las inconsistencias o redundancias existentes entre los diversos sistemas operacionales, como por ejemplo inconsistencias: en nomenclatura, en unidades de medida, en formatos de fechas, múltiples tablas con información similar.
- **Variante en el tiempo:** Los datos de cualquier organización son dinámicos, es decir, varían a lo largo del tiempo, en consecuencia, el datawarehouse está preparado para registrar estos cambios y no alterar la realidad que había en el momento en que se almacenaron, evitando así la problemática que ocurre en los sistemas operacionales, que reflejan solamente el estado de la actividad de negocio presente.
- **No volátil:** Es importante recalcar que la información es útil para el análisis y la toma de decisiones solo cuando es estable. Mientras que en un sistema transaccional los datos operacionales se reemplazan, en cambio, una vez que entran en el DW no cambian. Por ello, en la carga del datawarehouse solo se permiten dos tipos de operaciones: la carga de datos y el acceso a los mismos.

OLAP es un recurso utilizado en el área denominada Inteligencia de Negocios ,cuyo objetivo es agilizar la consulta de grandes cantidades de datos, para lo cual recurre a estructuras multidimensionales o cubos OLAP que contienen datos resumidos de grandes Bases de datos o Sistemas Transaccionales en línea OLTP (On-line Transaction Processing) (Feliciano Morales, Cuevas Valencia, & Martínez Castro, 2016).

Los sistemas OLAP se caracterizan por procesar grandes cantidades de información a partir de lo que resultaría ser a modo rutinario una serie de complejas consultas sobre una base de datos de forma sencilla. El tiempo de respuesta de las consultas es menor, ello debido a que estos sistemas poseen una enorme potencia de cálculo, técnicas de indexación especializadas y datos que se encuentran pre-compilados sobre una estructura intermedia (Feliciano Morales, Cuevas Valencia, & Martínez Castro, 2016).

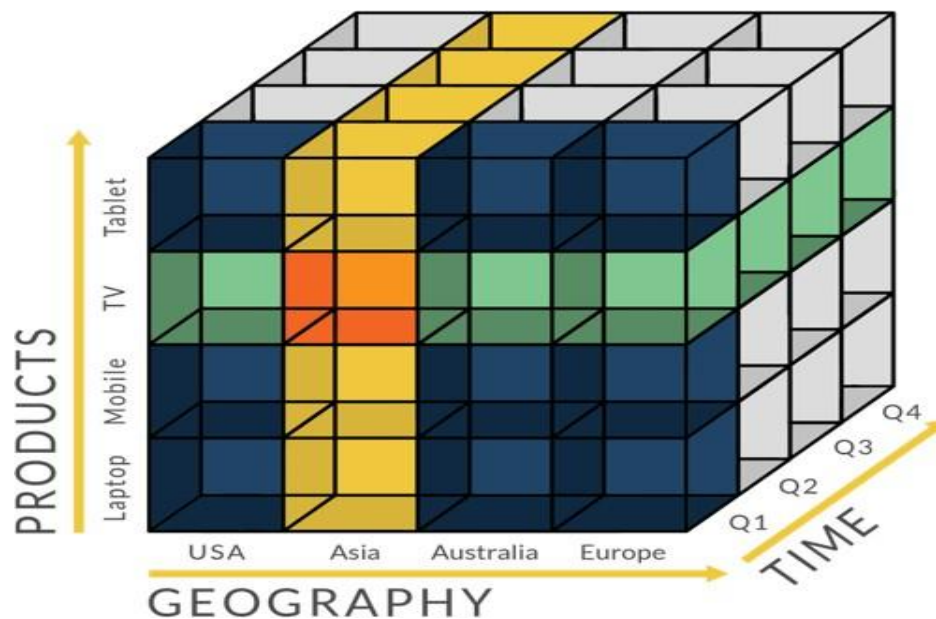
La tecnología OLAP soporta las operaciones analíticas comunes, por ejemplo, la consolidación, la profundización y la navegación. La consolidación implica agregar datos, por ejemplo,

totalizaciones simples o expresiones complejas que impliquen datos interrelacionados. Por otro lado, la profundización es la operación inversa de la consolidación e implica mostrar información detallada comprendida en los datos consolidados. La navegación hace referencia a la capacidad de examinar los datos desde diferentes puntos de vista (Feliciano Morales, Cuevas Valencia, & Martínez Castro, 2016).

En el análisis multidimensional presentado como ejemplo en la Figura 11, los datos se representan mediante dimensiones como Producto, Geografía, Tiempo. En general, las dimensiones se relacionan en jerarquías, por ejemplo, la dimensión Geografía se puede jerarquizar de la siguiente forma: Continente → País → Distrito. El tiempo también es una dimensión estándar con sus propias jerarquías tales como: día, semana, mes, trimestre y año.

Figura 11

Ejemplo de Cubo OLAP



Nota. Fuente: Adaptado de (Classora Technologies , 2012)

Para implementar procesamiento OLAP, generalmente integrado a las bodegas de datos, se han definido diversas arquitecturas: R-OLAP, M-OLAP y H-OLAP, las cuales coinciden en incluir un

esquema de metadatos, el cual define como están construidas las medidas y las dimensiones, describe sus características y las operaciones que se pueden llevar a cabo. A continuación, se detallan (Universidad Pontificia de Salamanca, 2006):

M-OLAP (Multidimensional OLAP):

Sistema OLAP que posee los datos almacenados en una base de datos multidimensional. Esta implementación mejora los tiempos de acceso a los datos ya que están pre calculados por ende requieren un mayor espacio para el almacenamiento.

R-OLAP (Relacional OLAP):

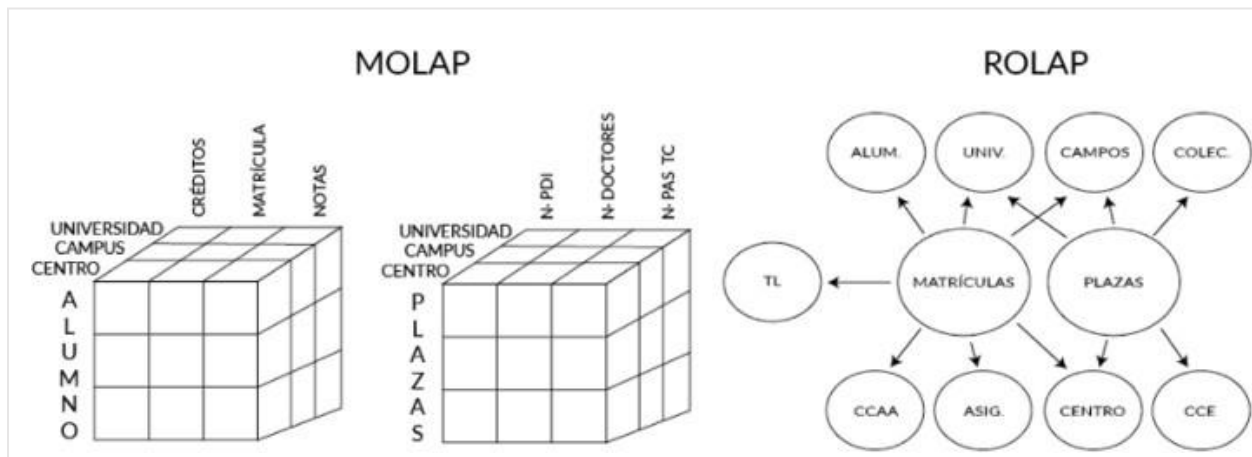
Sistema OLAP que mantiene los datos almacenados en una base de datos relacional. Para esta implementación se realiza un Cubo virtual o tablas en forma de estrella con lo que se consigue una mayor capacidad de almacenamiento sacrificando tiempo de respuesta.

H-OLAP (Hybrid OLAP): Combinación de los dos sistemas anteriores donde los datos se almacenan repartidos en implementaciones M-OLAP y R-OLAP. Esta combinación permite obtener ventajas de ambas implementaciones según donde se almacene el dato y las operaciones que se vayan a realizar sobre él.

Las arquitecturas presentadas anteriormente de OLAP, se esquematizan en la Figura 12 en la cual podemos observar la distribución y composición de los datos en cada arquitectura.

Figura 12

OLAP según su arquitectura de almacenamiento



Nota. Fuente: Elaboración propia

En cualquier herramienta OLAP encontramos dos tipos de variables características, independientemente del modo en que estén almacenados los datos: dimensión e indicador. Una dimensión contiene aspectos generales y aspectos específicos enfocados al desarrollo de un ambiente funcional informacional centrado en la vigilancia científica y tecnológica de las organizaciones; en el cual a cada uno de los criterios se puede asociar un indicador cuantitativo (Bouza Betancourt, Couto Torres, & Sosa Perez, 2017).

El indicador es un dato o una información que sirve para conocer o valorar las características de un hecho o para determinar su evolución futura. Los indicadores son puntos de referencia, que brindan información cualitativa o cuantitativa, conformada por uno o varios datos, constituidos por percepciones, números, hechos, opiniones o medidas, que permiten seguir el desenvolvimiento de un proceso y su evaluación, y que deben guardar relación con este. Igualmente puede tratarse como una magnitud utilizada para medir o comparar los resultados efectivamente obtenidos, en la ejecución de un proyecto, programa o actividad (Bouza Betancourt, Couto Torres, & Sosa Perez, 2017).

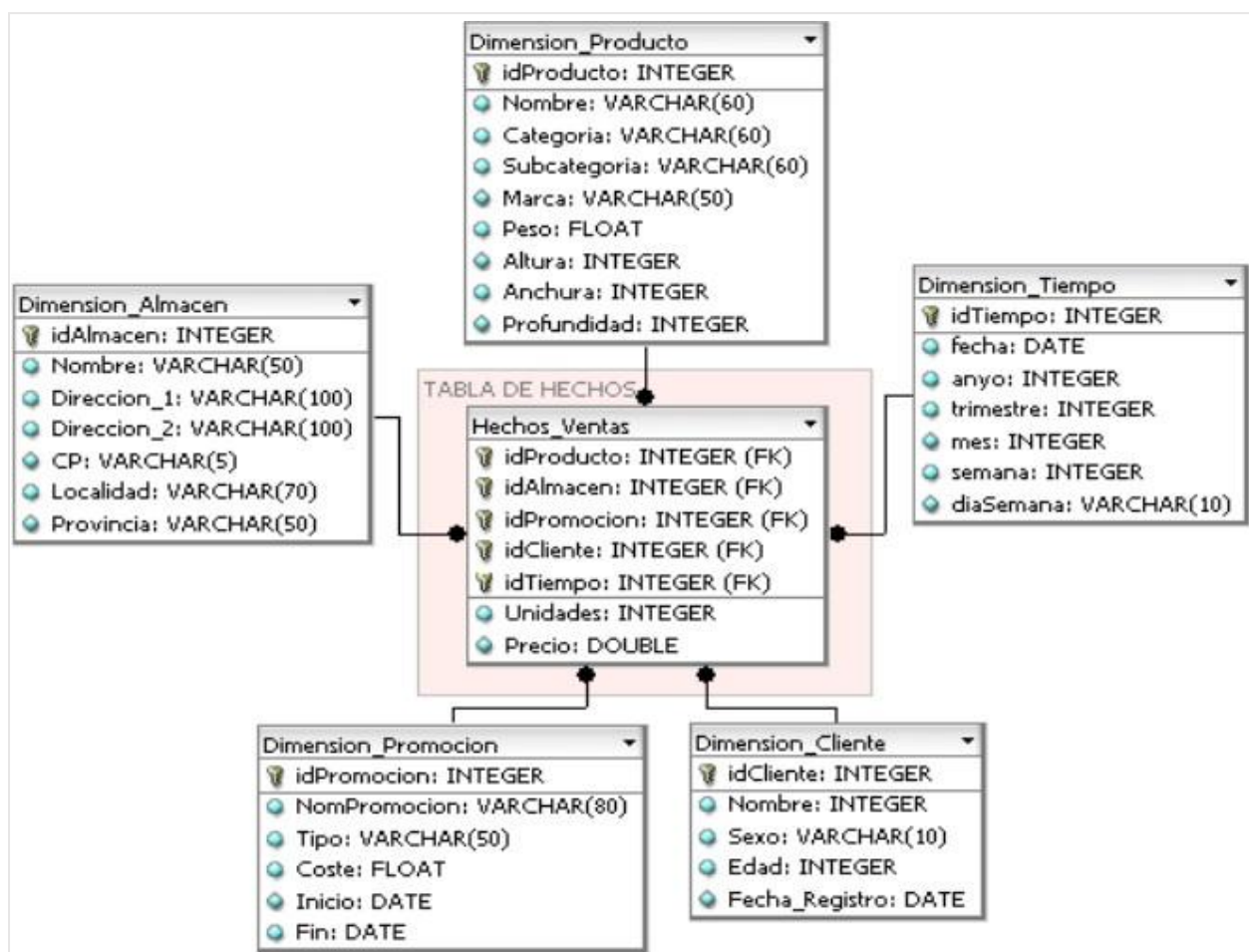
Por otro lado, las dimensiones nos indican los diferentes puntos de vista desde los que podemos analizar la información. Son usadas para seleccionar y agregar datos a un cierto nivel de detalle

deseado. Hablar de dimensiones, es hablar de jerarquías (las cuales refieren un nivel de profundidad en la información); permitiendo analizar la información de forma agrupada (ventas por año) o al detalle (ventas diarias). Ejemplo: La dimensión tiempo está formada por la jerarquía “Por Trimestre”, que tiene los niveles “Año > Trimestre > Mes > Día”.

En el modelado dimensional, podemos encontrar que los indicadores se encuentran en las Fact Tables, en español tabla de hechos; esta tabla se encuentra rodeada de dimensiones formando un esquema básico de Estrella. Por ejemplo la tabla de hechos “Hechos Venta”, contiene las métricas de unidades y precio; y se encuentra rodeada de las dimensiones las cuales agregan el nivel de detalle de la consulta (ver Figura 13) (Rodriguez Alfaro , 2021).

Figura 13

Ejemplo de Modelo Estrella



Nota. Fuente: Adaptado de (Rodriguez Alfaro , 2021)

La combinación de Reporting y Business Intelligence proporcionan información necesaria para ayudar a potenciar el rendimiento empresarial de las organizaciones, y así facilitar la toma de decisiones para mejorar los resultados de negocio a lo largo del tiempo. El Reporting contribuye con la generación de informes que posibilitan a los usuarios hacer seguimiento del negocio de forma sencilla, usualmente utilizados por usuarios gerenciales (Novoa Torres, Bermudez Huérfano, & Zamora Carrillo, Diciembre 2019).

En tal sentido, el Reporting debe permitir flexibilidad de acceso para elaborar informes y listados no demasiado complejos con los datos más usuales y más sencillos de analizar, tanto de manera agregada como detallada de la información. Al respecto, se recomienda el mantenimiento de al menos tres niveles de dificultad (Fernandez, 2021):

- Los usuarios poco expertos podrán solicitar la ejecución de informes o consultas predefinidas según unos parámetros predeterminados.
- Los usuarios con cierta experiencia podrán generar consultas flexibles mediante una aplicación que proporcione una interfaz gráfica de ayuda.
- Los usuarios altamente experimentados podrán escribir, total o parcialmente, la consulta en un lenguaje de interrogación de datos.

Hay una extensa gama de herramientas en el mercado para cumplir esta funcionalidad sobre entornos de tipo Data Warehouse, por lo que se puede elegir el software más adecuado para cada problemática empresarial concreta.

Marco Conceptual

Business Intelligence (BI): La inteligencia de negocios o business intelligence (BI) es el conjunto de procesos, aplicaciones y tecnologías que facilitan la obtención rápida y sencilla de datos provenientes de los sistemas de gestión empresarial para su análisis.

Data Warehouse (DW): Es un almacén o repositorio de datos que integra información de diferentes fuentes (base de datos, archivos de texto, hojas de cálculo, etc.) y permite un análisis para la toma de decisiones, un Data Warehouse se define un conjunto integrado de datamarts.

Data mart: Es un conjunto de datos estructurados de una forma que facilite su posterior análisis. Una data mart contiene la información referente a un área, un tema o una función en particular, con datos relevantes que provienen de las diferentes aplicaciones operacionales.

ETL (extract - transformation - load): Significa en español extracción, transformación y carga de los datos. Permite extraer datos de varias fuentes, transforma datos desde optimizado para transacciones a optimizado para informes y análisis.

Kimball: Metodología que determina que una data warehouse es la unión de todos los datamarts de una organización, sigue un enfoque de abajo hacia arriba (bottom –up) para almacenamiento de datos (DW) diseño de arquitectura en el que los mercados de datos se forman primero en función de los requisitos comerciales.

OLAP (On Line Analytical Processing): La funcionalidad de los sistemas OLAP se caracteriza por ser un análisis multidimensional de datos corporativos, que soportan los análisis del usuario y posibilidades de navegación, seleccionando la información a obtener.

Métrica: Es una medida cuantitativa del nivel en que un proceso posee cierto atributo, o propiedad medible; y, a su vez, una medida es el número asignado a dicho atributo como resultado de la aplicación de un procedimiento de medición. Por ejemplo, se puede tener métricas relacionadas a costos, plazos, entregables, etc.

OLTP: OLTP es la sigla en inglés de Procesamiento de Transacciones En Línea (On Line Transaction Processing). Es un tipo de procesamiento que facilita y administra aplicaciones transaccionales, usualmente para entrada de datos y recuperación y procesamiento de transacciones.

CAPÍTULO 3: REVISIÓN DE LA LITERATURA

Después de realizar una búsqueda en base de datos digitales como Science Direct, Google Académico y Scielo sobre artículos relacionados al tema de investigación y luego de una lectura exploratoria del contenido de cada documento, se ha seleccionado algunos trabajos que presentan propuestas sobre indicadores en el rubro de construcción, metodologías y herramientas aplicadas de manera exitosa en soluciones de inteligencia de negocios apoyada en indicadores de rendimiento.

Artículos Revisados

Artículo 1 (Holescher & Mortimer, 2018) Using Tableau to visualize data and drive decision-making

En el paper se enfatiza la importancia del proceso analítico de datos mediante el uso de herramientas de visualización de datos, con el objetivo de ayudar a los lectores a comprender los datos, así como el proceso de transformación de los mismos para mejorar el proceso de toma de decisiones. La investigación se centra en los resultados obtenidos de varios grupos de usuarios tras realizar un caso práctico con la información del proceso de ventas de una microempresa del sector producción.

El caso práctico consiste inicialmente en realizar la conexión a una fuente de datos en Access, la cual contiene información y transacciones de los últimos 6 meses de ventas de un pequeño fabricante de helados llamado Ivana. El software de visualización de datos utilizado es Tableau, con el cuál se realiza un análisis exploratorio para así identificar tendencias clave en los datos y así preparar la información para mejorar la toma de decisiones en dicho negocio. El caso fue probado en un campus residencial de una universidad pública en el Medio Oeste, para probar su eficacia se aplicó en tres secciones de un curso de sistemas de información contable de pregrado durante la primavera de 2017 y semestres de otoño de 2017. Estas secciones, en general, no tienen estudiantes familiarizados con el software de visualización, por lo cual se opta por realizar la primera mitad del caso guiado con instrucciones paso a paso con el fin de familiarizar al usuario con el uso; para luego la segunda mitad lograr el desenvolvimiento de cada uno.

Los resultados obtenidos en la primera mitad, de forma guiada, muestran que solo el solo el 5,94% de los estudiantes no cumple con las expectativas de aplicar las técnicas de análisis de datos explicadas y responder preguntas relacionadas a la operación comercial del negocio. Por último, en la segunda mitad del caso se obtuvo que menos del 8,64% de los estudiantes no cumplen las expectativas de preparar un informe profesional que permita responder preguntas del negocio y así ayudar a la toma de decisiones. En general, luego de aplicar encuestas, los resultados muestran que los usuarios corroboran la eficacia de Tableau ya que utilizan una variedad de habilidades, comprenden mejor como analizar los datos y el impacto en la toma de decisiones.

Artículo 2: (Parama Fadli Kurnia, 2018) Business Intelligence Model to Analyze Social Media Information

El paper parte de la importancia que hoy en día tienen las redes sociales, ya que estas se han vuelto un canal de comunicación digital permitiéndonos compartir documentos, socializar e incluso crear comunidades. Esta investigación tiene como propósito crear un tablero basado en inteligencia de negocios para monitorear el desempeño de las publicaciones en redes sociales como Facebook y Twitter. El rendimiento de las publicaciones está relacionado al número de interacciones que generan como: las reacciones, comentarios, publicaciones, etc.

En general, la solución involucra los siguientes pasos: la recolección de datos, el análisis de datos, el diseño del almacén de datos, el diseño de inteligencia empresarial, la clasificación de textos y los métodos de evaluación. Enfocándonos en la parte de construcción del almacén de datos, propone diseñar un Datawarehouse basado en el método Kimball, el cual considera 4 fases: seleccionar el proceso de negocio, declarar la granularidad de la información, identificar dimensiones e identificar tablas de hechos. El almacén de datos se realiza mediante el uso de un modelo de esquema en estrella que consta de una tabla de dimensiones y una tabla de hechos. El esquema de estrella en los datos de Facebook consta de fact_page_topic como tablas de hechos y dim_page, dim_topics, dim_time como tablas de dimensiones. En este esquema puede mostrar la cantidad de publicaciones, comentarios, compartir y me gusta en artículos que se agrupan nuevamente en función de temas relacionados con los artículos publicados en Facebook. El dashboard se construye basado en el framework CodeIgniter el cual proveerá data insights, como temas más populares, usuarios activos los cuales podrían transformarse en cuentas publicitarias.

Tras la implementación del sistema basado en inteligencia de negocios, se concluye que el Datawarehouse creado como fuente de datos para el sistema BI puede ejecutar los cálculos y resúmenes de datos automáticamente, por lo que ya no es necesario realizar cálculos manualmente. Además, la aplicación Business Intelligence es muy útil para monitorear el desempeño de las noticias publicadas en las redes sociales tanto desde la cuenta interna de la empresa como de la competencia en tiempo real, por lo que ya no es necesario visitar una a una las cuentas en redes sociales.

Artículo 3: (Santos, y otros, 2017) A Big Data system supporting Bosch Braga Industry 4.0 strategy

El paper enfatiza la nueva problemática de las grandes empresas producto del constante desarrollo de las Tecnologías de la Información y las Comunicaciones, lo que viene impulsando la cuarta revolución industrial. Hoy en día, se obtienen datos de diversas fuentes, como por ejemplo de: personas, dispositivos, sensores, entre otros tornando a los datos como un desafío clave. En tal sentido los sistemas de Big Data se tornan necesarios para recolectar, integrar y almacenar los datos con el objetivo de disponibilizar información oportuna en toma de decisiones.

Por lo tanto, este documento aborda este desafío clave, proponiendo e implementando una arquitectura de Big Data Analytics en la organización multinacional (Bosch Car Multimedia - Braga). Como parte de la solución, se extraen los datos almacenados en el ERP SAP, los cuáles pasaran por el proceso ETL mediante la herramienta Talend. La información resultante de este proceso debido al volumen y variedad se almacenará en HDFS y Hive; esta información será accedida en tiempo real con el respaldo analítico que brinda Presto al ejecutar las consultas en lenguaje SQL y finalmente Tableau permite la visualización de los datos obtenidos.

Tras la implementación del sistema de Big Data, se mostró como diversas tecnologías se complementan entre sí con el objetivo de apoyar el proceso de toma de decisiones. Se validó que, en los dos procesos de negocio seleccionados, el de calidad y el de costos, los datos se modelaron, limpiaron, transformaron de una forma correcta, para posteriormente ser entregados a los componentes de almacenamiento de datos, capaces de lidiar con históricos datos, lo que permite una respuesta de la empresa Bosch a tiempo y de calidad ante análisis decisorios.

Artículo 4: (Cheung, Suen, & Cheung, 2004) PPMS: a Web-based construction Project Performance Monitoring System

Este paper describe el desarrollo de un Sistema de Monitoreo del Desempeño de Proyectos de Construcción con el propósito de ayudar a los gerentes de proyecto a ejercer el control en los proyectos de construcción. Con la ayuda de un panel de especialistas en administración de proyectos, se identifican las siguientes categorías de medidas de desempeño del proyecto para su inclusión en el sistema de monitoreo: Personas, Costo, Tiempo, Calidad, Seguridad

y Salud, Medio Ambiente, Satisfacción del Cliente y Comunicación. Para cada una de las categorías de medidas de desempeño, también se establecen indicadores de desempeño y sus mediciones.

Como parte de la solución, el proceso de seguimiento se automatiza mediante el uso de Tecnología de base de datos y World Wide Web. La recopilación y difusión de datos están igualmente automatizadas. Los componentes computacionales utilizados son: PHP, base de datos backend MySQL y un módulo de presentación gráfica. PHP fue seleccionado debido al bajo consumo de recursos y especialmente por ser preciso para los desarrollos web. MySQL complementa a PHP permitiendo gestionar las relaciones de información. El sistema de monitoreo del desempeño de los proyectos operará mediante el uso de una interfaz, el cual básicamente tiene dos grandes funciones: (1) permite recopilar información relacionada a los indicadores seleccionados de los proyectos (2) el sistema tiene la capacidad de mostrar información gráfica.

Tras la implementación del sistema, se concluye que ayuda a los directores de proyectos y administradores en la evaluación del desempeño del proyecto de una manera oportuna. Las categorías de medida se pueden agregar y eliminar según los objetivos del proyecto. Más importante aún, el uso del sistema permite que los usuarios comparen y presenten gráficos fáciles de usar y entender, permitiendo comparar las opiniones perceptivas de los distintos proyectos y permitiéndoles identificar áreas donde el desempeño este fallando y así abordar la solución del mismo con prontitud.

Artículo 5: (Ling, 2004) How project managers can better control the performance of design-build projects

El propósito de este estudio es identificar los factores clave que afectan 11 áreas del desempeño del proyecto de diseño y construcción. Usando datos de 42 proyectos públicos y privados, este estudio encuentra que las características del contratista son los determinantes clave que afectan más número de métricas de rendimiento. La variable más importante es el historial del contratista para completar proyectos dentro del presupuesto, en horario y a un nivel aceptable de calidad. Otros determinantes importantes son la capacidad de los contratistas y la adecuación de sus recursos.

Para llevar a cabo este estudio, primero se realizó una revisión de la literatura, en tal sentido se obtuvo un listado de 65 factores potenciales que afectan el desempeño del proyecto de las 11 áreas más relevantes de proyectos de construcción. El método de investigación tuvo como objetivo identificar factores importantes que afectan el rendimiento del proyecto, apoyado en un cuestionario de estudio de caso retrospectivo, el cual fue diseñado en base a los factores y métricas de desempeño descubiertas a partir de la revisión de la literatura. Este cuestionario pedía al encuestado que proporcionara información específica sobre el proyecto final e inicial costos, duraciones finales y planificadas, superficie bruta y fecha de cierre de licitación. Los encuestados también tenían que indicar la calidad del proyecto y el nivel de satisfacción del propietario utilizando escalas.

Los datos recopilados en el cuestionario fueron sometidos a diversos análisis de correlación y varianzas, con lo cual se obtienen las categorías más resaltantes a monitorear en los proyectos de construcción: gestión de costos, gestión del tiempo, gestión de calidad. Otro hallazgo significativo de este estudio es que la mayoría de los factores que afectan el éxito de proyectos de construcción son variables relacionadas con los contratistas, es por ello que recomiendan que los gerentes de proyecto presten mucha atención a la selección de contratistas para asegurarse de que los diferentes aspectos del desempeño del proyecto se cumplan.

Artículo 6: (Suganda Girsang, y otros, 2018) Business Intelligence for Construction Company Acknowledgement Reporting System

El paper enfatiza el rol que lleva a cabo una de las asociaciones de empresas indonesias de servicios de construcción, la cual se encarga de registrar a todas las empresas de servicios de construcción para que trabajen de acuerdo con la ley y la constitución del país. Debido a que el trabajo requiere brindar informes detallados al gobierno, a inversionistas nacionales o internacionales, y en general a cualquier institución que necesite la información, surge el problema de crear consultas para cada fuente de datos desde tablas en MySQL, las cuales se exportan manualmente en archivos Excel para crear los informes solicitados; lo que causa una gran labor de análisis y demoras de los mismos.

La solución recomendada es crear un Data Warehouse, el cual posee almacenamiento estructurado y puede analizar datos con OLAP que tiene capacidad para procesar y visualizar datos a alta velocidad. Para construir el almacén de datos, el diseño del mismo se basó en el método de Kimball, por lo que esta asociación tendrá la capacidad de generar informes de acuerdo con las necesidades de cada instancia. En primera instancia se realizó la recopilación de datos relevantes de la base de datos transaccional MySQL, la cual se analiza para producir medidas y cálculos a utilizarse en el análisis OLAP y en indicadores a mostrarse en el dashboard. Luego de ello se realizó el modelado dimensional de la tabla de hechos y dimensiones; ya listo el modelo de estrella, la data proveniente de MySQL pasará por el proceso ETL utilizándose la herramienta Pentaho con lo cual se podrá cargar datos al Datawarehouse. Finalmente se realizó la creación de cubos, los cuales fueron conectados a Qlik Sense para generar los reportes requeridos.

Se validó que al usar los dashboards y los informes, el proceso de análisis de datos se vuelve mucho más fácil de realizar, ello debido a que el Data Warehouse brinda la capacidad de procesar datos de hechos con múltiples dimensiones en la base de datos OLAP. El resultado del procesamiento de datos se puede presentar con muchos tipos de diagramas, tablas y cuadros de mando. La combinación de informes y cuadros de mando facilitó a los analistas empresariales la toma de decisiones adecuadas y rápidas para resolver los problemas.

Artículo 7: (Boulila, Al-kmali, Farid, & Mugahed, 2018) A business intelligence based solution to support academic affairs: case of Taibah University

El artículo recalca el importante desafío que muchas universidades experimentan al tomar decisiones estratégicas con el objetivo de mejorar la calidad educativa. Estas universidades recopilan datos de estudiantes, tanto en el proceso de admisión, en el transcurso y la salida de su vida universitaria; por ello el objetivo es desarrollar una herramienta que proporciona información rápida y nueva para respaldar el proceso de los asuntos académicos. Sin embargo, debido a la gran cantidad de datos que provienen, generalmente, de múltiples fuentes de datos, la necesidad de un sistema automatizado se vuelve más que necesaria.

Como parte de la solución se propone realizar un sistema basado en BI; en tal sentido, se llevaron a cabo tres tareas esenciales: (1) recopilación de datos de diferentes fuentes utilizando el proceso ETL (Extraer, Transformar y Cargar), (2) propuesta de una solución multidimensional que describa los procesos académicos y (3) visualización de resultados a través de un conjunto de paneles e informes. Los experimentos se realizaron utilizando las herramientas de datos de SQL Server, las cuales son utilizadas en el siguiente orden: SQL Server Integration Services (SSIS), SQL Server Analysis Services y SQL Server Reporting Services y Power BI. La solución proporcionó muchos indicadores estadísticos y predictivos necesarios en las tareas académicas.

Los resultados obtenidos tras la implementación de BI en la Universidad ayudaron a los administradores a comprender mejor diferentes operaciones relacionadas con el proceso académico; ya que se pudo disponibilizar datos históricos y predictivos relacionados a los indicadores seleccionados ayudando a comprender muchas acciones frecuentes, las cuales hasta dicho momento no contaban con justificaciones claras listadas a continuación: (1) Eliminación de cursos, (2) Desaprobación repetida de cursos sin conocer el motivo, (3) Demoras de estudiantes en el plan de estudio, (4) Indicadores para seleccionar cursos que permitan elevar la calificación del estudiante.

Resumen de Revisión de la Literatura

A continuación, se muestra un resumen de indicadores, metodologías y herramientas encontrados en la revisión de la literatura.

Tabla 2*Resumen de Revisión de Literatura*

N °	Autor(es)	Indicador	Herramienta	Metodología
1	(Holescher & Mortimer, 2018)	No especifica	Tableau	No especifica
2	(Parama Fadli Kurnia, 2018)	No especifica	CodeIgniter	Kimball
3	(Santos, y otros, 2017)	No especifica	Tableau	No especifica
4	(Cheung, Suen, & Cheung, 2004)	Personas, Costo, Tiempo, Calidad, Seguridad y Salud, Medio Ambiente, Satisfacción del Cliente y Comunicación	No especifica	No especifica
5	(Ling, 2004)	Costo, Tiempo, Calidad	No especifica	No especifica
6	(Suganda Girsang, y otros, 2018)	No especifica	Qlik Sense	Kimball
7	(Boulila, Al-kmal, Farid, & Mugahed, 2018)	No especifica	SQL Reporting Services / Power BI	Inmon

Nota. Fuente: Elaboración propia

En los que se refiere a los indicadores utilizados en proyectos de construcción se utilizan los provenientes de las categorías de: Personas, Costo, Tiempo, Calidad, Seguridad y Salud, Medio Ambiente, Satisfacción del Cliente y Comunicación; y Calidad. En lo que respecta a herramientas

de visualización de datos se observó el empleo de Tableau, Qlik Sense y Power BI. A nivel de Metodologías en el mundo de la Inteligencia de Negocios, sobresalen 2: Kimball e Inmon.

Indicadores de proyectos de Construcción

De la revisión de la literatura se ha identificado a dos autores: Cheung, Suen, & Cheung (2004) y Ling (2004). Ambos proponen indicadores de desempeño para proyectos de construcción. Estos indicadores se basan en el denominado triángulo de hierro (coste, plazo y calidad) propuesto por el doctor Martín Barnes en 1969. A continuación, se detallan las propuestas analizadas de los dos autores.

Cheung, Suen, & Cheung (2004)

Apoyados con especialistas de control de proyectos, miembros de la asociación de gestores de proyectos en Honk Kong, identificaron ocho categorías críticas de medida de rendimiento en la gestión de proyectos, las cuales son: personas, costes, tiempo, calidad, seguridad y salud, ambiente, satisfacción del cliente y comunicación. Comentan que las categorías de tiempo, costo, calidad; y seguridad y salud son la base de medición. A su vez incluyen medidas del medio ambiente y personas como respuesta al impacto que generan los temas medioambientales y la responsabilidad de que el trabajador se sienta a gusto en el proyecto. A continuación, se detalla el concepto referido a cada categoría propuesta.

Categoría 1 - Personas

La categoría personas se enfoca en el sentir de los miembros del Proyecto, acerca del rendimiento del Proyecto con respecto al tiempo, costo, calidad, entre otros en un periodo de tiempo.

Categoría 2 – Costo

La categoría costo se encarga de mostrar que tan bien el proyecto realiza los gastos en función al presupuesto, esto es de vital importancia ya que permite ajustar los recursos y evitar sobrecostos. Existen cuatro áreas que son relevantes para el monitoreo de costos dentro del proyecto: Pagos internos, Ordenes, Costos y Reclamaciones y Proyecciones de Pagos.

Categoría 3 – Tiempo

El monitoreo de la categoría tiempo es un desafío grande que tienen los gerentes, ya que se debe monitorear como el proyecto cumple los entregables y horarios planeados dentro del periodo de ejecución. Existen tres temas relevantes en los que se centra esta categoría: fechas críticas, hitos y tiempo de respuesta.

Categoría 4 – Calidad

La categoría Calidad apoya en asegurar que los proyectos logren los estándares de calidad detallados en el contrato, para ello se deben monitorear áreas como Control de Calidad, el reporte de inconformidades y tasas de rechazo en entregables.

Categoría 5 – Seguridad y Salud

La categoría seguridad y salud se enfoca en monitorear el cumplimiento de medidas de seguridad con el objetivo de prevenir y reducir siniestros y accidentes en el trabajo.

Categoría 6 – Medio Ambiente

La categoría medio ambiente es la destinada a monitorear las repercusiones causadas por las actividades de la construcción, como lo son el tema ambiental de contaminación y gestión de residuos.

Categoría 7 – Satisfacción del Cliente

Esta categoría está enfocada en revelar el grado de satisfacción del cliente relacionado al rendimiento del proyecto.

Categoría 8 – Comunicación

La categoría comunicación es utilizada para evaluar la eficacia de la comunicación entre los participantes del proyecto, considerando, por ejemplo: el número de reuniones sostenidas, el número de solicitudes de información, entre otros.

Ling (2004)

Enfatiza dos roles importantes en el monitoreo de proyectos de construcción, el primero es el cliente que solicita el proyecto y el segundo es al contratista quien realizará la construcción de la obra. Ambas partes apuntan a un resultado exitoso, por ello propone evaluar métricas de las siguientes categorías: Costo, Tiempo, Calidad y Propietarios.

Categoría 1 – Gestión de Costos

Los gerentes de proyectos se interesan en dos temas relacionados a la gestión de costos: (1) minimizar costos unitarios y (2) minimizar el crecimiento de costos. Estos contruirán a que el proyecto no sobrepase los costos estimados.

Categoría 2 – Gestión de Tiempo

Los gerentes de proyectos se interesan en tres temas relacionados a la gestión de tiempo: (1) maximizar la velocidad de la construcción, (2) maximizar la velocidad de entrega y (3) control del cronograma del proyecto.

Categoría 3 – Gestión de Calidad

Los gerentes de proyectos se interesan principalmente en dos temas relacionados a la gestión de calidad: (1) calidad de equipos y (2) personal con alto expertis técnico.

Categoría 4 – Gestión de Carga Administrativa

Los gerentes de proyecto deben minimizar la carga administrativa del cliente, por lo que en la búsqueda del mejor contratista deben considerar los siguientes temas: (1) contratistas con capacidad de diseño, (2) alto nivel de personal, (3) historial de entregas. Con ello se logrará minimizar la carga del cliente y del director del proyecto.

En conclusión, tomando como base ambas propuestas, para esta tesis se ha seleccionado las categorías de: Costo, Tiempo y Seguridad y Salud; ya que una de las prioridades de la gerencia de Control de Proyectos de Construcción es asegurar que los proyectos se lleven a cabo en el plazo estipulado y bajo los costos planificados y finalmente asegurar la seguridad y salud de sus

trabajadores. A continuación, se sintetizan los indicadores que serán considerados para cada una de las categorías de métricas seleccionadas:

Tabla 3

Categorías e indicadores de rendimiento en Proyectos de Construcción

Categoría de medidas de rendimiento	Indicadores de rendimiento
Costo	<i>Productividad de Equipos</i> Costo real de equipos Costo previsto de equipos <i>Costo del Proyecto</i> Costo actual Costo Previsto <i>Margen</i> Costo total del proyecto Venta total del proyecto
Tiempo	<i>Cumplimiento del Plan</i> Numero de asignaciones cumplidas Numero de asignaciones programadas <i>Avance</i> Porcentaje de avance real Porcentaje de avance planeado
Seguridad y Salud	<i>Estadísticas de accidentabilidad</i> Número de accidentes reportados Número de horas hombres perdidas

Herramientas de Inteligencia de Negocios

De la revisión de la literatura, se encontró que Tableau, Power BI y Qlik han sido utilizados como herramientas de visualización en muchos proyectos de Inteligencia de Negocios e inclusive Big Data; ello debido a una buena razón, acorde con Gartner, estas 3 herramientas se encuentran en el cuadrante superior derecho de herramientas líderes. En la figura 14 se puede observar el análisis

ampliado en Business Intelligence conocido como el Cuadrante Mágico para Plataformas de Inteligencia de Negocios propuesto por Gartner (2020). Estas herramientas tienen como objetivo común el crear aplicaciones que ayuden a las organizaciones en el proceso de aprendizaje del negocio.

Figura 14

Cuadrante Mágico de Gartner para Analytics y Plataformas de Inteligencia de Negocios



Nota. Fuente: Adaptado de Gartner (febrero 2020)

A continuación, se describen de manera resumida cada una de estas herramientas.

a) Tableau:

Tableau es una plataforma de análisis visual que transforma la manera en la que se utilizan los datos para resolver problemas. Esta plataforma se fundó en 2003 como resultado de un proyecto de ciencias de la computación de la Universidad de Standfort desarrollado por Chris Istulte, Christian Chabot y Pat Hanraham (Tableau, 2021).

La plataforma Tableau está basada en dos tecnologías innovadoras: VizQL, esta tecnología permite la creación de visualizaciones complejas tan solo con un clic del mouse; y el motor de datos, el cual permite procesar información de millones de filas en tan solo un par de segundos. Con lo que respecta, a orígenes de datos ofrece una amplia conexión a diversas fuentes datos estructuradas y no estructuradas a los cuales se puede acceder en pocos minutos debido a conexiones predefinidas; adicional a ello permite realizar un intercambio fácil de datos entre sistemas web y móviles. Finalmente, con lo que respecta al aspecto de visualización, posee una gran variedad de tipos gráficos sumamente llamativos y muy interactivos(Bocevska, Savoska, & Milevski, 2017).

Tableau es una herramienta que permite conectar y explotar los datos en pocos minutos creando visualizaciones de alto impacto de manera intuitiva al arrastrar componentes y al customizarlos. Los tableros o informes construidos, permiten que el usuario genere una mejor interrelación con los datos, facilitando el proceso de análisis y permitiendo responder y generar preguntas relacionadas al negocio de aplicación. Tableau posee el componente Prep diseñado para preparar la data fácilmente, puede ser utilizado para combinar, limpiar, agregar la data entre otras funciones apoyando al proceso ETL. (Bocevska, Savoska, & Milevski, 2017)

Tableau hoy es utilizado por decenas de miles de usuarios en más de 100 países, incluidos: Amazon, Apple, Walmart, Starbucks, Pfizer, Coca-Cola, Mazda. Un gran número de empresas peruanas también la utilizan, como: Telefónica, Ministerio de Salud, El Comercio, Ministerio del Trabajo entre otras.

A nivel internacional, la empresa DataSift ha implementado Tableau integrada a su plataforma Datasift plataforma líder de datos, la cual proporciona acceso abierto a datos de varias redes sociales (Facebook, Twitter, Instagram, etc.). Como sabemos, hoy en día las

b) Power BI:

Power BI es una plataforma SaaS basada en la nube que permite analizar datos y compartir información a través de la web. Power BI, es una combinación de la tecnología de base de datos más utilizada (SQL Server) y el software de datos Excel; es decir, Power BI es un resultado de los últimos lanzamientos de Microsoft, lanzado por primera vez en el año 2015. (Power Pro Consulting, 2021)

Power BI está integrado con una amplia gama de fuentes de datos, que incluyen la nube y soluciones locales. Con una amplia variedad de fuentes de datos, puede conectarse fácilmente a soluciones SaaS, datos locales en SQL Server Analysis Services, servicios Azure, archivos Excel, entre otros. La tecnología subyacente de Power BI es un motor de análisis en memoria y una base de datos en formato columnar que admite estructuras de almacenamiento de datos tabulares que utiliza Power Pívor. Esto logra un equilibrio entre rendimiento y facilidad de uso (en comparación con los cubos tridimensionales que requieren consultas de lenguajes más complejos como como MDX (expresiones multidimensionales). (Boulila, Al-kmal, Farid, & Mugahed, 2018)

Una de las ventajas de Power BI es que proporciona acceso a todos los datos de una única ubicación, independientemente de dónde residan los datos. Este enfoque híbrido ofrece una serie de beneficios: (1) tiempo rápido para obtener información con conexiones directas; (2) conectividad segura y en tiempo real con los datos y (3) escalable, ya que los datos migran de soluciones locales a la nube. (Bocevska, Savoska, & Milevski, 2017)

Microsoft (2021) en su página oficial reconoce que muchas organizaciones utilizan las soluciones brindadas por Power BI para resolver tareas analíticas con el objetivo de preparar estrategias que apoyen la toma de decisiones. Algunas de las organizaciones más conocidas son las siguientes: Hewlett Packard, Nestlé, Adobe, Nokia entre otros. La figura 16 muestra el flujo de trabajo aplicado en la implementación de soluciones con Power BI, como se puede observar la conexión a fuentes de datos se realiza desde Power BI Desktop, para la publicación de proyectos se utiliza Power BI Service y con ello se logra alojar los proyectos en la nube organizativa de tal forma que los interesados acceden al contenido.

Figura 16

Flujo de trabajo de la Plataforma Power BI



Nota. Fuente: Adaptado de (Microsoft, 2021)

c) QlikView:

QlikView es una plataforma analítica de datos desarrollada por la empresa sueca Qlik Tech Inc. en el año 1995.

Uno de los principios clave de la herramienta QlikView es el uso de una amplia gama de fuentes de datos estructuradas y no estructuradas y las asociaciones comunes entre ellas. QlikView permite conexiones ODBC / OLE a través de un sencillo asistente para extraer rápidamente datos de los sistemas de origen. Esta plataforma ha sido desarrollada para ser intuitiva y fácil de usar. El acceso instantáneo a los datos, la flexibilidad en el uso y la actualización en tiempo real son características que hacen que el uso de esta herramienta sea eficaz. QlikView utiliza consultas SQL estándar para leer datos y procedimientos almacenados, por lo que el usuario debe estar familiarizado con este lenguaje. (Shukla & Dhir, 2016)

Una de las grandes ventajas de Qlikview es la carga de datos en memoria, con ello se logra disponibilidad para realizar búsquedas asociativas instantáneas y análisis en tiempo real al hacer un par de clics. Esta ventaja se logra gracias al motor de búsqueda asociativa, el cual

es una memoria de próxima generación que ofrece la arquitectura QlikView la cual permite cargar datos en tiempo real comprimiendo los datos hasta un 10% del tamaño original, eliminando los problemas y la complejidad de herramientas BI tradicionales basadas en procesamiento en disco con datos estáticos y empaquetados. (Bocevska, Savoska, & Milevski, 2017)

A modo de conclusión, se puede mencionar que estas herramientas brindan aspectos comunes y también aspectos particularidades los cuales se resumen en la Tabla 4. En la que se observa que las tres son soluciones que dominan el mercado. La diferencia entre ellas cada vez se está haciendo más pequeña, ello debido al desarrollo y la evolución de estas plataformas garantizan mejores posibilidades con los datos. La elección de una de estas depende de las necesidades y ambiciones que se tenga con la herramienta.

Para el presente trabajo de investigación se seleccionó Tableau, porque tiene mejor nivel de usabilidad, ya que esta herramienta posee la ventaja de arrastrar y soltar los componentes con el objetivo de crear y desarrollar fácilmente sus tableros de control, reportes, entre otros, lo anteriormente comentado es de vital importancia ya que se busca que los usuarios auto gestionen la creación de reportes. Adicional a ello es de gran ayuda los recursos de ayuda y la comunidad que tiene Tableau, ya que permite la colaboración entre usuarios.

Tabla 4*Cuadro comparativo de herramientas de visualización BI*

Características	Tableau	Power BI	Qlik
Adaptabilidad del Usuario	Sin experiencia de programación previa	Sin conocimiento específico	Requiere conocimiento de creación de Scripts Usuarios competentes
Productos	Tableau Desktop, Tableau Server y Tableau Online.	Power BI Desktop, Power BI Service y Power BI Mobile	Qlik Sense, Qlik Analytics Platform, Qlik Core o Qlikview
Visualización	Facilidad de exploración visual	Potencia en reporting y cuadros de mando.	Ofrece plantillas de informes
Aprendizaje	Cuenta con una gran cantidad de recursos	Microsoft cuenta con una gran red de partners que brindan recursos de aprendizaje	Ofrece recursos orientados a perfiles con conocimiento de manipulación de datos
Precio	7 veces más caro que Power BI	Más asequible que sus competidores	Según necesidades ofrece variedad en modelo de precios.
Alojamiento	On Premise y Cloud	En la nube (SaaS) o local (On Premise).	EN la nube, On Premise o API abierta

Nota. Fuente: Elaboración propia**Metodologías de Inteligencia de Negocios**

De la revisión de literatura se pudo observar que existen dos metodologías BI referentes en el modelado del datawarehouse, estas son la de Inmon y Kimball; ambos proponen diversos enfoques de creación y modelado de la estructura de datos multidimensional. Por ello a continuación,

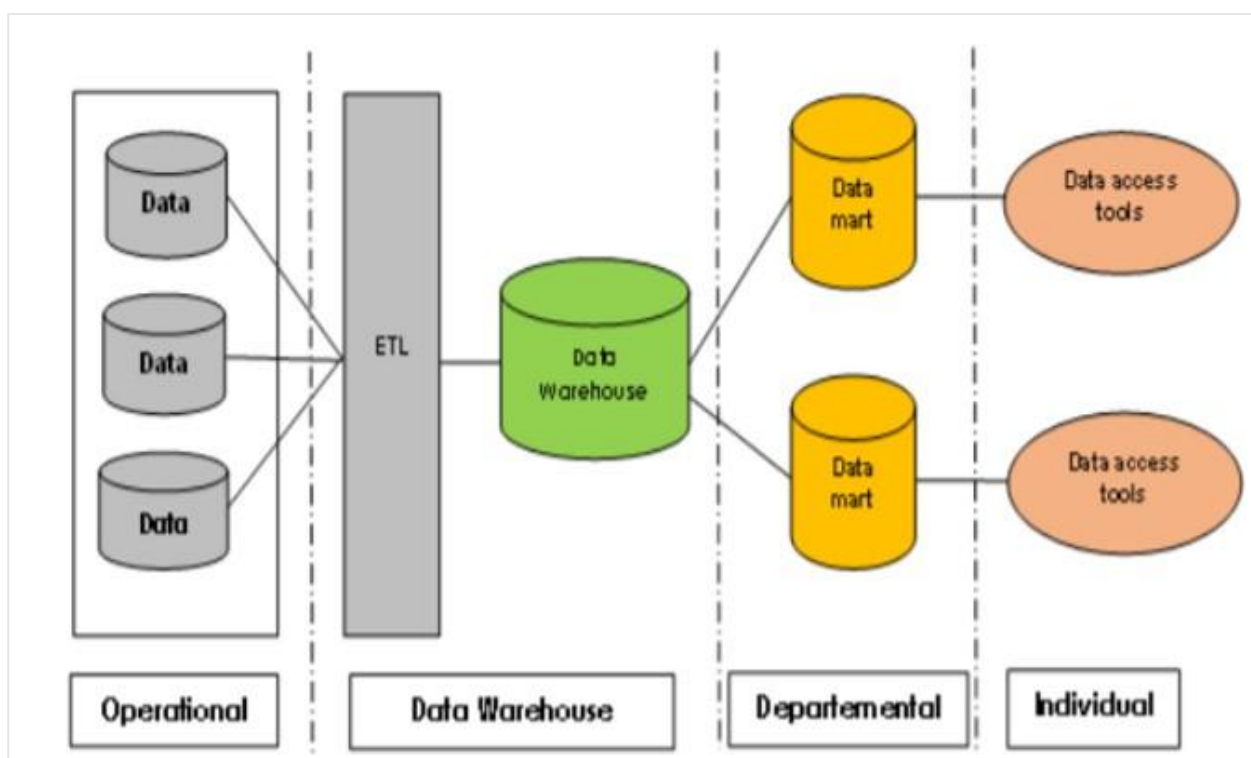
primero se presentan cada propuesta por separado para luego mostrar un comparativo de los aspectos más relevantes de ambas metodologías.

a) Inmon

Desarrollado por Bill Inmon en los 90's. Su enfoque se denomina “basado en datos”; ya que los datos de la empresa se cargan a priori sin conocer los requisitos del usuario. La arquitectura del almacén de datos de Inmon mostrado en la figura 17, incluye todos los sistemas de información de la empresa con sus bases de datos en lugar de considerar solo fragmentos de información, conocida como top down; propone crear el datawarehouse y que los data marts se segmentan en base a este (Labioud & Yessad, 2016)

Figura 17

Arquitectura Datawarehouse propuesta por Bill Inmon



Nota. Fuente: Adaptado de (Labioud & Yessad, 2016)

Inmon divide el entorno de las bases de datos de la empresa en cuatro niveles que incluyen: operacional, atómico (a nivel de dato), departamental (data marts) e individual. La información

operativa, se transforma y se cargan en los datos atómicos apoyada en los procesos ETL. Estos datos atómicos, según requisitos de los departamentos de la empresa modelan los datamarts orientados a temáticas específicas y siempre consistentes por la atomicidad del depósito de datos (Labiod & Yessad, 2016).

Inmon considera que un datawarehouse y un datamart están físicamente separados. El almacén de datos tiene su propia existencia física y está orientada al almacenamiento, trazabilidad y escalabilidad en respuesta a nuevos requisitos. Por otro lado, el datamart tiene su propia existencia física, su estructura está orientada al rendimiento de nuevos requisitos por parte del usuario (Labiod & Yessad, 2016)

b) Kimball:

Ralph Kimball creó su enfoque en los años 90 ofreciendo una nueva arquitectura, una nueva visión y un modelado innovador de un almacén de datos. Su enfoque se opone a la planteado por Inmon, al involucrar a los usuarios finales en las primeras etapas del proyecto. Considera que el data warehouse puede verse como un conjunto de mercados de datos consistentes basados en dimensiones; su enfoque se conoce como “bottom-up”, lo que significa que parte de un conjunto de datamarts y posteriormente se integra en un data warehouse centralizado (Rivadera, 2010). En una vista multidimensional de datos se destaca las diferentes perspectivas de análisis mostradas en la figura 18.

Rivadera (2010) detalla que la metodología se basa en lo que Kimball denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Este ciclo de vida del proyecto de DW, está basado en cuatro principios básicos:

Centrarse en el negocio: Referido a involucrar a usuarios conocedores del negocio en etapas iniciales, en conjunto con ellos se definirán los requerimientos del negocio y su valor asociado apoyando en el entendimiento y análisis de los implementadores.

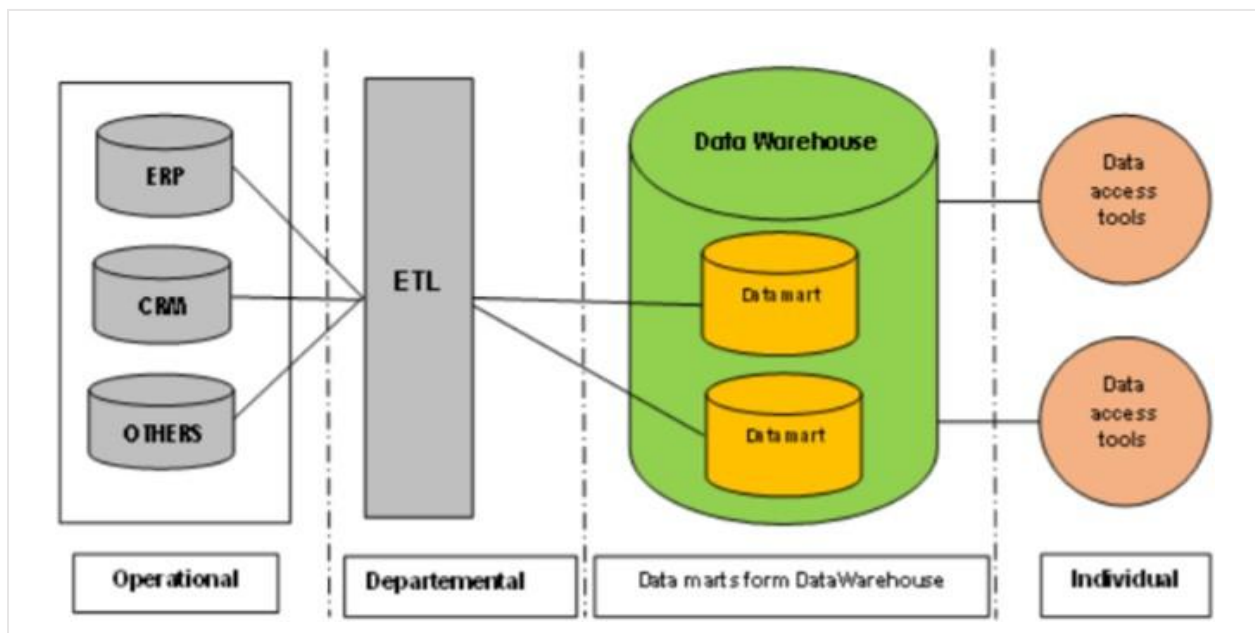
Infraestructura de información adecuada: El objetivo es diseñar una base de información única, integrada, fácil de usar, de alto rendimiento donde se reflejará la amplia gama de requerimientos de negocio identificados en la empresa.

Entregas significativas: El objetivo es crear el almacén de datos de forma incremental y significativa en plazos de 6 a 12 meses. La significancia de los mismos se evalúa en base al valor de cada elemento identificado en el negocio, con ello se determina el orden de aplicación de los incrementos (muy similar a metodologías ágiles en construcción de software).

Solución completa: Se debe proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios, referido a la entrega de: herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

Figura 18

Arquitectura Data Warehouse propuesta por Ralph Kimball



Nota. Fuente: Adaptado de (Labiod & Yessad, 2016)

La Tabla 5 muestra una comparación de las dos propuestas, la cual considera características como: la arquitectura del desarrollo del almacén de datos, la metodología, la complejidad general del enfoque, el costo y el tiempo de implementación del almacén de datos.

Tabla 5*Comparación de los enfoques acorde a la metodología y la arquitectura*

	Inmon	Kimball
Estructura de la arquitectura	Data a nivel atómico alimenta a los data marts.	El conjunto de data marts constituye el data warehouse.
Complejidad	Bastante complejo.	Simple.
Desarrollo de la metodología	Inspirada en la metodología espiral.	Basada en las 4 etapas del ciclo de vida del negocio.
Costos de despliegue	Altos costos iniciales y los siguientes más bajos.	Bajo costo inicial y en los siguientes cuesta un poco más.
Tiempo de despliegue	Duración larga	Duración corta

Nota. Fuente: Elaboración propia**Benchmarking de Metodologías de Inteligencia de Negocios**

Para realizar el benchmarking de Metodologías, se han seleccionado ciertos criterios que el autor identificó de la revisión de literatura, principalmente del paper realizado por Labiod & Yessad (2016) en donde se contrasta en el desarrollo de dicha investigación los siguientes criterios enunciados a continuación.

a) Criterios de Comparación

Al examinar las 3 metodologías de inteligencia de negocios, podemos notar que los enfoques son muy diferentes, pero con el mismo propósito de proveer un marco de desarrollo en el modelado y construcción del datawarehouse, es así que se identificaron los siguientes criterios de evaluación.

Usuarios finales: Se busca que los usuarios finales pueden gestionar sus propias necesidades de información, y que en base al repositorio puedan explotar la información requerida.

Implementación: Es importante que la metodología sea entendible y fácil de usar.

Rendimiento: Se busca que el tiempo de respuesta en la ejecución de consultas sea el menor posible.

Estructura de la Arquitectura: Se requiere que la metodología permita crear un modelado puntual y que permite ir añadiendo más modelos para construir el repositorio unificado.

Tiempo de despliegue: Se requiere que la metodología nos permita hacer el despliegue de manera rápida y ordenada.

Acorde a nuestra necesidad de información, los factores se ponderan de la siguiente forma:

Tabla 6

Ponderación de factores - Benchmarking de metodologías

Factor	Peso	Justificación
Usuarios finales	0.3	Autogestión de información
Implementación	0.1	Facilidad de implementación
Rendimiento	0.3	Consultas rápidas
Estructura arquitectura	0.2	Modelado incremental
Tiempo despliegue	0.1	Rapidez en despliegue

Nota. Fuente: Elaboración propia

b) Evaluación

Para completar el Benchmarking de Metodologías se realizó lo siguiente: La Tabla 7 muestra el valor de puntuación asociado a la afinidad que tendrá cada metodología con los criterios, por ejemplo: Si la metodología No tiene afinidad con un criterio, se le puntuará como 0; si tiene Bajo nivel de afinidad, puntuará como 1; si tiene Medio nivel de afinidad, puntuará como 2 y si tiene Alto nivel de afinidad, se le puntuará con 3. La Tabla 8, muestra un resumen de esto último, es decir, un cuadro de doble entrada por criterio y metodología y en donde se ve la afinidad existente.

Finalmente, en la Tabla 9 se muestra el Benchmarking final de las Metodologías, en donde se observa el ponderado de cada criterio por metodología, obtenido por la multiplicación del peso porcentual de cada criterio con el valor de afinidad.

Tabla 7

Matriz de puntuación por afinidad de criterio- Benchmarking Metodologías

OPCION	PESO
BAJO nivel de para el criterio	1
MEDIO nivel de para el criterio	2
ALTO nivel de para el criterio	3

Nota. Fuente: Elaboración propia

Tabla 8

Matriz de criterios – Benchmarking de Metodologías

	Inmon	Kimball
Usuarios finales	Medio	Alto
Implementación	Bajo	Alto
Rendimiento	Bajo	Alto
Estructura arquitectura	Bajo	Alto
Tiempo despliegue	Bajo	Medio

Nota. Fuente: Elaboración propia

Tabla 9*Benchmarking de metodologías de inteligencia de negocios*

METODOLOGÍAS		Inmon		Kimball	
CRITERIOS	PESO	Valor	Ponderado	Valor	Ponderado
Usuarios finales	0.3	2	0.6	3	0.9
Implementación	0.1	1	0.1	3	0.3
Rendimiento	0.3	3	0.9	3	0.9
Estructura arquitectura	0.2	2	0.4	3	0.6
Tiempo despliegue	0.1	1	0.1	2	0.2
Total	1	9	2.0	14	2.9

Nota. Fuente: Elaboración propia**c) Resultado**

Luego del análisis expuesto y de los resultados mostrados en la Tabla 6, se opta por aplicar la metodología de Kimball para el desarrollo del proyecto de inteligencia de negocios de la presente investigación, ya que se validó que es una metodología de fácil uso, de desarrollo incremental con un tiempo de despliegue corto y sobre todo por involucrar a los usuarios finales desde etapas tempranas y en cada etapa de la metodología. A su vez el enfoque de Kimball se adapta fácilmente a la forma de trabajo y experiencia de la autora de esta investigación.

CAPÍTULO 4: APLICACIÓN DE LA METODOLOGÍA

La metodología seleccionada para esta investigación fue la de Ralph Kimball (Kimball, 2008), la cual se basa en lo que Kimball denomina el Ciclo de Vida Dimensional del Negocio, caracterizado por realizar entregas iterativas, centrarse en el negocio, brindar una infraestructura adecuada y finalmente por ofrecer una solución completa abarcando desde la planificación hasta el crecimiento y mantenimiento de la solución (Rivadera, 2010).

Descripción de la metodología

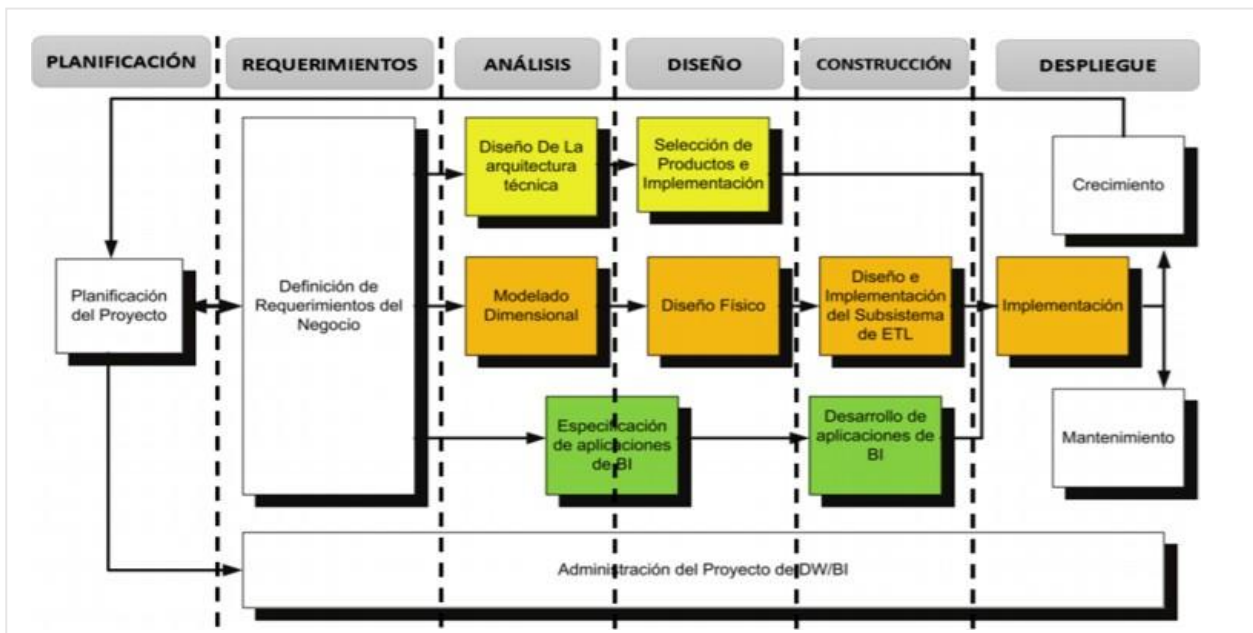
La metodología considera seis fases como se puede apreciar en la Figura 19, estas son: Planificación, Definición de Requerimientos, Análisis, Diseño, Construcción y Despliegue. Cada fase se divide en tareas. Las tareas asociadas a las fases de Planificación y Definición de Requerimientos son la base que soportaran todo el desarrollo del proyecto. Las tareas asociadas a las siguientes fases podrán iniciar en forma paralela siguiendo tres rutas (Figura 19): la primera corresponde a las tareas relacionadas con Tecnología (color amarillo), la segunda corresponde a las tareas relacionadas con Datos (color anaranjado) y la tercera corresponde a las tareas relacionadas con Aplicaciones de Inteligencia de Negocios (color verde).

Fase Planificación

La planificación es la fase inicial de un proyecto de desarrollo de solución BI. Incluye acciones típicas como: definición del proyecto, definición del alcance, la programación de actividades y el desarrollo del plan. En la definición del proyecto se evalúa la demanda de información y si esta es viable de obtenerse. El alcance define los límites del proyecto. En la programación se asignarán los roles a los miembros del equipo de acuerdo a los requerimientos del negocio. Finalmente se define el desarrollo del plan proporciona suficiente detalle para realizar el seguimiento del proyecto.

Figura 19

Tareas de la Metodología de Ralph Kimball



Nota. Fuente: Adaptado de (Rivadera, 2010)

Fase de Requerimientos

Para identificar los requerimientos del negocio se realizan entrevistas a los empleados del negocio con el fin de conocer el giro del negocio y las necesidades de información. Esta etapa es de suma importancia y requiere que se pueda realizar una adecuada traducción de las necesidades de la organización de ello depende el éxito de la solución planteada. En esta actividad se deben considerar los siguientes puntos:

- Entendimiento de terminologías del negocio.
- Analizar la situación actual de la empresa.
- Identificar idóneamente a los entrevistados.
- Revisar documentación interna de la empresa.

Fase de Análisis

Esta fase de análisis considera las tareas de Diseño de la Arquitectura técnica y el Modelado Dimensional, las cuales se detallan a continuación.

Diseño de la arquitectura técnica

El diseño de la arquitectura técnica cubre los procesos y herramientas que se aplican a los datos. La arquitectura técnica de Kimball separa la gestión de la data en dos grandes contenedores. El primero de ellos es el Back Room, el cual cubre los procesos ETL y el Front Room, el área de presentación; y las aplicaciones y servicios de Inteligencia de Negocios.

El Back Room ETL se conforma por el proceso de extracción, limpieza y transformaciones el cual es entregado al servidor de presentación. El Front Room Área de Presentación se encarga de realizar la estructura dimensional organizada en base a los procesos de negocios. El Front Room Aplicaciones y Servicios BI es la vista pública del sistema DW/BI, incluyen los reportes estándares y aplicaciones analíticas para los usuarios del negocio.

Modelado dimensional

La creación de un modelo dimensional es un proceso dinámico y altamente iterativo. Se da inicio a este proceso con el diseño de un modelo dimensional de alto nivel de procesos priorizados. Este proceso consta de cuatro pasos:

i. Elegir el proceso de negocio

Consiste en elegir el área a modelar. Esta decisión depende principalmente del análisis de requerimientos y de los temas analíticos anotados en la etapa anterior.

ii. Establecer la granularidad

La granularidad consiste en establecer el nivel de detalle de la información. La elección del nivel de granularidad depende de los requerimientos del negocio y de la factibilidad de los datos actuales. La sugerencia general es comenzar a diseñar el datawarehouse al mayor nivel de detalle, ya que en base a este se pueden realizar agrupamientos posteriores al nivel deseado.

iii. Elegir las dimensiones

Las dimensiones surgen de las reuniones del equipo, facilitadas por el nivel de granularidad. Las tablas de dimensiones poseen un conjunto de atributos (generalmente textuales) que

brindan una perspectiva de análisis sobre alguna medida de una tabla de hechos. Una forma de identificarlas es que sus atributos son candidatos a encabezados de informes, tablas y visualizaciones.

iv. Identificar medidas y tablas de hechos

El último paso del modelado dimensional, consiste en identificar las medidas provenientes de los procesos de negocios. Las medidas habitualmente se vinculan con el nivel de granularidad y se encuentran en tablas que denominamos tablas de hechos (fact en inglés). Cada tabla de hechos tiene como atributos una o más medidas de un proceso organizacional, de acuerdo a los requerimientos. Un registro contiene una medida expresada en números, como ser cantidad, tiempo, dinero, etc., sobre la cual se desea realizar una operación de agregación (suma, promedio, conteo, etc.) en función de una o más dimensiones.

Fase de Diseño

Esta fase de análisis considera las tareas de Selección del producto e implementación, el Diseño Físico y especificación de aplicaciones BI las cuales se detallan a continuación.

Selección del producto e implementación

Utilizando como marco el diseño de arquitectura técnica es necesario evaluar y seleccionar componentes específicos de la arquitectura como ser la plataforma de hardware, el motor de base de datos, la herramienta de ETL o el desarrollo pertinente, herramientas de acceso, etc. Una vez evaluados y seleccionados los componentes determinados se procede con la instalación y prueba de los mismos en un ambiente integrado de Data Warehousing.

Diseño Físico

En el diseño físico se determinan cuestiones y configuraciones técnicas del motor de almacenamiento, se considera cuán grande será el sistema DW, la configuración del sistema, determinación de memoria y asignación de servidores, configuraciones de indexación y particiones

Especificación de aplicaciones BI

Una parte fundamental de todo proyecto de inteligencia de negocios reside en proporcionar a la comunidad de usuarios una forma más estructurada y fácil de acceder al repositorio de datos. Este

acceso se proporciona a través de las aplicaciones BI, las cuales son la cara visible de al proveer los informes y aplicaciones de análisis proporcionan información útil a los usuarios. Las aplicaciones de BI incluyen un amplio espectro de tipos de informes y herramientas de análisis. Kimball divide a estas aplicaciones en dos categorías basadas en el nivel de sofisticación, y les llama informes estándar y aplicaciones analíticas.

(1) Informes estándar

Los informes estándar son la base del espectro de aplicaciones de BI. Son informes relativamente simples, de formato predefinido, y parámetros de consulta fijos, los cuales son utilizados por usuarios del negocio en el día a día brindando información de lo que sucede en un área determinada de la empresa. Algunos informes estándares típicos podrían ser:

- Ventas del año actual frente a previsión de ventas por vendedor.
- Tasa de renovación mensual por plan de servicio.
- Tasa quinquenal de deserción por unidad académica.

(2) Aplicaciones analíticas

Son más complejas que los informes estándar. Estas aplicaciones pueden incluir algoritmos y modelos de minería de datos, que ayudan a identificar oportunidades o cuestiones subyacentes en los datos, y el usuario puede pedir cambios en los sistemas transaccionales basándose en los conocimientos obtenidos del uso de la aplicación de BI. Algunas aplicaciones analíticas comunes incluyen:

- Análisis de la eficacia de las promociones.
- Análisis de rutas de acceso en un sitio Web.
- Detección de fraude.

Fase de Construcción

Esta fase de análisis considera las tareas de Diseño e Implementación del ETL y Desarrollo de Aplicaciones BI, las cuales se detallan a continuación.

Diseño e implementación del sistema de Extracción, Transformación y Carga (ETL)

El sistema ETL es la base sobre la cual se alimenta el repositorio de datos datawarehouse. Este sistema se implementa para extraer los datos alojados en los sistemas orígenes, aplicar diferentes reglas para aumentar la calidad y consistencia de los datos, consolidar la información y finalmente cargar (grabar) la información en el DW en un formato acorde para la utilización por parte de las herramientas de análisis.

Desarrollo de aplicaciones BI

Luego de haber identificado las aplicaciones o informes requeridos por los usuarios finales, se procede con el desarrollo en sí de las aplicaciones lo cual involucra configuraciones de la meta data y construcción de reportes específicos. Una vez que se ha cumplido con todos los pasos de la especificación y se tiene la posibilidad de trabajar con algunos datos de prueba, comienza el desarrollo de la aplicación.

Fase de Despliegue

Esta fase de análisis considera las tareas de Implementación y Mantenimiento y Crecimiento del Data warehouse, las cuales se detallan a continuación.

Implementación

La implementación representa la convergencia de la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio. Existen varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación y la retroalimentación.

Mantenimiento y Crecimiento del Data Warehouse

El mantenimiento consiste en administrar el entorno del Data Warehouse existente considerando las siguientes pautas: enfocarse en los usuarios de negocio, gestionar adecuadamente las operaciones del Data Warehouse, medir y proyectar su éxito y comunicarse constantemente con los usuarios para establecer un flujo de retroalimentación. Finalmente, es importante sentar las bases para el crecimiento y evolución del Data Warehouse en donde el aspecto clave es manejar el crecimiento y evolución de forma iterativa utilizando el Ciclo de Vida propuesto, y establecer las oportunidades de crecimiento y evolución en orden por nivel prioridad.

Aplicación de las fases I y II de la Metodología

En este apartado se describe la aplicación de las fases 1 y 2 de la metodología de Kimball en el proyecto objeto de esta investigación, las cuales son fases comunes a cualquier proyecto.

Aplicación de Fase I: Planificación

La empresa en estudio ofrece al mercado servicios de ingeniería y construcción, servicios mineros, negocios en concesiones de infraestructura y desarrollos inmobiliarios. Posee un equipo competente y especializado, conformado por 1600 empleados y cerca de 6 mil obreros. Los proyectos que gestiona este grupo están distribuidos en tres Unidades de Negocio, tales como: Edificaciones, Inmobiliaria y Plantas Industriales. El área encargada del monitoreo de los proyectos es la de Control de Proyecto conjuntamente con las gerencias de Cumplimiento, Recursos Humanos, Ventas y Logística de Equipos.

i) Definición del proyecto

El objetivo de este proyecto de BI fue analizar el desempeño de los proyectos de construcción a través de sus indicadores claves más relevantes que dispone la Gerencia de Control de Proyectos de la empresa constructora. Esta información esta almacenada en archivos Excel y en base de datos SQL Server.

ii) Definición del alcance

El alcance de este proyecto considera la creación de un Datawarehouse que almacene información unificada sobre los proyectos de construcción de la empresa que permita generar indicadores de desempeño a través de reportes analíticos. En lo que se refiere a las actividades a realizar se consideró las siguientes:

(1) Análisis de requerimientos: Se establece priorizar los requerimientos asociados a las sub áreas del área de Control de proyectos; estos son los relacionados a: ventas, personal, equipos y cumplimiento.

(2) Modelamiento de la solución: Se trabajó con la información contenida en la base de datos transaccional de SQL, y los registros de fuentes Excel.

(3) *Desarrollo e implementación*: Se realiza el proceso ETL para la carga del repositorio y se desarrollarán reportes de seguimiento.

(4) *Pruebas*: Se realiza las pruebas correspondientes a cada etapa de desarrollo.

Aplicación de Fase II: Requerimientos del Proyecto

Para la captura de los requerimientos del presente trabajo de investigación, se ha tenido varias reuniones con las siguientes áreas: Gerencia de Cumplimiento, Gerencia de Recursos Humanos, Gerencia de Ventas, Gerencia de Logística de Equipos. Las reuniones ayudaron a conocer el giro del negocio, y visibilizar las necesidades de la Gerencia de Control de Proyectos, con ello se han analizado los requerimientos e identificado las preguntas del negocio:

- a. ¿Cuál es la venta presupuestada y la real de cada uno de los proyectos de la gerencia?
- b. ¿Cuál es el avance de los proyectos en un determinado periodo de tiempo respecto a lo planificado?
- c. ¿Cuántos días de retraso se presentan entre el requerimiento y asignación de equipos asignados a los proyectos respecto a lo planificado?
- d. ¿Cuál es la pérdida de horas hombre debido a los accidentes reportados en un determinado proyecto?
- e. ¿Cuál es el cumplimiento de las actividades semanales programadas?

Las preguntas del negocio identificadas evidenciaron que la Gerencia de Control de Proyectos requiere el monitoreo de ciertos indicadores, para la solución propuesta se han seleccionado seis indicadores pertenecientes a la categoría Costo, Tiempo, Seguridad y Salud mostrados a continuación en la Tabla 10.

Tabla 10*Tabla de indicadores a monitorear en Gerencia de Control de Proyectos*

Número	Categoría	Indicador	Objetivo	Gerencia responsable	Meta
1	Costo	Venta	Medir el cumplimiento del monto de venta presupuestada	Gerencia de Ventas	>venta presupuestada
2	Costo	Margen	Medir el cumplimiento del margen de ganancia presupuestado	Gerencia de Ventas	>margen presupuestado
3	Tiempo	Índice de avance	Medir el cumplimiento del cronograma general en los proyectos	Gerencia de Cumplimiento	>100%
4	Tiempo	Cumplimiento de planes semanales	Medir el cumplimiento de los planes semanales en los proyectos	Gerencia de Cumplimiento	>80%
5	Tiempo	Atención de equipos	Medir los días de atraso respecto a la atención planificada de equipos.	Gerencia de Equipos.	<20 días
6	Seguridad y Salud	Índice de frecuencia ATP ¹	Medir la ocurrencia de accidentes con tiempo perdido en los proyectos	Gerencia RRHH.	<20%

Nota. Fuente: Elaboración propia¹ ATP: Accidentes con Tiempo perdido

CAPÍTULO 5: SOLUCIÓN TECNOLÓGICA

En este capítulo se detallan las actividades de las fases III a la VI, que corresponden con el Análisis, Diseño, Construcción y Despliegue, respectivamente, de la metodología de Ralph Kimball, cuyo resultado es la solución de Inteligencia de Negocios.

Aplicación de Fase III: Análisis

En esta fase se elabora la arquitectura técnica y el modelado dimensional. Como se explicó en el capítulo de marco teórico, la arquitectura técnica se refiere al esqueleto de la solución que incluye capas de fuentes de datos, su transformación y presentación; mientras que, el modelado dimensional se refiere al modelado de datos de la solución comprendido por dimensiones y tablas de hechos.

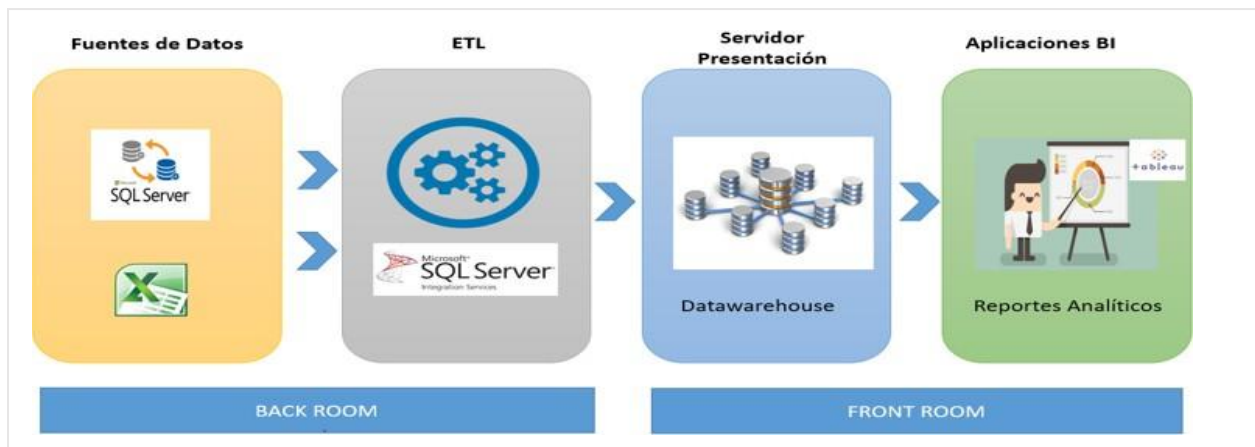
A) *Arquitectura técnica de la solución*

Para elaborar la arquitectura técnica se ha seguido el esquema típico

La figura 21 muestra la arquitectura planteada para la solución de Inteligencia de Negocios, se presentan los dos componentes del Back Room y Front Room con sus respectivas vistas y herramientas utilizadas en cada una de ellas.

Figura 21

Arquitectura de la Solución propuesta



Nota. Fuente: Elaboración propia

A continuación, se explicarán las vistas contenidas en la arquitectura presentada:

Vista Fuentes de Datos

Los datos que se utilizarán para alimentar el Datawarehouse corresponden a fuentes de datos transaccionales del sistema operacional de la empresa constructora, el cual está almacenado en SQL Server, a su vez se tienen fuentes de datos en hojas Excel.

Vista ETL

En la sección ETL se llevará a cabo las tareas relacionadas con la extracción, manipulación, control, integración, limpieza de datos, carga y actualización del Datawarehouse. Se le aplicará el proceso ETL a las fuentes de datos (archivos planos y base de datos SQL Server), el desarrollo de los paquetes se realizará en la herramienta SQL Server Integration Services.

Vista Servidor de Presentación

Los diversos datamarts que conformarán el repositorio central datawarehouse serán diseñados bajo la arquitectura ROLAP haciendo uso de modelos de datos en estrella y copo de nieve. Estas estructuras almacenarán la metadata la cual nos brinda información descriptiva sobre el contexto, calidad y características de los datos.

Vista Aplicaciones BI

Se explotará la información del datawarehouse implementado a través de reportes analíticos, los cuales serán compartidos a usuarios finales con el objetivo de mostrar datos almacenados de forma útil y transparente. Para la presentación se utilizará la herramienta de visualización Tableau Desktop.

B) Modelado Dimensional

Para que las tablas del modelado sean entendibles y de fácil lectura se utilizaron las siguientes consideraciones:

(1) Se utilizó el prefijo “BI_L” para las dimensiones y para las tablas de hecho “BI_F”, a su vez se optó por colorear las tablas de dimensiones relacionados a ciertas perspectivas del mismo color haciendo más fácil su asociación.

(2) Todo campo a emplear en el modelo del Data Warehouse se encuentra en grupos de tipos de datos como: números enteros, números decimales, tipo fecha y cadena de caracteres. Cabe recalcar que para los identificadores de las tablas se utiliza el tipo de dato Integer con la finalidad de hacer consultas más eficientes y rápidas.

El modelado se realizó en la herramienta Erwin, la cual nos permite construcción vistas de diseño lógico y físico.

Para cumplir con los requerimientos solicitados para el monitoreo de los proyectos, se modelaron 4 datamarts: Modelo CP² Avance, Modelo CP Margen, Modelo Requerimiento Equipo y Modelo ATP los cuales constituyen el repositorio unificado datawarehouse. A continuación, se detallarán los modelos con sus respectivas dimensiones y tablas de hechos.

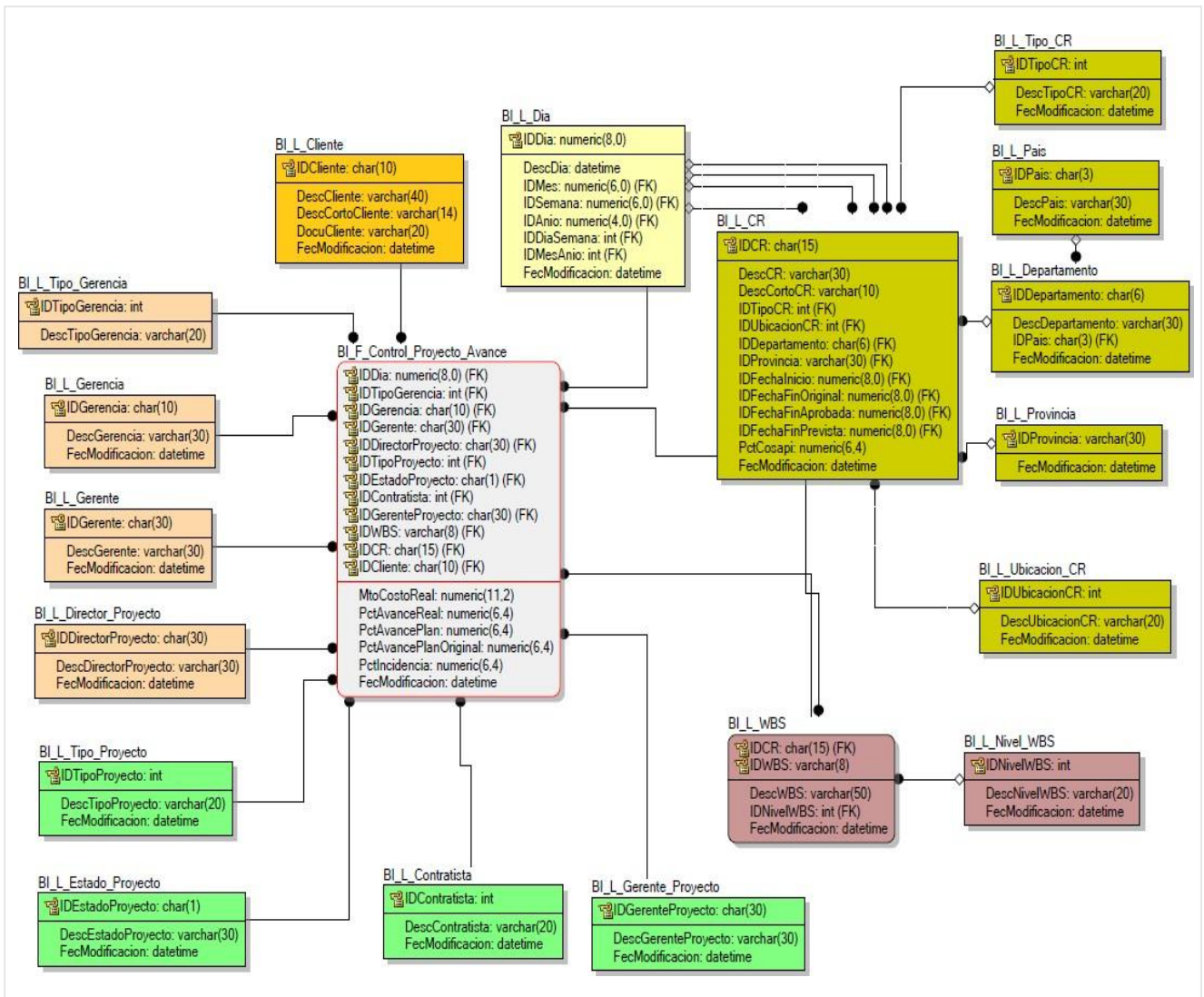
Modelo CP Avance

Este modelo mostrado en la Figura 22 contiene la información necesaria para monitorear el avance de los proyectos, posee medidas que ayudan a conocer el avance versus lo planeado. El modelo está formado por 18 dimensiones relacionadas a las siguientes perspectivas del negocio: Gerencia, Centro de Responsabilidad, Cliente y Día, las cuáles serán descritas a continuación.

² CP=Control de Proyectos

Figura 22

Modelo Constructora CP Avance



Nota. Fuente: Elaboración propia

Perspectiva Gerencia

Las dimensiones que conforman la perspectiva Gerencia muestran información relacionada a los roles ejecutores de los proyectos y a las unidades responsables. La tabla 11 detalla las 6 dimensiones pertenecientes a la perspectiva Gerencia.

Tabla 11*Dimensiones relacionadas a la Perspectiva Gerencia*

Nombre Dimensión	Descripción	Atributos	Tipo Dato
BI_L_Gerencia	Tabla maestra de gerencias de la compañía.	IdGerencia(pk) DescGerencia FecModificacion	Entero Texto Fecha
BI_L_Tipo_Gerencia	Relacionada al tipo descriptivo de cada gerencia.	IdTipoGerencia(pk) DescTipoGerencia	Entero Texto
BI_L_Gerente	Tabla maestra de gerentes de la compañía	IdGerente(pk) DescGerente FecModificacion	Entero Texto Fecha
BI_L_Director_Proyecto	Tabla maestra de directores que supervisan proyectos	IdDirectorProyecto DescDirectorProyecto FecModificación	Entero Texto Fecha
BI_L_Gerente_Proyecto	Tabla maestra de gerentes que supervisan proyectos	IdGerenteProyecto(pk) DescGerenteProyecto FecModificación	Entero Texto Fecha
BI_L_Contratista	Tabla maestra de los responsables de ejecución de la obra.	IdContratista DescContratista FecModificación	Entero Texto Fecha

Nota. Fuente: Elaboración propia*Perspectiva Centro de Responsabilidad*

Información relacionada al centro de responsabilidad, entendido como el centro que monitorea la obra y es el encargado de planificar la utilización del presupuesto y aprovisionar de recursos para la ejecución de la obra. La tabla 12 muestra las dimensiones que componen esta perspectiva.

Tabla 12

Dimensiones relacionadas a la Perspectiva Centro de Responsabilidad

Nombre Dimensión	Descripción	Atributos	Tipo Dato
BI_L_CR	Relacionada al CR muestra ubicación y descriptivos adicionales.	IDCR DesCR IDTipoCR IDUbicacion IDFechaInicio IDFechaFin	Entero Texto Fecha Entero Numérico Numérico
BI_L_Tipo_CR	Relacionada a la clasificación del CR.	IdTipoCR (pk) DescTipoCR FecModificación	Entero Texto Fecha
BI_L_Departamento	Tabla maestra de ubicaciones departamentales	IdDepartamento (pk) DescDepartamento IdPais (fk) FecModificacion	Entero Texto Entero Fecha
BI_L_Pais	Tabla maestra de ubicaciones de países	IdPais (pk) DescPais FecModificación	Entero Texto Fecha

BI_L_Provincia	Tabla maestra de ubicaciones de provincias	IdProvincia(pk)	Entero
		DescProvincia	Texto
		FecModificación	Fecha

Nota. Fuente: Elaboración propia

Perspectiva Cliente y Día

Se presentan 2 dimensiones, la de la cliente referida a la persona o entidad solicitante del proyecto y la dimensión Día contiene la granularidad y jerarquías que podemos encontrar en un calendario.

Tabla 13

Dimensiones relacionadas a la perspectiva Cliente y Tiempo

Nombre Dimensión	Descripción	Atributos	Tipo Dato
BI_L_Cliente	Relacionada al cliente	IDCliente (pk)	Int
		DescCliente	Texto
		DescCortoCliente	Texto
		DocuCliente	Texto
		FecModificación	Fecha
BI_L_Dia	Relacionado a la granularidad de la fecha	IDDia	Numérico
		DescDia	Fecha
		IDMes	Numérico
		IDSemana	Numérico
		IDAnio	Numérico
		IDDiaSemana	Entero
		IDMesAnio	Entero
		FecModificacion	Fecha

Nota. Fuente: Elaboración propia

Tabla de hechos

La tabla central del modelo se denomina BI_F_Control_Avance, la cual se conecta a las dimensiones detalladas anteriormente, a su vez esta tabla contiene medidas, por ejemplo: MtoCostoReal, PctAvanceReal, entre otras y atributos los cuales son llaves foráneas de las dimensiones relacionadas, la estructura de esta tabla se muestra en la figura 23.

Figura 23

Fact Table del modelo CP Avance

BI_F_Control_Proyecto_Avance	
	IDDia: numeric(8,0) (FK)
	IDTipoGerencia: int (FK)
	IDGerencia: char(10) (FK)
	IDGerente: char(30) (FK)
	IDDirectorProyecto: char(30) (FK)
	IDTipoProyecto: int (FK)
	IDEstadoProyecto: char(1) (FK)
	IDContratista: int (FK)
	IDGerenteProyecto: char(30) (FK)
	IDWBS: varchar(8) (FK)
	IDCR: char(15) (FK)
	IDCliente: char(10) (FK)
	MtoCostoReal: numeric(11,2)
	PctAvanceReal: numeric(6,4)
	PctAvancePlan: numeric(6,4)
	PctAvancePlanOriginal: numeric(6,4)
	PctIncidencia: numeric(6,4)
	FecModificacion: datetime

Nota. Fuente: Elaboración propia

Medidas

La tabla BI_F_Control_Avance contiene cinco medidas las cuales nos permiten visibilizar los costos y porcentajes de avance planeado y real asociados a los proyectos de construcción.

1. *MtoCostoReal*: Muestra el costo asignado a cada proyecto en ejecución.
2. *PctAvanceReal*: Muestra el porcentaje de avance real de cada uno de los proyectos.

3. *PctAvancePlanOriginal*: Muestra el porcentaje de avance planeado originalmente de cada proyecto.
4. *PctAvancePlan*: Muestra el porcentaje de avance planeado de cada proyecto
5. *PctIncidencia*: Muestra el porcentaje de incidencia de un proyecto, si lo hubo.

Los tres modelos restantes que componen el datawarehouse fueron desarrollados bajo los lineamientos comentados. Por ello no se hará mucho énfasis en cada uno de ellos y se mostrará solo el modelo.

Modelo CP Margen

El modelo de Control de Proyecto Margen mostrado en la Figura 24, nos visibilizará información relevante de los costos y ventas asociados en la ejecución de cada proyecto. Ello nos ayudará a ver cuál es el margen de ganancia respecto a lo planteado en el resultado operativo de la empresa.

Modelo Requerimiento Equipo

El modelo de Requerimiento de Equipos mostrado en la Figura 25, nos visibilizará información relevante a la gestión de los equipos desde el aprovisionamiento, planificación, espera y atención.

Modelo ATP

El modelo ATP mostrado en la figura 26, nos permite conocer el comportamiento del trabajo del capital humano en el proyecto. Podemos visibilizar la cantidad de horas trabajadas y las horas afectadas por los accidentes en el trabajo tanto en horas como el costo que implica en el proyecto.

Figura 24 Modelo CP Margen

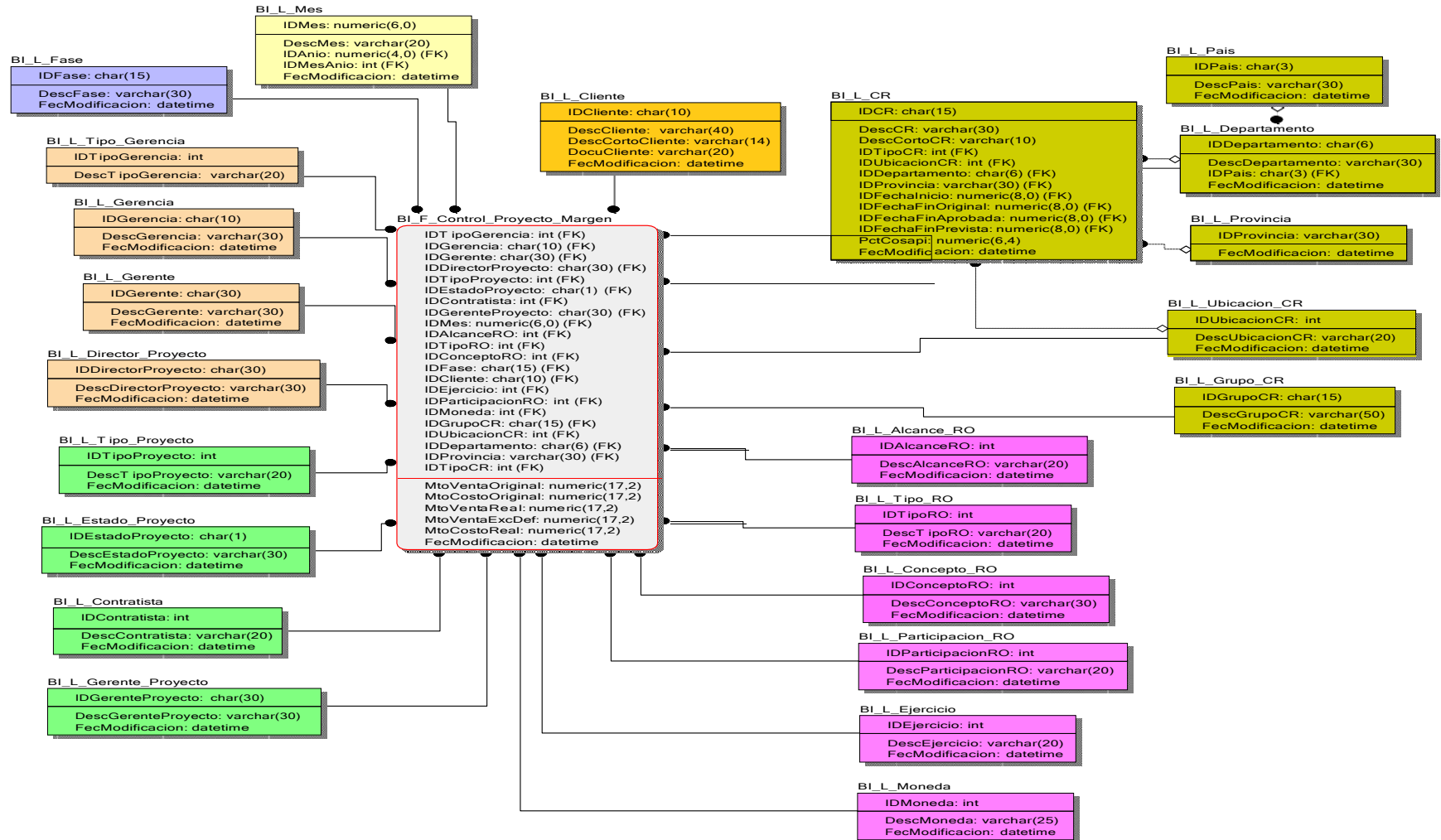


Figura 25 *Modelo CP Requerimiento Equipo*

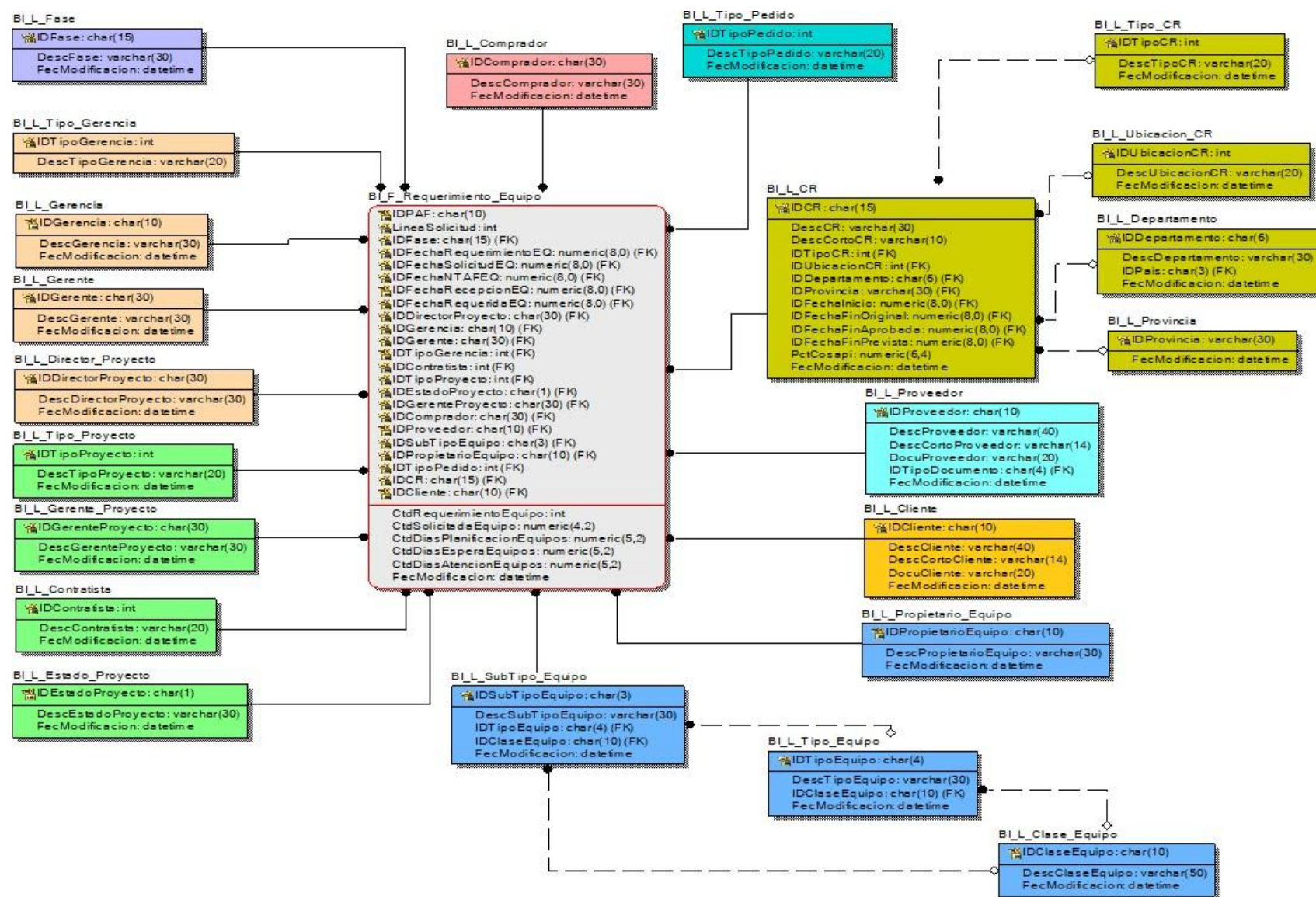
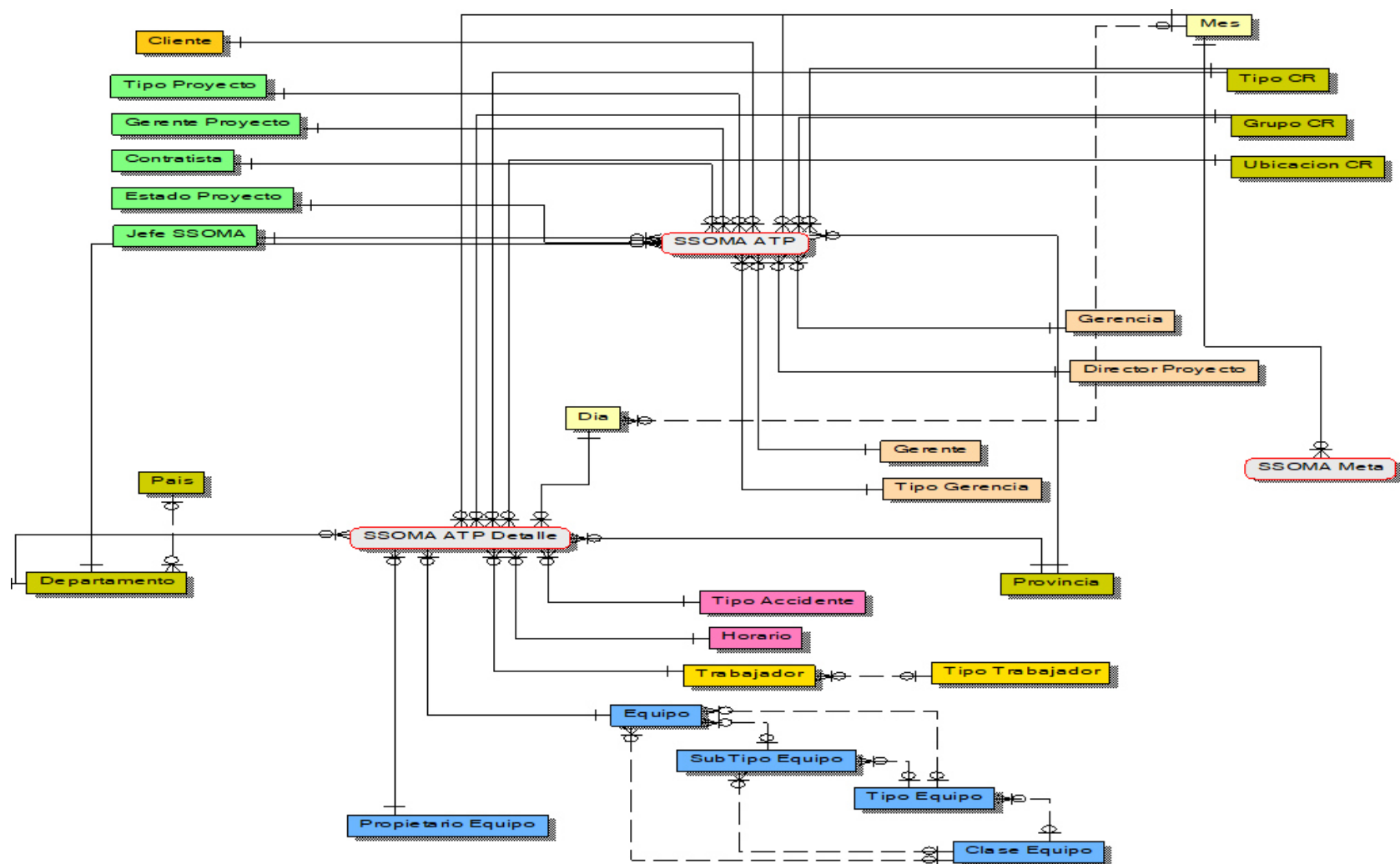


Figura 26 Modelo SSOMA - Accidentabilidad



Aplicación de Fase IV: Diseño

En la fase de diseño se desarrollaron las siguientes tareas a realizar: Selección del producto de implementación y Especificación de aplicaciones BI.

A) Selección del producto de implementación

La empresa constructora en estudio, como se comentó anteriormente, tiene parte de la data transaccional en SQL Server, por ello se opta por almacenar el modelo dimensional en SQL Server respetando las tecnologías manejadas por el área. Para el proceso ETL se seleccionó Integration Services, ya que al ser un producto de Microsoft posee los conectores necesarios y a su vez por la familiaridad de la herramienta. Finalmente, para la parte de explotación de datos, basados en el comparativo de herramientas líderes mostrado en el capítulo de Revisión de Literatura, se seleccionó Tableau por ofrecer las características de facilidad de uso, variedad de gráficos y por ser intuitiva.

B) Especificación de aplicaciones BI

En la presente tesis, se ha considerado la construcción de un datawarehouse el cuál contendrá información relevante de los proyectos de construcción como el monitoreo del avance, los costos, la utilidad y la accidentabilidad de los mismos. Esta información relevante será explotada a través de reportes analíticos en la herramienta Tableau, los cuales estarán disponibles en el servidor web y serán accedidos por los usuarios del negocio.

Aplicación de Fase V: Construcción

En la fase de Construcción se desarrollaron las siguientes tareas: Diseño e Implementación del ETL y Desarrollo de Aplicaciones BI.

A) Diseño e Implementación del ETL

El proceso de transformación de datos nos permitirá llevar los datos del ambiente transaccional al ambiente Datawarehouse, se realizaron dos procesos de transformación de datos. Para el primero se realizó un ETL con la herramienta Integration Services con el fin de poder llevar la data a un ambiente intermedio (Stagging Area); la segunda transformación se realizó mediante los Stores Procedures con el fin de llevar la información al Datawarehouse (ver figura 27).

Figura 27

Esquema de Transformación de Datos



Nota. Fuente: Elaboración propia

Proceso ETL - Staging Area

Desarrollado en el paquete SSIS (Integration Services) de SQL Server, en este proceso se transforman tablas que usaremos en la construcción del Data warehouse de los orígenes de archivos planos y del ERP Peoplesoft (información almacenada en SQL Server), estos destinos modificados se cargan en una base intermedia llamada Staging Area.

Proceso ETL- Datawarehouse

Desarrollado en el ambiente de SQL Server, nos apoyamos de los Stores Procedures para la carga definitiva de la data en las tablas hechos y dimensiones al ambiente productivo Warehousing. Básicamente en los Stores se aplicaron los filtros y selección en base a la información relevante a obtener para cada fuente.

En pasos anteriores se identificó las fuentes de información y se diseñó la arquitectura de nuestras tablas del destino del datawarehouse, para ello se realizó el mapeo de la base de datos fuente con la base de datos del destino. De esta manera se establece la relación de campos migrados. Luego se realizaron las transformaciones adecuadas que sufrirá cada campo de las tablas para el eficiente ingreso de datos.

A continuación, se detalla cada una de las etapas anteriormente comentadas del proceso ETL.

Etapas 1: Carga al Staging Area

En esta etapa vamos a definir el estándar ETL utilizado, el procedimiento general y un ejemplo de carga del modelo CP Avance.

Estándar de ETL

El estándar utilizado para la asignación de nombres de los 38 paquetes ETL construidos en SSIS nombra a los procesos con el prefijo “CargaDatos” seguido del origen del sistema origen (los utilizados (1) PS: Peoplesoft, (2) RRHH: Recursos Humanos (3) XLS: archivo Excel) y finalmente un nombre que ayude a identificar el proceso.

Figura 28

Carga Staging Area – Lista de Procesos ETL

Nombre	Fecha de modificación	Tipo	Tamaño
CargaDatos_PS	10/04/2015 17:51	Archivo DTSX	2.981 KB
CargaDatos_RRHH	25/04/2021 10:56	Archivo DTSX	215 KB
CargaDatos_XLS_CPAvance_WBS	25/04/2021 10:56	Archivo DTSX	60 KB
CargaDatos_XLS_CPMargen_Alcance	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_CPMargen_Concepto	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_CPMargen_Moneda	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_CPMargen_Originales	25/04/2021 10:56	Archivo DTSX	72 KB
CargaDatos_XLS_CPMargen_Participacion	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_CPMargen_Tipo	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_CPPAC_TipoCausa	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_CPPAC_WBS	25/04/2021 10:57	Archivo DTSX	60 KB
CargaDatos_XLS_CPPEQ_CPProdFaseInicial	25/04/2021 10:57	Archivo DTSX	68 KB
CargaDatos_XLS_CPPEQ_Fase	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_CPPMO_Fase	25/04/2021 10:57	Archivo DTSX	50 KB
CargaDatos_XLS_CPPMO_Inicial	25/04/2021 10:57	Archivo DTSX	86 KB
CargaDatos_XLS_CR_Especiales	25/04/2021 10:57	Archivo DTSX	55 KB
CargaDatos_XLS_GrupoCR	25/04/2021 10:57	Archivo DTSX	61 KB
CargaDatos_XLS_Mensual	25/04/2021 10:56	Archivo DTSX	415 KB
CargaDatos_XLS_Proyecto	25/04/2021 10:57	Archivo DTSX	78 KB
CargaDatos_XLS_Requerimiento_Equipo	25/04/2021 10:57	Archivo DTSX	91 KB
CargaDatos_XLS_RRHH_Analista	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_RRHH_AreaProyecto	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_RRHH_CargoObrero	25/04/2021 10:57	Archivo DTSX	51 KB
CargaDatos_XLS_RRHH_CategoriaEmpleado	25/04/2021 10:57	Archivo DTSX	52 KB

Nota. Fuente: Elaboración propia

La Figura 28 muestra los archivos que componen el proceso ETL. En rojo se muestra los procesos consolidadores de extracción de fuentes de datos transaccionales tanto del ERP PeopleSoft como

del sistema de RRHH; mientras que en morado se muestran los procesos extractores de las fuentes Excel.

Procedimiento General

Las transformaciones generadas se realizaron para estandarizar las consultas hacia el cubo. Dentro de estas transformaciones se encuentran: Estandarización de nombres, formato de campos, lenguaje o idioma, tamaño de campos permitido y campos pre-calculados. Estas operaciones se realizan en las distintas secuencias que pasa la tabla fuente, este flujo por el que pasa la información se diseña por medio de características ofrecidas por la herramienta de SQL Integration Services, el cual permite analizar el flujo de los datos y las distintas agregaciones y filtros por la que recorre la información antes de residir en las tablas destino.

El mapeo de los datos se valida en cada fase de los procesos gracias a las distintas características que ofrece la herramienta. En el caso de cargar de las tablas destino de las tablas de Oracle Peoplesoft se desarrolla una selección de campos y transformaciones necesarias de los mismos que nos sean útiles para la solución.

La Figura 29 muestra el esquema de carga de las tablas del ERP, en el cuál observamos dos pasos principales en la estructura: la limpieza de datos y la carga de datos.

Figura 29

Esquema de Flujo de Control de Fuentes PeopleSoft

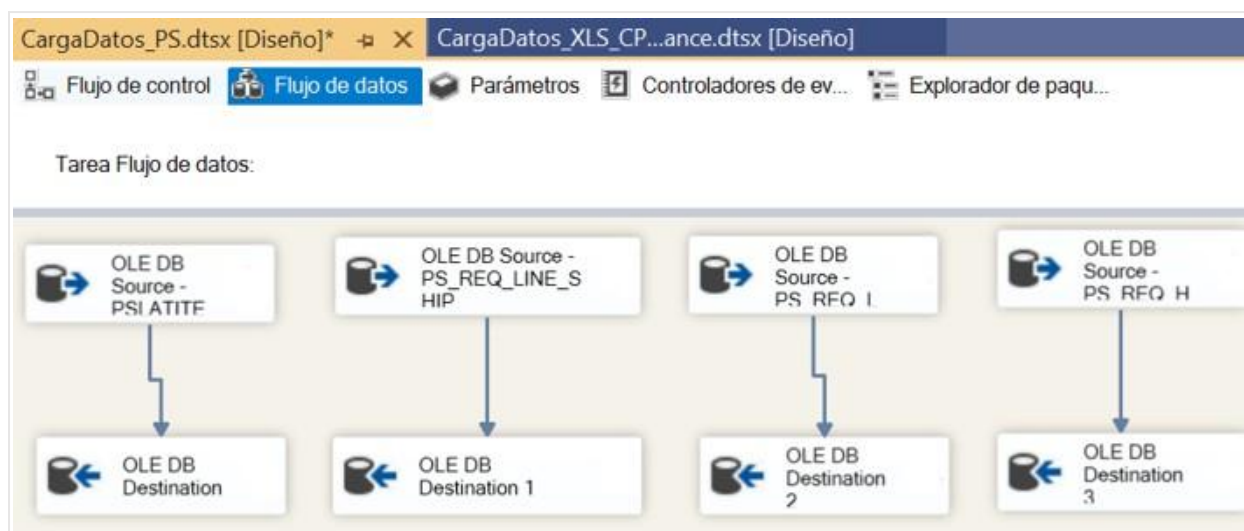


Nota. Fuente: Elaboración propia

La tarea *BorrarTablasStageDWConstructora* tal como su nombre lo indica realiza la limpieza de todas las fuentes destino, con ello las fuentes estarán preparadas para cargarse. La segunda tarea *Carga de Datos* se encargará de mapear los orígenes de datos y cargar la data a las tablas Stage destino. La figura 30 muestra a detalle la *Carga de Datos* de todas las fuentes de interés del sistema transaccional PeopleSoft, como se observa, se realiza el mapeo del origen de datos y luego se mapean los campos requeridos a la conexión destino de cada fuente.

Figura 30

Esquema de Flujo de Datos de Fuentes PeopleSoft



Nota. Fuente: Elaboración propia

Modelo CP Avance

Acorde a lo detallado en el modelado, donde se realizó énfasis en el modelo CP Avance, se va a detallar el proceso ETL relacionado a sus fuentes. Se detallará el mapeo (origen de datos – tabla Staging – tabla DWH) de las fuentes que permiten poblar las tablas Hechos BI_F_Control_Proyecto_Avance y dimensiones como BI_L_CR, BI_L_Gerente_Proyecto, entre otras miembros de este modelo (ver Tabla 14).

Tabla 14*Fuentes en construcción de CP Avance Stagging Area*

Tabla DWH (DWConstructora)	Tabla Intermedia (StageDWConstructora)	Origen de datos (MDConstructora)
BI_L_CR	PS_CO_CR_VW PS_CO_PROJECT	PS_CO_CR_VW PS_CO_PROJECT
BI_L_Gerente_Proyecto	PS_PERSONAL_DATA	PS_PERSONAL_DATA
BI_L_Cliente	PS_CUSTOMER PS_CUST_ID_NBRS	PS_CUST_ID_NBRS
BI_F_Control_Proyecto_Avance	XLS_CPAvance	Archivo XLS CP Avance

Nota. Fuente: Elaboración propia

A continuación, se detalla el procedimiento de carga de las fuentes Stagging XLS_CPAvance, PS_CO_PROJECT y PS_CO_CR_VW.

Carga XLS_CPAvance:

Se muestran los pasos necesarios para poblar la tabla XLS_CP_Avance. El flujo de datos mostrado en la Figura 31 realiza las siguientes tareas:

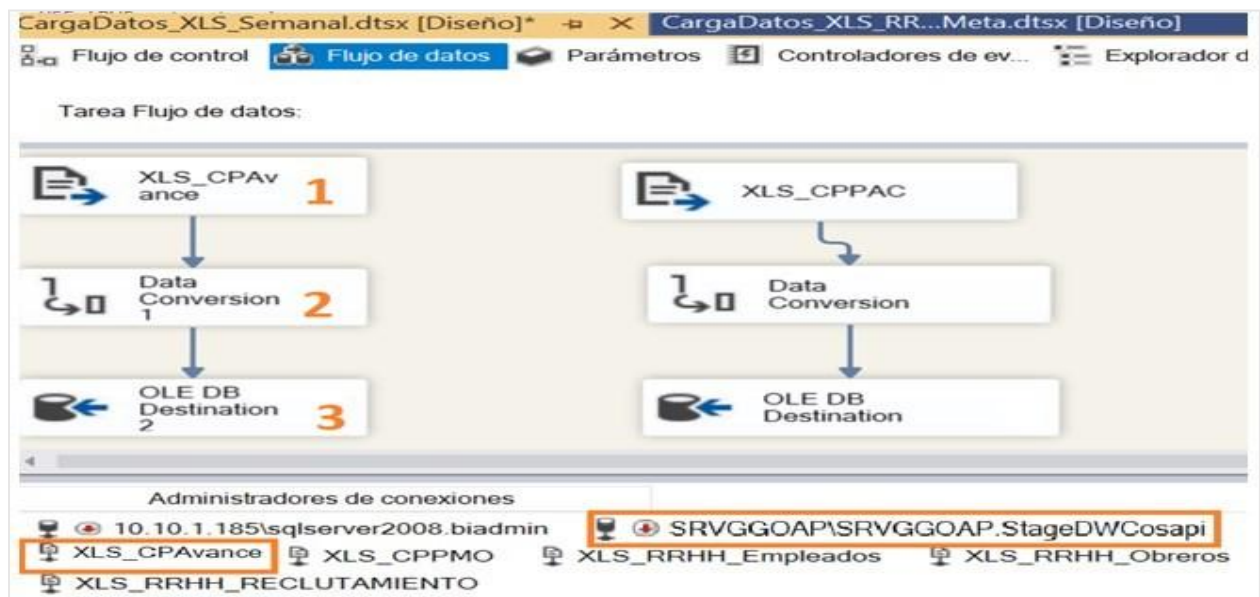
(1) *Flat File Sources:* Componente que permite el manejo de archivos planos. Aquí se mapea el archivo XLS_CPAvance.xls

(2) *Data Conversion:* Componente utilizado para realizar conversiones a la data origen y así dejarla lista para el formato destino.

(3) *Comando de OLEDB:* Nos permite mapear la conexión al destino. Aquí se mapea la conexión al servidor Stage.

Figura 31

Flujo de Datos de la Carga CP Avance



Nota. Fuente: Elaboración propia

El convertor de datos se utilizó para realizar el mapeo de los campos requeridos. La Figura 32 muestra la selección y definición de estructura de los campos.

Figura 32

Convertor de Datos de XLS_CP_Avance

Configure las propiedades utilizadas para convertir el tipo de datos de una columna de entrada a otro tipo. Configure la escala y la página de códigos de la columna en función del tipo de datos al que se convertirá la columna.

Columna de entrada	Alias de salida	Tipo de datos	Longitud	Precisión	Escala
Dia	Copy of Dia	fecha [DT_DATE]			
CR	Copy of CR	cadena [DT_STR]	15		
WBS	Copy of WBS	cadena [DT_STR]	20		
Incidencia	Copy of Incidencia	cadena [DT_STR]	8		
% Avance Real	Copy of % Avance Re...	cadena [DT_STR]	8		
% Avance Planeado	Copy of % Avance Pl...	cadena [DT_STR]	8		
% Avance Planeado ...	Copy of % Avance Pl...	cadena [DT_STR]	8		

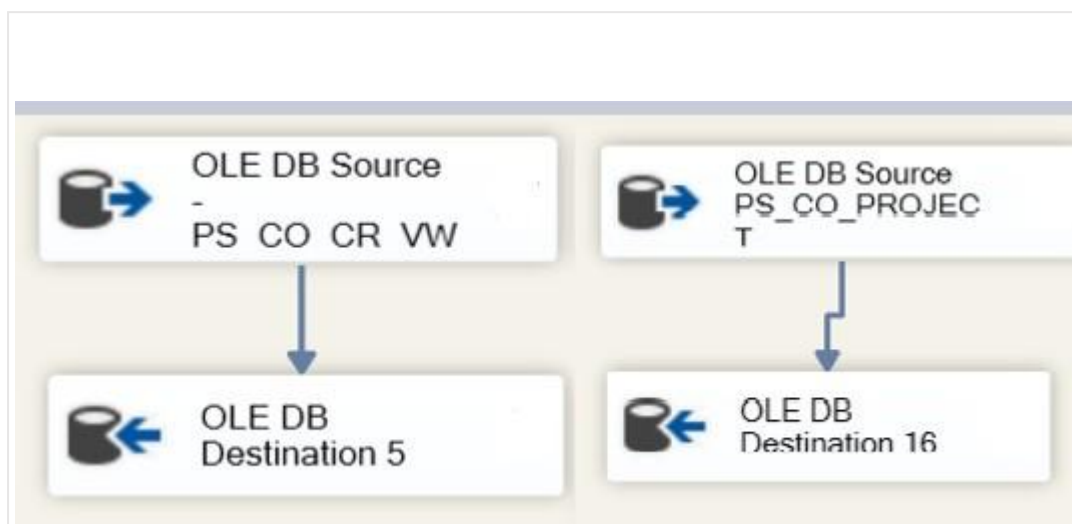
Nota. Fuente: Elaboración propia

Carga PS_CO_CR_VW y PS_CO_PROJECT

La vista del centro de responsabilidad y la del proyecto son una de las fuentes principales que servirán para poblar la tabla BI_L_CR del ambiente Warehousing. Como se observa en la figura 33 al ser los orígenes provenientes de la base de datos no requieren de casteo de datos, por ello solo se utiliza los componentes OLE DB para mapear conexiones origen y destino.

Figura 33

Flujo de Datos PS_CO_CR_VW y PS_CO_PROJECT



Nota. Fuente: Elaboración propia

Etapas 2: Carga al Data Warehouse

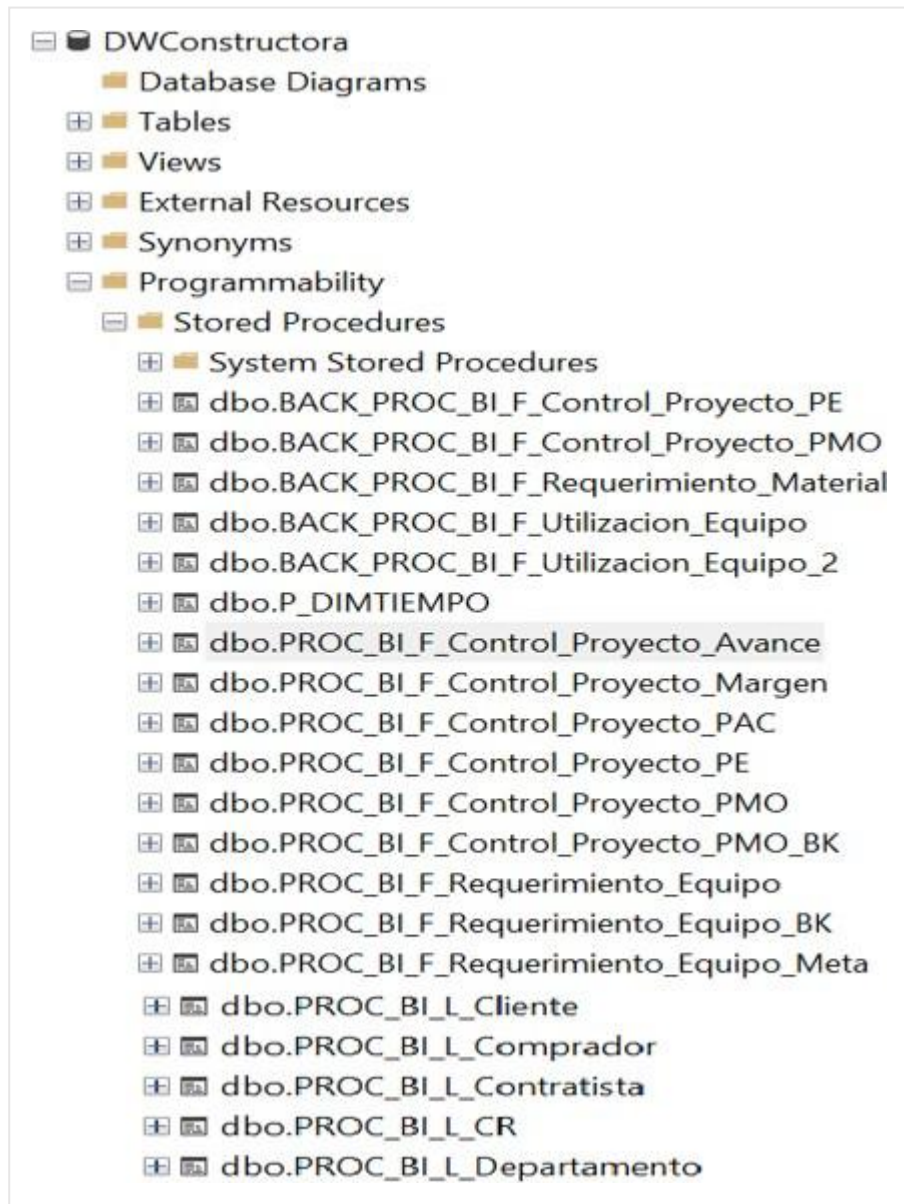
En esta etapa vamos a definir el estándar ETL utilizado, y un ejemplo de carga del modelo CP Avance.

Estándar de ETL

El estándar utilizado para la asignación de nombres de los 25 procedimientos almacenados desarrollados en SQL Server, nombra a los procesos con el prefijo “PROC_BI” seguido del nombre de la tabla a cargar. La Figura 34 muestra los procedimientos que componen el proceso ETL para la carga al ambiente Datawarehousing tanto de las tablas de hechos como dimensiones. Nótese que estos son creados en el mismo gestor de base de datos SQL Server en la parte de Programmability.

Figura 34

Carga Datawarehousing - lista de procedimientos almacenados



Nota. Fuente: Elaboración propia

Modelo CP Avance

Acorde a lo detallado en el modelado, donde se realizó énfasis en el modelo CP Avance, se va a detallar el proceso ETL “PROC_BI_F_Control_Proyecto_Avance”. Iniciamos con la limpieza de la tabla, seguido de la inserción de registros en base a los cruces de información de fuentes de interés. Como se observa en la Figura 35, la tabla Fact se arma en base al cruce principal de la

fuentes Staging XLS_CP_Avance, con sus diversas dimensiones para obtener las llaves primarias de esta tabla.

Figura 35

Procedimiento de carga a fuente BI_F_Control_Proyecto_Avance

```
USE [DWConstructora]
GO
/***** Object: StoredProcedure [dbo].[PROC_BI_F_Control_Proyecto_Avance]    Script Date: 25/04/2021 13:53:26 *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER PROCEDURE [dbo].[PROC_BI_F_Control_Proyecto_Avance]
DECLARE DIAEJECUCION AS DATE;
AS

SELECT MAX(IDDIA) INTO DIAEJECUCION FROM StageDWCosapi.dbo.XLS_CPAdvance;
DELETE BI_F_Control_Proyecto_Avance WHERE IDDIA>=DIAEJECUCION;

INSERT INTO BI_F_Control_Proyecto_Avance(IDDIA, IDTipoGerencia,
IDGerencia, IDGerente, IDDirectorProyecto, IDTipoProyecto,
IDEstadoProyecto, IDContratista, IDGerenteProyecto, IDWBS, IDCR,
IDCliente, MtoCostoReal, PctAvanceReal, PctAvancePlan, PctAvancePlanOriginal,
PctIncidencia, FecModificacion)
SELECT ISNULL(CONVERT(Numeric(8, 0), CONVERT(CHAR(8), T1.Dia, 112)), 19900101) Dia,
-1 IDTipoGerencia,
ISNULL(T2.PROJECT_TYPE, ISNULL((SELECT IDGerencia from StageDWCosapi.dbo.XLS_Proyecto where IDCR=T1.CR), '.')) IDGerencia,
',' IDGerente,
ISNULL((SELECT TOP 1 EMPLID
FROM StageDWCosapi.dbo.PS_CO_CARGO_PROJ
WHERE BUSINESS_UNIT = T2.BUSINESS_UNIT
AND PROJECT_ID = T2.PROJECT_ID
AND CO_TIPO_CARGO = '1'
WHERE BUSINESS_UNIT = T2.BUSINESS_UNIT
AND PROJECT_ID = T2.PROJECT_ID
ORDER BY EFFDT DESC), ',') IDEstadoProyecto,
ISNULL(T2.INTEGRATION_TMPL, ',') IDContratista,
ISNULL((SELECT TOP 1 EMPLID
FROM StageDWCosapi.dbo.PS_CO_CARGO_PROJ
WHERE BUSINESS_UNIT = T2.BUSINESS_UNIT
AND PROJECT_ID = T2.PROJECT_ID
AND CO_TIPO_CARGO = '2'
ORDER BY EFFDT DESC), ',') IDGerenteProyecto,
T1.WBS IDWBS,
T1.CR IDCR,
ISNULL(T3.CUST_ID, ',') IDCliente,
T4.MtoCostoReal MtoCostoReal,
(CASE WHEN RTRIM(LTRIM(T1.PorcAvanceReal)) = '' THEN NULL ELSE CONVERT(NUMERIC(8, 4), T1.PorcAvanceReal) END),
(CASE WHEN RTRIM(LTRIM(T1.PorcAvancePlaneado)) = '' THEN NULL ELSE CONVERT(NUMERIC(8, 4), T1.PorcAvancePlaneado) END),
(CASE WHEN RTRIM(LTRIM(T1.PorcAvanceOriginal)) = '' THEN NULL ELSE CONVERT(NUMERIC(8, 4), T1.PorcAvanceOriginal) END),
(CASE WHEN RTRIM(LTRIM(T1.PorcIncidencia)) = '' THEN NULL ELSE CONVERT(NUMERIC(8, 4), T1.PorcIncidencia) END),
GETDATE()
FROM StageDWCosapi.dbo.XLS_CPAdvance T1 LEFT JOIN StageDWCosapi.dbo.PS_PROJECT T2
ON T2.PROJECT_ID = T1.CR
LEFT JOIN StageDWCosapi.dbo.PS_CO_PROJECT T3
ON T3.BUSINESS_UNIT = T2.BUSINESS_UNIT
AND T3.PROJECT_ID = T2.PROJECT_ID
LEFT JOIN (SELECT LEFT(IDGrupoCR, 4) IDGrupoCRX,
CONVERT(CHAR(4), IDEjercicio) IDEjercicioX,
IDMes,
SUM(MtoCostoReal) MtoCostoReal
FROM dbo.BI_F_Control_Proyecto_Margen XX
WHERE IDParticipacionRO = 1
```

Nota. Fuente: Elaboración propia

B) Desarrollo de aplicaciones BI

Los análisis previamente documentados en la toma de requerimientos hacen posible que se tenga una idea de los inputs y los posibles outputs de cada reporte; pero principalmente es la manera como presentarlo en el diseño lo que hará que los reportes tengan un valor agregado para la toma de decisión. Acorde a las metas del proyecto, se construirán los siguientes tableros:

Tabla 15

Listado de tableros construidos en la herramienta Tableau

Nro	Nombre	Descripción	Información	Filtros Especiales
1	RO Proyectos	Permite conocer el resultado operativo de los proyectos manejados en la empresa, permitiendo analizar la venta adjudicada, el margen de ganancias anual de los proyectos.	- Periodo. - Unidad de Negocio. - Venta en Proyectos. - Utilidad. - Margen	- Unidad de Negocio. - Año y Mes. - Moneda.
2	Avance	El reporte permite conocer información sobre el cumplimiento del avance físico planeado en los proyectos de la empresa en un determinado periodo.	- Indicadores de avance acumulado. - Indicadores de avance Mensual.	- Centro de Responsabilidad.
3	Accidentabilidad	Este reporte permite revisar información sobre medidas de seguridad tomadas en los proyectos, se muestran los niveles de accidentabilidad.	- Indicadores de frecuencia ATP.	- Centro de Responsabilidad.

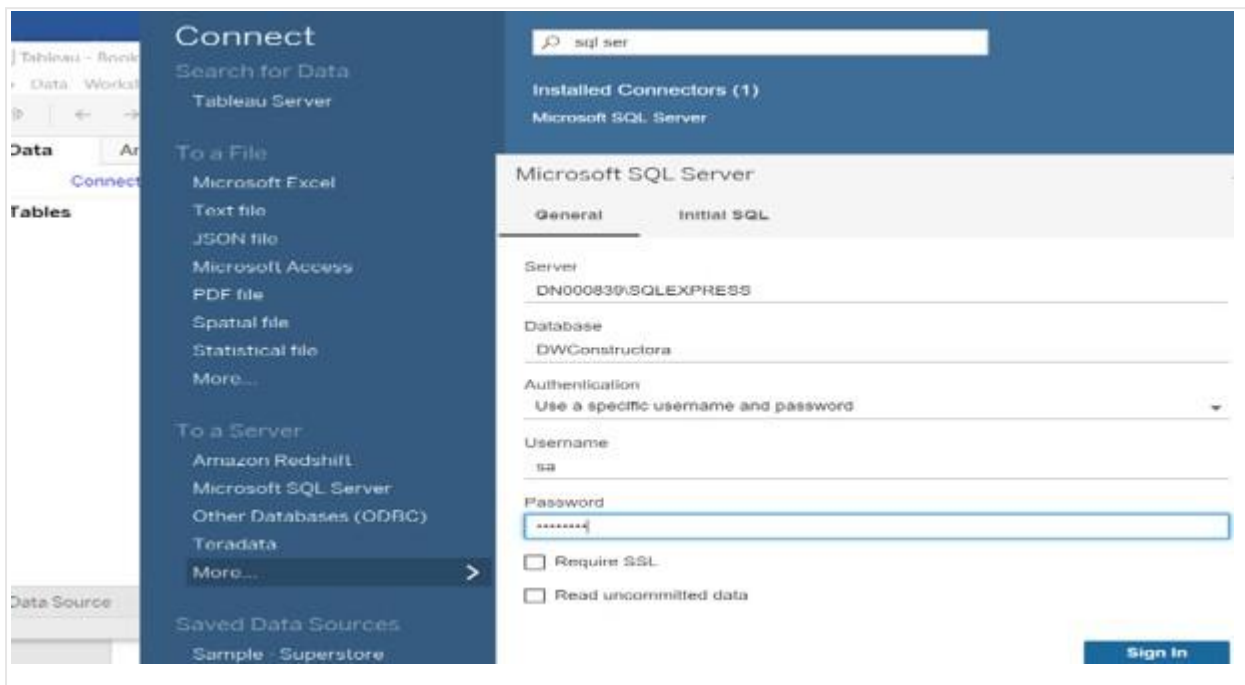
Nota. Fuente: Elaboración propia

Procedimiento de Construcción

Una vez que se tiene poblado el Datawarehouse, se utilizaron estas fuentes para la creación de los tableros comentados. La construcción se realizará en Tableau, esta herramienta ya tiene configurada la conexión al SQL Server, tan solo nos basta elegir la opción y autenticar las credenciales de conexión (ver Figura 36).

Figura 36

Tableau – Conexión de Orígenes de Datos

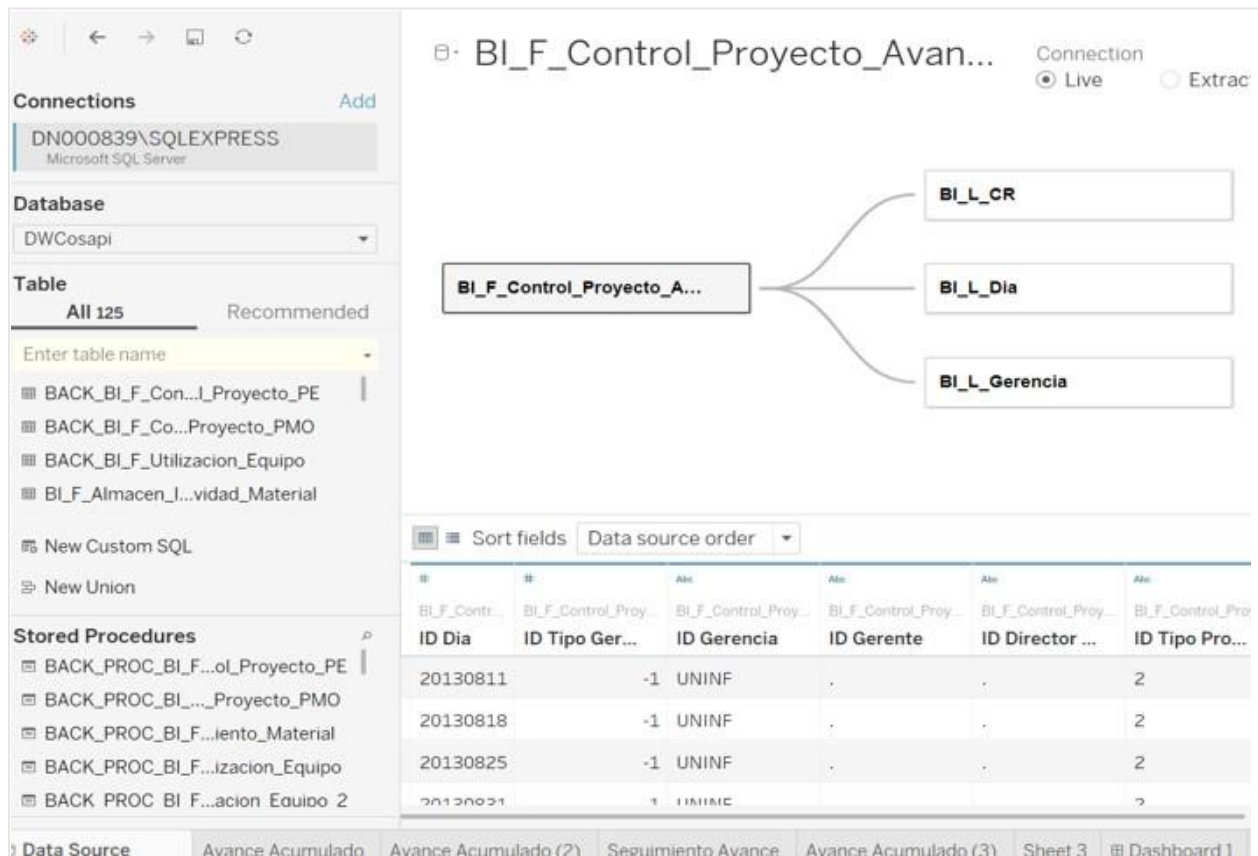


Nota. Fuente: Elaboración propia

Una vez autenticados, se seleccionó la base de datos y las tablas involucradas a usar, estas son arrastradas al lienzo de Data Source donde se creó las relaciones de tablas en el esquema de trabajo de Tableau mostrado en la Figura 37; de este modo se tiene la data consolidada para una posterior explotación en reportes y tableros de control.

Figura 37

Tableau – Esquema de Trabajo Data Source



Nota. Fuente: Elaboración propia

En pasos anteriores se explicó la carga de la tabla de hechos BI_F_Control_Proyecto_Avan..., fuente input para el tablero de Avance. Una vez importada la data, se selecciona la opción “New Worksheet” mostrada en la Figura 38 donde crearemos cada uno de los reportes que incluirá el tablero.

Figura 38

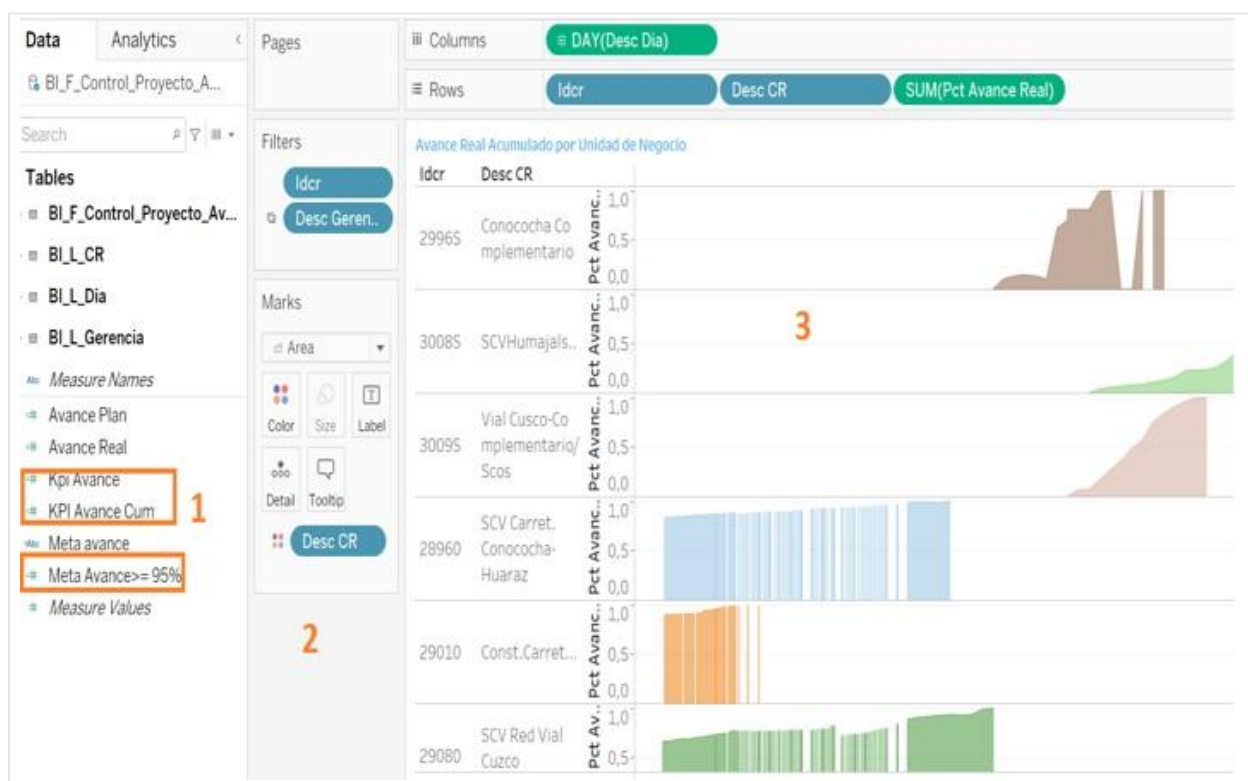
Tableau- Creación de nueva hoja de trabajo Worksheet



La Figura 39 muestra a continuación el flujo de trabajo realizado. Aquí se seleccionaron los campos de interés, creación de nuevos campos relacionados al cálculo del indicador y se seleccionan los gráficos de visualización.

Figura 39

Tableau – Hoja de trabajo Avance Proyecto

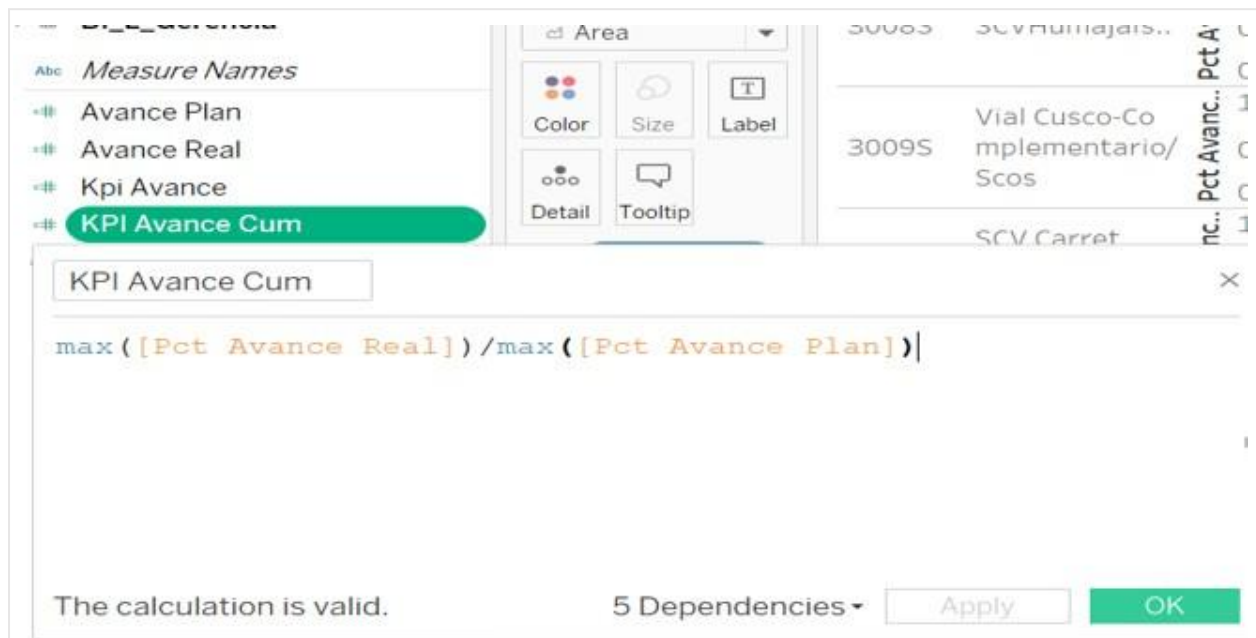


Nota. Fuente: Elaboración propia

(1) *Gestión de Data:* Aquí buscamos los campos que van a utilizarse en la hoja de trabajo. Apoyados en las medidas de la fuente BI_F_Control_Avance, construiremos los indicadores Kpi Avance, Kpi AvanceCum y la medida Meta Avance. A continuación, se detalla la fórmula del indicador Avance Cumplimiento. Por ejemplo, la Figura 40 muestra la fórmula del KPI Avance Cum.

Figura 40

Tableau – Creación de indicador Avance Cumplimiento



Nota. Fuente: Elaboración propia

(2) *Configuración de Diseño:* En esta sección se podrá personalizar las etiquetas de datos, elegir las formas, el color y determinar el nivel de detalle a mostrar. Aquí se ha seleccionado que el color se distingue en base al Centro de Responsabilidad.

(3) *Lienzo de trabajo:* Se especifican los campos que irán en las filas y columnas, así como la jerarquía con la que se mostrará. La Figura 41 muestra que la fecha se mostrará al nivel de detalle día y la medida a mostrar será el acumulativo del avance real.

Figura 41

Tableau – Campos en filas y columnas



Nota. Fuente: Elaboración propia

Dependiendo de la cantidad de atributos y medidas, Tableau nos sugerirá los gráficos y/o seleccionar el que se adecue mejor a lo que queremos visualizar.

Figura 42

Tableau – tipos de visualizaciones disponibles



Nota. Fuente: Elaboración propia

Tableros construidos en Tableau

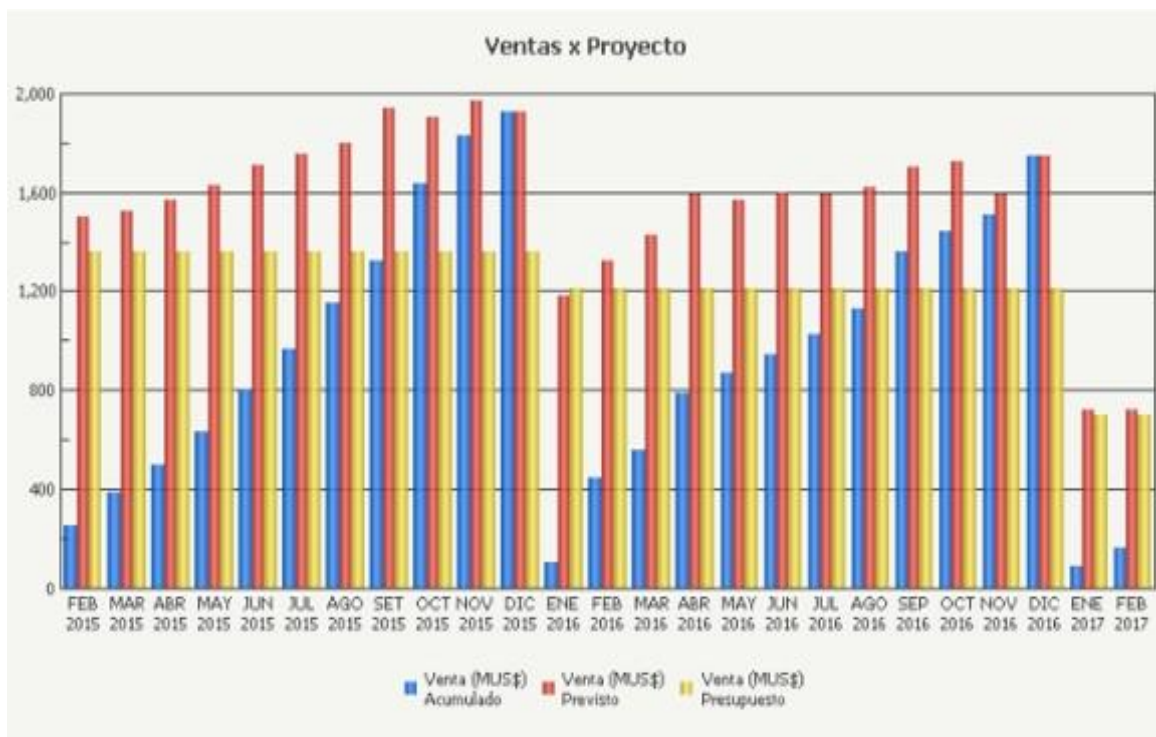
Como se comentó anteriormente se han construido 3 tableros, los cuales apoyan al monitoreo de proyectos de construcción aportando información de las categorías de indicador: Costo, Tiempo y Seguridad y Salud respectivamente.

1. Tablero RO (Resultado Operativo)

Este tablero contiene información de un evolutivo de ventas mostrado a través de un gráfico de barras con sus diversos conceptos de venta prevista, venta presupuestada y acumulado. A su vez se muestra un resultado consolidado del resultado operativo anual por venta, utilidad y margen de ganancia aperturado por los diversos conceptos. La figura 43 muestra el evolutivo de ventas comentado, el cual es mostrado a nivel de millones de dólares.

Figura 43

Ventas por Proyecto - Gráfico de barras evolutivo por tipo de venta



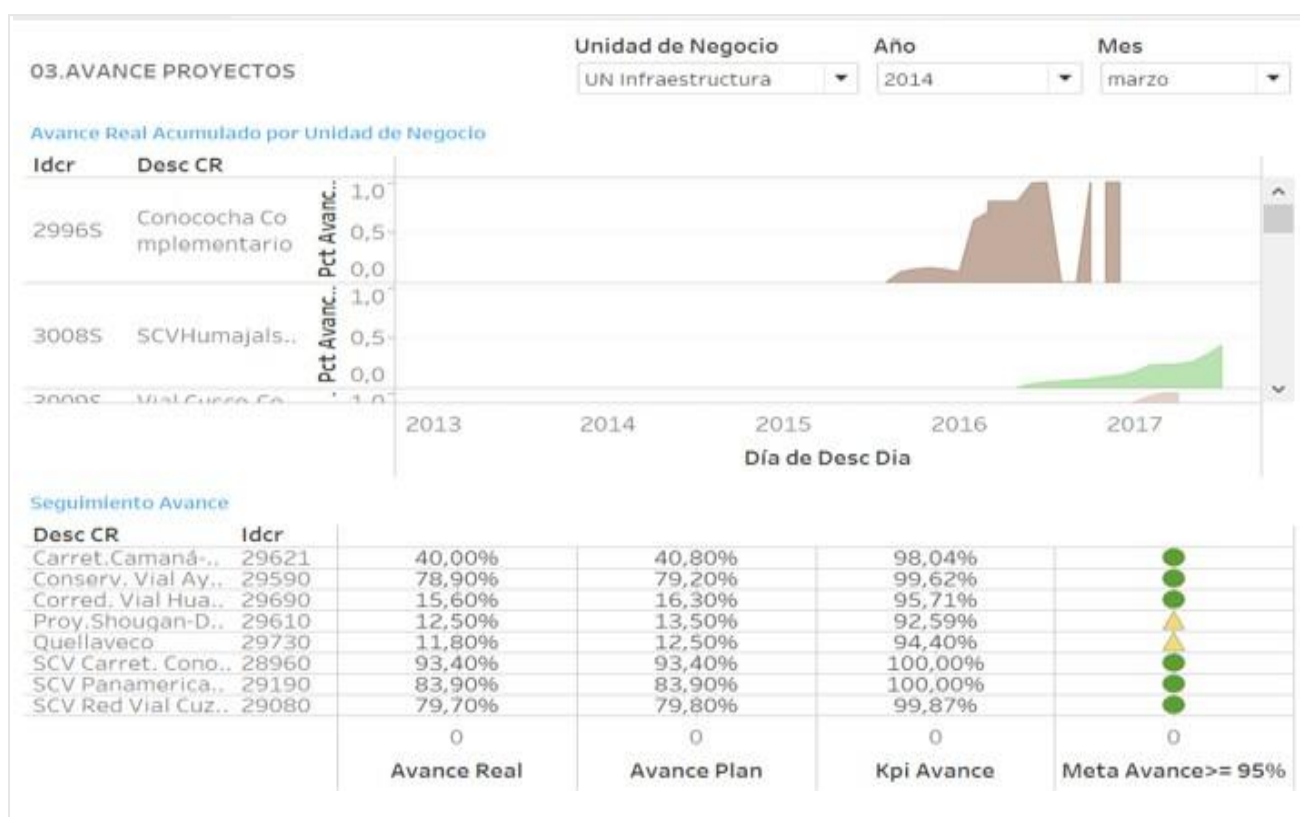
Nota. Fuente: Elaboración propia

2. Tablero Avance

El tablero Avance Proyectos mostrado en la Figura 44 contiene dos reportes importantes, el primero de ellos muestra el avance real de los proyectos a lo largo de su vida útil. El segundo es un reporte más específico que permite monitorear el avance respecto a lo planeado, apoyado en semáforos indica si se cumple lo planteado.

Figura 44

Tablero Avance Proyectos



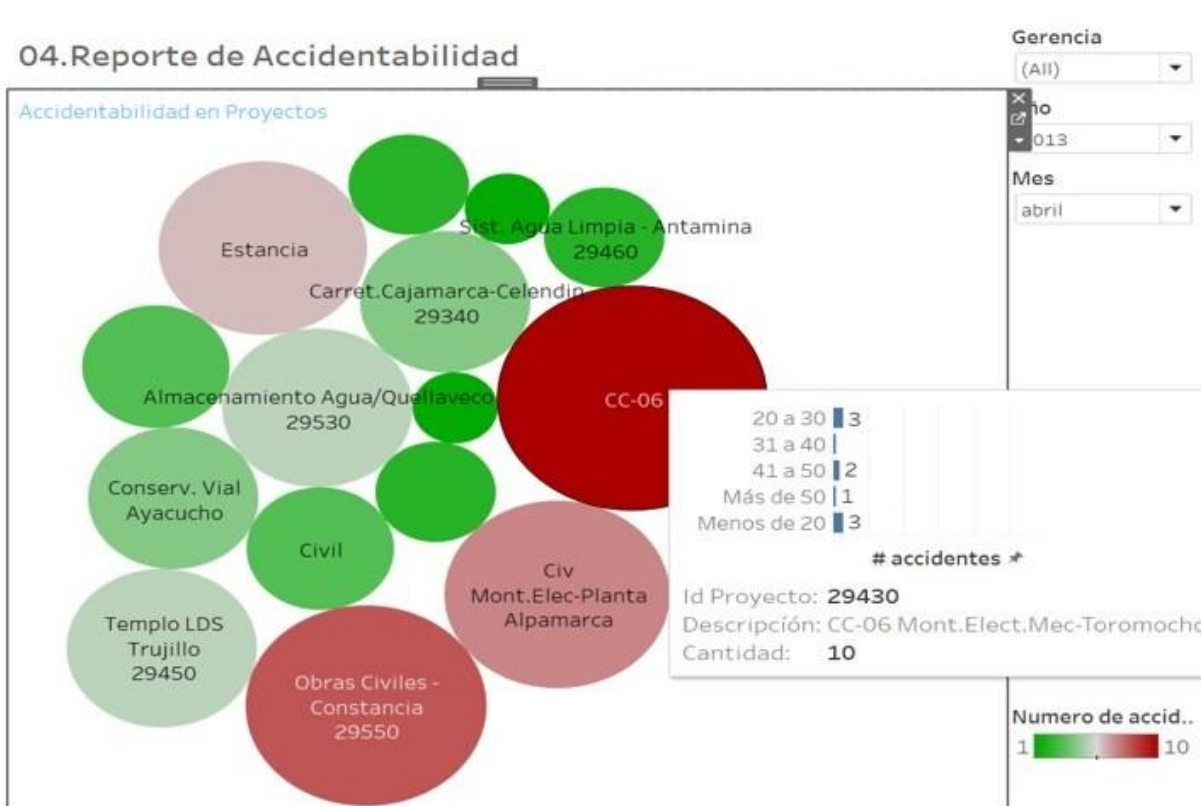
Nota. Fuente: Elaboración propia

3. Tablero Accidentabilidad

El tablero mostrado en la Figura 45 se enfoca en mostrar la accidentabilidad de los proyectos, apoyados en un gráfico de burbujas se identifica con facilidad cuál es el proyecto con mayor accidentabilidad, a su vez se obtiene el detalle adicional de rango de edades de cada accidente sucedido en el proyecto.

Figura 45

Tablero de Accidentabilidad en Proyectos



Nota. Fuente: Elaboración propia

Aplicación de Fase VI: Despliegue

Para la implementación del proyecto se realizó un plan de pruebas que asegure el correcto funcionamiento de los componentes de la solución de inteligencia desarrollada a lo largo de este capítulo.

Plan de Pruebas

Se definieron los casos de prueba, estos fueron registrados y se comparó el resultado obtenido con la salida proyectada para dicha entrada, de esta manera se determinó si cada objeto de prueba cumple con los criterios de calidad. Los criterios de aceptación para los casos de prueba se muestran en la Tabla 16.

Tabla 16

Plan de Pruebas - Criterios de aceptación

Objeto de Prueba	Proceso Involucrado	Criterio
Procesos ETL y Base de Datos	Extracción	Los datos extraídos de la base de datos fuente deben ser el mismo que los datos cargados en la base del Datawarehouse
Servicios OLAP	Explotación	La información del análisis final del usuario debe ser igual a los datos del datawarehouse.

Nota. Fuente: Elaboración propia

Determinación de los casos de prueba

Tabla 17

Plan de Pruebas – Requerimiento Equipo

Caso de Prueba : Requerimiento Equipo	
Objetivo de la prueba:	Revisar ETL y carga de Datamart
Objetivo de la prueba:	Verificar que la tabla [BI_F_Requerimiento_Equipo]teniendo como tablas fuente las tablas de Proyectos y tabla de requerimientos.
Precondición	Disponibilidad de la base de datos fuente
Descripción de la Prueba	Probar la cantidad de registros de las tablas fuente frente a la cantidad de registro del datamart
Resultados Esperados	La tabla BI_F_Requerimiento_Equipo muestra la misma cantidad de documentos que la tabla de detalle de importaciones expresa en la tabla de Requerimientos.
Resultado Obtenido	Se cargaron los datos correctamente.

Nota. Fuente: Elaboración propia

Se diseñó la plantilla mostrada en la Tabla 17, la cual ayudó a especificar los casos de prueba que conforman el plan de pruebas del proyecto. A modo de ejemplo se realiza la prueba del modelo Requerimiento Equipo.

Ejecución y análisis de los casos de prueba

Una vez definidos los casos de prueba, se ejecutó cada una de ellos. Se registraron los resultados positivos o negativos de la prueba, aquí cabe recalcar que cualquier error podría ser sujeto a correcciones hasta lograr el resultado esperado.

Continuando con el ejemplo de carga de la tabla BI_F_REQUERIMIENTO_EQUIPO. Se muestra en la Figura 46 el proceso de carga satisfactoria ETL.

Figura 46

Plan de Pruebas - Carga de datos al Stage mediante ETL



Nota. Fuente: Elaboración propia

Luego de correr el paquete ETL seguido de la ejecución de PROC_BI_F_Requerimiento_Equipo que cruza información de la carga del ETL mostrada y de los proyectos obtenidos de la base operativa. Se observa en la Figura 47 que la tabla Fact Requerimiento Equipo fue cargada satisfactoriamente.

Plan de Pruebas - Carga de datos al DW

`SELECT * from [dbo].[BI_F_Requerimiento_Equipo]`

100 %

Results Messages

	IDPAF	LineaSolicitud	IDFase	IDFechaRequerimientoEQ	IDFechaSolicitudEQ	IDFechaNTAFEQ
1	562803	0		2013-03-28 05:00:00.000	2013-03-28 05:00:00.000	1901-01-01 05:00:00.000
2	650305	0		2013-10-25 05:00:00.000	2013-05-07 05:00:00.000	1901-01-01 05:00:00.000
3	701	0		2013-07-02 05:00:00.000	2013-07-02 05:00:00.000	1901-01-01 05:00:00.000
4	1	0		2013-08-10 05:00:00.000	2013-08-02 05:00:00.000	1901-01-01 05:00:00.000
5	20	0		2013-04-25 05:00:00.000	2013-04-22 05:00:00.000	1901-01-01 05:00:00.000
6	363	0		2013-06-06 05:00:00.000	2013-06-06 05:00:00.000	1901-01-01 05:00:00.000
7	652504	0		2013-04-28 05:00:00.000	2013-04-28 05:00:00.000	1901-01-01 05:00:00.000
8	698437	0		2013-04-23 05:00:00.000	2013-04-23 05:00:00.000	1901-01-01 05:00:00.000
9	583105	0		2013-05-31 05:00:00.000	2013-05-31 05:00:00.000	1901-01-01 05:00:00.000
10	451	0		2013-07-26 05:00:00.000	2013-06-17 05:00:00.000	1901-01-01 05:00:00.000
11	422006	0		2013-06-20 05:00:00.000	2013-06-20 05:00:00.000	1901-01-01 05:00:00.000

Nota. Fuente: Elaboración propia

CAPÍTULO 6: ANÁLISIS DE RESULTADOS

En esta capítulo se presenta los resultados obtenidos a partir de la data consolidada, relacionada a los proyectos de la empresa constructora, mediante la solución de inteligencia de negocios desarrollada en los capítulos anteriores. Estos resultados son comparados con los resultados obtenidos de forma tradicional, es decir se hace una comparación del antes y después de la implementación de inteligencia de negocios. Esta comparación está basada en el tiempo de procesamiento de los indicadores establecidos en los capítulos anteriores.

Resultado de los indicadores por categoría

Categoría Costo

En esta categoría se ha considerado el monitoreo de métricas correspondientes a la venta, utilidad y margen de los proyectos y su gestión de la operación total de la empresa constructora. Si bien la información de costos y ventas estaba disponible antes de la implementación de la solución de inteligencia de negocios, esta no era presentada a este nivel de detalle comparativo con aperturas de seguimiento, en consecuencia, es un aporte que permite determinar ventas y ganancias en las diversas unidades de negocio de la empresa en un determinado periodo y en diferentes monedas; con un impacto importante para la toma de decisiones financieras de los proyectos.

Figura 48

Resultados del índice de margen de ganancia

01.RO GENERAL

Unidad .. (All) Moneda Soles Año 2017 Mes abril

Resultado Operativo(RO) Anual

	Venta Anual			Utilidad Bruta Anual			Margen Bruto		
	Acumulado	Previsto	Presupue..	Acumulado	Previsto	Presupue..	Acumulado	Previsto	Presupue..
Proyectos en Ejecución,,	345,078	1,220,314	832,778	12,750	71,104	62,630	3,69%	5,83%	7,52%
Proyectos Adjudicados,,	0	27,618	136,420	0	2592	10002		9,39%	7,33%
Servicios a Consorcios,,	27,607	62,487	54,329	7,777	10,727	10,223	28,17%	17,17%	18,82%

Nota. Fuente: Elaboración propia

En la Figura 48, podemos observar que al corte de Abril del 2017, el resultado operativo de los proyectos en ejecución tiene un margen bruto acumulado de 3,69%, es decir estamos a un 51% de alcanzar el margen presupuestado anual de 7.55%. Esta información se puede desplegar también por proyecto también, teniendo la opción de hacer un análisis más enfocado en proyectos requeridos.

Categoría Tiempo

En esta categoría se ha considerado el índice de avance en los proyectos por unidad de negocio considerando las medidas del avance real y planificado. Esta información no estaba disponible antes de la implementación de la solución de inteligencia de negocios, en consecuencia, es un aporte que permite determinar el avance de un determinado proyecto en el tiempo, con un impacto importante para la toma de decisiones del plan de actividades, permitiendo visibilizar alertas de retraso y cumplimiento, a su vez visibilizar el ciclo de vida de los proyectos a lo largo de su ejecución.

Figura 49

Resultados del índice de avance de proyecto

Seguimiento Avance				
Desc CR	Idcr			
Ampl. Fase 2A DPWo...	30410	7,90%	8,90%	88,76%
CentralTerm Obras...	INT1003	32,10%	39,60%	81,06%
Clínica Internacional	30020	96,40%	97,70%	98,67%
Cons. Muro Anclado...	30380	3,00%	3,20%	93,75%
Conserv. Vial Ayacu...	29590	98,30%	98,40%	99,90%
ConstCaminoAcceso...	30320	69,00%	81,80%	84,35%
Corred. Vial Huanca...	29690	88,80%	89,80%	98,89%
Edificio El Parque	29980	96,70%	99,20%	97,48%
Edificio Primera Visi...	30220	33,00%	33,80%	97,63%
Mejoram. Av. Sanch...	30130	5,00%	5,30%	94,34%
Metro Línea 2-Serv...	29783S	6,70%	51,00%	13,14%
Montaje Area1 PMR...	3014S	37,20%	60,20%	61,79%
OCC y OEM HPGR y...	30350	18,30%	22,50%	81,33%
Obra Civil Area Seca...	30210	79,90%	86,80%	92,05%
Pachacutec - Etapa 2...	30100	90,20%	90,80%	99,34%
Quellaveco	29730	99,90%	85,90%	116,30%
REUB INTE RED EST...	30360	3,70%	4,80%	77,08%
Reub.&Alcan.Est.1,2...	30390	3,20%	3,70%	86,49%
Reub.&Alcan.Est.5,6...	30440	1,30%	1,30%	100,00%
Reub.Agua&Alcant...	30260	71,50%	75,00%	95,33%
SCV Chivay - Arequip...	30050	22,70%	24,70%	91,90%
SCVHumajalso-Tacn...	3008S	25,80%	38,40%	67,19%
Templo Arequipa-Pe...	30170	0,20%	0,10%	200,00%
Templo Arequipa-Pe...	30171	2,20%	1,30%	169,23%
		0	0	0
		Avance Real	Avance Plan	Kpi Avance
				Meta Avance>= 95%

Nota. Fuente: Elaboración propia

En la Figura 49 se presentan los proyectos gestionados en el periodo de Abril 2017 por la totalidad de unidades de negocios de la empresa, cada uno de estos detalla su respectivo índice de seguimiento de avance. De los 24 proyectos vigentes, solo un 16% cumplió la meta de avance, con ello los gerentes se enfocan en la revisión de las alertas.

Categoría Seguridad y Salud

En esta categoría se ha considerado el índice de accidentabilidad en los proyectos por unidad de negocio en el año 2017. En la tabla 18 se presentan los proyectos con mayor nivel de índice de accidentabilidad por cada unidad. Esta información no estaba disponible antes de la implementación de la solución de inteligencia de negocios, en consecuencia, es un aporte que permite determinar el costo asociado al número de accidentes por proyecto con un impacto importante para la toma de decisiones.

Tabla 18
Resultados del Índice ATP

Unidad de Negocio	Proyecto	# Accidentes	Edad lesionada
Infraestructura	29530 - Quellaveco	29	< 20 años
Edificaciones	29450- Templo Trujillo	27	41 - 50 años
Plantas Industriales	CC-06 Mon.Elec- Toromocho	54	20-30 años

Nota. Fuente: Elaboración propia

Análisis comparativo

En esta sección comparamos el tiempo de procesamiento tradicional, el cual solo consideraba el cálculo de indicadores de la categoría costo, con el tiempo de procesamiento luego de implementarse la solución de inteligencia de negocios desarrollada a lo largo de esta tesis. Cabe recalcar que aparte de los indicadores de la categoría costo, la solución BI incluye categorías de tiempo y seguridad y salud.

La Tabla 19 muestra el comparativo del tiempo usado para obtener los indicadores de monitoreo de proyectos de construcción, antes y después de la solución de inteligencia de negocios.

Tabla 19

Tiempo usado para obtener los indicadores, antes y después de la solución BI

(Data en minutos)		
Meses	Antes	Después
1	620	25
2	510	22
3	630	20
4	660	28
5	720	35
6	600	30
Promedio	623	26

Nota. Fuente: Elaboración propia

Como se observa con el procesamiento tradicional en promedio se invertía unos 623 minutos en promedio, mientras que con la solución de inteligencia de negocios se obtienen los indicadores de desempeño de proyectos en un tiempo promedio de 26 minutos.

De igual modo, se va a comparar el costo asociado a las horas hombre que se utilizaba antes y después de la solución de inteligencia de negocios, cabe recalcar que el costo no considera licencias ni herramientas de desarrollo, ya que la empresa ya contaba con dichos recursos.

Antes del sistema de inteligencia de negocios, este trabajo en promedio era realizado por 6 personas, las cuales trabajaban de manera simultánea con costo promedio de hora hombre es de S/. 23.80.

$$\text{Costo anterior a solución} = 23.80 * 6 * \frac{1h}{60 \text{ min}} * 623 \text{ min} = \text{S/. } 1482$$

Tras la implementación del sistema de inteligencia de negocios, solo se necesita a una persona que supervise la ejecución trasladándose al área de operaciones con costo promedio de S/. 120 por hora hombre.

$$\text{Costo posterior a solución} = 120 * 1 * \frac{1h}{60 \text{ min}} * 26 \text{ min} = \text{S/. } 52$$

Como se observa con el procesamiento tradicional en promedio se invertía en promedio S/.1482, mientras que con la solución de inteligencia de negocios se obtienen los indicadores de desempeño de proyectos con un costo de S/52.

En conclusión, el tiempo de procesamiento se redujo en 95,8 % tras la implementación del sistema basado en business intelligence y el costo asociado a horas hombre se redujo en 96,5 %, existiendo diferencias significativas en la obtención de los indicadores de monitoreo de proyectos de la Gerencia de Control de Proyectos. Esta mejora permite que los ejecutivos se centren en diversos análisis y trabajos de mayor valor que apoyen la toma de decisiones de forma efectiva y eficiente.

CONCLUSIONES, RECOMENDACIONES Y TRABAJOS FUTUROS

Conclusiones

a) Conclusión de Objetivo 1

Como resultado del Benchmarking de Metodologías de Business Intelligence realizada en esta investigación, se optó por utilizar la Metodología de Kimball, ya que se validó que es una metodología de fácil uso, de desarrollo incremental con un tiempo de despliegue corto y sobre todo por involucrar a los usuarios finales desde etapas tempranas y en cada etapa de la metodología. A su vez el enfoque de Kimball se adapta fácilmente a la forma de trabajo y experiencia de la autora de esta investigación.

b) Conclusión del Objetivo 2

Como resultado del cuadro comparativo de Herramientas de Business Intelligence realizada en esta investigación, se optó por utilizar la herramienta Tableau, ya que tiene mejor nivel de usabilidad, posee la ventaja de arrastrar y soltar los componentes con el objetivo de crear y desarrollar fácilmente tableros de control, reportes, entre otros. Posee una gran comunidad que permite la colaboración eficaz entre usuarios y generación de nuevos recursos.

c) Conclusión del Objetivo 3

Como resultado de la revisión de la literatura, para esta tesis se seleccionó las categorías de indicadores: Costo, Tiempo y Seguridad y Salud; ya que una de las prioridades de la gerencia de Control de Proyectos de Construcción es asegurar que los proyectos se lleven a cabo en el plazo estipulado y bajo los costos planificados. A su vez cumplir con los estándares de calidad de maquinarias y equipos utilizados en los proyectos de construcción; y finalmente asegurar la seguridad y salud de sus trabajadores.

d) Conclusión del Objetivo 4

En base a las fuentes de información alojadas en SQL Server y Excel, se modeló e implementó 4 datamarts: Modelo CP Avance, Modelo CP Margen, Modelo Requerimiento Equipo y Modelo ATP, los cuales constituyen el repositorio unificado datawarehouse. Cada modelo contiene métricas relacionadas a las categorías de indicadores a monitorear los cuales y dimensiones que

permitieron analizar de diversas perspectivas el monitoreo de proyectos de construcción. A su vez se implementaron dashboards que presentan estos indicadores en el tiempo y al adicionarle un semáforo que permite visualizar alertas, se ha mejorado el análisis de la información de los directivos, al realizar un diagnóstico rápido de la situación de los proyectos.

e) Conclusión del Objetivo 5

Se optó por validar los tiempos de procesamiento tanto de forma tradicional y los obtenidos con la solución de business intelligence, para ello se tomó una muestra de seis meses con ello se pudo comparar los tiempos obtenidos. Se comprobó que tras la implementación de business intelligence, el tiempo de procesamiento tradicional se redujo en 95,8% permitiendo a los ejecutivos realizar trabajos de mayor valor para la toma de decisiones eficientes y efectivas.

Recomendaciones

Recomendación de Conclusión 3: Se recomienda considerar la adición de nuevas categorías de indicadores, de tal modo que el Datawarehouse abarque información más rica de los proyectos. En base a la información existente en la base de datos se sugiere la adición de las categorías Gestión de la Calidad y Gestión de Carga Administrativa.

Recomendación de Conclusión 4: Se recomienda repotenciar los servidores de la compañía, ya que según necesidades específicas se irán incorporando más datamarts, a su vez los ordenadores del Área de Control de Proyectos para que puedan construir nuevos reportes oportunamente.

Recomendación de Conclusión 5: Se recomienda con respecto al datawarehouse, se recomienda probar con diferentes herramientas para procesar los datos en la fase de ETL, de tal forma se pueda mejorar el tiempo de procesamiento de la información.

Trabajos Futuros

Capacitar a los usuarios del negocio con el uso de la herramienta Tableau, de tal forma que puedan autogestionar su información al construir nuevos reportes analíticos de la información ya contenida en el repositorio dimensional.

Realizar un sistema de registro de información relacionada a los avances y consumo de materiales de proyectos, cuyos datos actualmente se registran en Excel, de tal modo que la información se registre en la base de datos transaccional corporativa.

BIBLIOGRAFÍA

1. Acosta, V. M. (2019). Descubre la principal diferencia entre datamart y datawarehouse. *Revista digital INESEM - Informatica y tics*.
2. Albarran, T. S., & Salgado Gallegos, M. (2013). La inteligencia analitica y la competitividad de las empresas. *RECAI Rev. Estud. en Contaduría, Adm. e Informática*.
3. Andreau, R., Ricart, J., & Valor, J. (1991). *Estrategia y Sistemas de Información*. Madrid: Mc Graw-Hill.
4. Bernabeu, R. D. (Enero de 2009). *DataPRIX*. Obtenido de DATA WAREHOUSING: Investigación y Sistematización de Conceptos: dataprix.com/files/DWH_Metodologia_HEFESTO-V1.0.pdf
5. Bocevska, A., Savoska, S., & Milevski, I. (2017). BI Tools Analysis According to Business Criteria as Data Integration Possibilities, Hardware Specification, Tools for Data Visualization and Comparison of Used Technologies. *Proceedings of the 11th International Conference on Information Systems and Grid Technologies*, (págs. 79-90). Macedonia.
6. Boulila, W., Al-kmali, M., Farid, M., & Mugahed, H. (2018). A business intelligence based solution to support academic affairs: case of Taibah University. *Springer Science+Business Media*, 1-8.
7. Bouza Betancourt, O., Couto Torres, D., & Sosa Perez, N. (2017). Evaluación del estado de la gestión de la información científica y tecnológica: dimensiones e indicadores. *Revista Cubana de Información en Ciencias de la Salud* 2017, 1-17.
8. Chalco Valencia, & Galo Andrés. (2018). *Implementación de Business Intelligence en la Empresa Corporación Líderes*. Tesis de grado, Pontificia Universidad Católica del Ecuador, Ecuador.
9. Cheung, S., Suen, H., & Cheung, K. (2004). PPMS: a Web-based construction Project Performance. *Automation in Construction*, 361-376.
10. Contreras de La Fuente, R. (Agosto , 2018). “CMMI DEV 2 COMO BASE DEL PROCESO ETL PARA LA GENERACIÓN DE CONOCIMIENTO EN LA INTELIGENCIA DE NEGOCIOS. APIZACO TLAXCALA , Mexico: TECNOLÓGICO NACIONAL DE MÉXICO INSTITUTO TECNOLÓGICO DE APIZACO.
11. COPECO. (Noviembre de 2018). *Proyecto Gestión Copeco*. Obtenido de <https://www.pgrd-copeco.gob.hn/wp-content/uploads/2018/11/05-04-2016-manual-constructor.pdf>
12. Cordero-Guzmán, D., & Rodríguez-López, G. (2017). La inteligencia de negocios: una estrategia para la gestión de las empresas productivas. *Revista Ciencia UNEMI*, 40-48.
13. Diez-Silva, M., Pérez-Ezcurdia, A., Gimena Ramos, F., & Montes-Guerra, M. (2019). Medición del desempeño y éxito en la dirección de proyectos. *Rev. esc.adm.neg*, 60-79.
14. Espinoza, R. (25 de Febrero de 2010). *DataPRIX*. Obtenido de Herramientas ETL. ¿Que son, para que valen?. Productos mas conocidos. ETL's Open Source.:

<https://www.dataprix.com/es/blog-it/respinosamilla/herramientas-etl-son-valen-productos-mas-conocidos-etls-open-source>

15. Feliciano Morales, A., Cuevas Valencia, R., & Martínez Castro, J. (2016). Procesamiento Analítico con Minería de Datos. *Revista Iberoamericana de las Ciencias Computacionales e Informática*(ISSN: 2007-9915), 1-22.
16. Fernandez, C. (Agosto de 2021). *Data Prix*. Obtenido de Query y reporting en un Data Warehouse: <https://www.dataprix.com/es/query-y-reporting-un-data-warehouse>
17. García Orozco, Javier Francisco. (2010). *Gestión de la información y el conocimiento Observatorio para la educación en ambientes virtuales*. México: UDG Virtual.
18. Gil-Montelongo, M., López-Orozco, G., Molina-García, C., & Bolio-Yris, C. (2011). LA GESTIÓN DE LA INFORMACIÓN COMO BASE DE UNA INICIATIVA DE GESTIÓN DEL CONOCIMIENTO. *Redalyc.org*, 231-237.
19. Gordillo-Otárola, V. (2014). *Evaluación de la Gestión de los proyectos en el Sector Construcción del Perú*. Piura,Perú: Universidad de Piura (PIRHUA). Obtenido de https://pirhua.udep.edu.pe/bitstream/handle/11042/2051/MAS_PRO_006.pdf%3Bsequence=1
20. Hernandez Trasobares, A. (2013). LOS SISTEMAS DE INFORMACIÓN: EVOLUCIÓN Y DESARROLLO. *Dialnet*, 149-165.
21. Holescher, J., & Mortimer, A. (September de 2018). Using Tableau to visualize data and drive decision-making. *Journal of Accounting Education*, págs. 49-59. doi:<https://doi.org/10.1016/j.jaccedu.2018.05.002>
22. INEI. (2021). *Informe Técnico- Producto Bruto Interno Trimestral*. Lima: INEI
23. Kimball, R. (2008). *The Data Warehouse Lifecycle Toolkit. 2nd Edition*. New York: Wiley Publishing.
24. Labiod, A., & Yessad, L. (2016). Comparative Study of Data Warehouses Modeling Approaches: Inmon, Kimball. *2016 International Conference on System Reliability and Science* (págs. 95-99). Algeria: IEEE.
25. Ling, F. (2004). How project managers can better control the performance. *International Journal of Project Management*, 477-488.
26. M.Strong, D., W.Lee, Y., & Y.Wang, R. (August 2012). "Beyond Accuracy : What Data Quality Means to Data Consumers". *I2*(4), 103-110.
27. Mazon, R., Pardillo, L., & Trujillo, J. C. (2011). *Diseño y Explotación de almacenes de datos: Conceptos*. Editorial Club Universitario.
28. Microsoft. (6 de 6 de 2021). *Power BI*. Obtenido de <https://powerbi.microsoft.com/es-es/customer-showcase/>
29. MINSA. (06 de 06 de 2021). *Plataforma Única del Estado Peruano*. Obtenido de <https://www.gob.pe/institucion/minsa/normas-legales/356904-642-2019-oga-minsa>

30. Mora, G. (2018). Siglo XXI economía de la información: gestión del conocimiento y Business Intelligence, el camino a seguir hacia la competitividad. *Signos*, 161-174.
31. Muñoz Hernandez , H., Osorio-Mass, R., & Zuñiga Pérez, L. (2016). Inteligencia de los negocios. Clave del éxito en la era de la información. Clio América.
32. Novoa Torres, N., Bermudez Huérfano, D., & Zamora Carrillo, H. (Diciembre 2019). Nociones, consideraciones y ventajas de la inteligencia de negocios BI. *Revista Vinculos*, 280-281.
33. Ormeño, Y. (2 de Mayo de 2020). *Informe Situacional del Sector Construcción en el Perú Post Covid-2019*. Obtenido de Project Management Institute Lima, Perú: <https://pmi.org.pe/informe-situacional-del-sector-construccion-en-el-peru-post-covid-2019/>
34. Parama Fadli Kurnia, S. (2018). Business Intelligence Model to Analyze Social Media Information. *3rd International Conference on Computer Science and Computational Intelligence 2018*, (págs. 5 - 14). Jakarta,Indonesia.
35. PMI. (2021). *Project Management Institute*. Recuperado el 8 de Mayo de 2021, de <https://www.pmi.org/>
36. Power Data. (6 de Junio de 2017). *Power Data*. Obtenido de El valor de la gestión de los datos: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/qu-son-los-procesos-etl>
37. *Power Pro Consulting*. (07 de 07 de 2021). Obtenido de <https://datadogconsulting.wixsite.com/website/post/2017-04-21-breve-historia-de-microsoft-power-bi>
38. Puentes Neira, A., & Guevara Romero, C. (2015). *Indicadores de Desempeño en la gestión de proyectos, un análisis del estado del arte basado en las publicaciones científicas actuales*. Bogota,Colombia: UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS.
39. Rimarachin, W. (2015). SISTEMA DE INFORMACIÓN EJECUTIVO BASADO EN BUSINESS INTELLIGENCE Y LA CALIDAD DE INFORMACIÓN DE LOS INDICADORES. 98.
40. Rivadera, R. (2010). La metodología de Kimball para el diseño de almacenes de datos (Data Warehouses). *Revista Ucasal*, 56-71.
41. Rodriguez Alfaro , F. (22 de 5 de 2021). *Data Management*. Obtenido de <https://datamanagement.es/2019/06/27/business-intelligence-modelo-estrella/>
42. Rudeli, N., Vilez, E., Gonzales, J., & Santilli, A. (2018). Causas de Retrasos en Proyectos de Construcción: Un análisis cualitativo. *Memoria Investigación de Ingeniería*, 78-80.
43. Rimvydas Skyrius, G. (2012). Business Intelligence in the Process of Decision Making: Changes and Trends. *Ekonomika*, 147-157.
44. Santos, M., Oliveira e Sá, J., Andrade, C., Vale Lima , F., Costa, E., Costa , C., . . . Galvao, J. (2017). A Big Data system supporting Bosch Braga Industry 4.0 strategy. *International Journal of Information Management*, 750-760.
45. Shukla, A., & Dhir, S. (2016). Tools for Data Visualization in Business Intelligence: Case Study Using the Tool Qlikview. *S.C. Satapathy et*, 319-326. doi: 10.1007/978-81-322-2752-6_31.

46. Suarez Alfonzo, A., Cruz Rodriguez, I., & Perez Macias, Y. (2015). INFORMATION MANAGEMENT: AN ESSENTIAL TOOL FOR SKILLS DEVELOPMENT IN. *Revista Universidad y Sociedad [seriada en linea]*, pp. 72-79. Obtenido de <http://rus.ucf.edu.cu/>
47. Suganda Girsang, A., Muhamad Isa, S., Saputra, H., Apriadin Nuriawan, M., Putra Ghazali, R., & Kaburuan, E. (2018). Business Intelligence for Construction Company Acknowledgement Reporting System. *The 1st 2018 INAPR International Conference* (págs. 113-122). Jakarta, Indonesia: IEEE.
48. Tableau. (07 de 07 de 2021). Obtenido de <https://www.tableau.com/es-es/why-tableau/what-is-tableau>
49. Universidad Pontificia de Salamanca. (2006). IV SIMPOSIO INTERNACIONAL DE SISTEMAS DE INFORMACIÓN E INGENIERIA DE SOFTWARE EN LA SOCIEDAD DEL CONOCIMIENTO. *SISOFT2006* (pág. 183). Colombia: Universidad Pontifica de Salamanca campus de Madrid (España).