

Universidad Católica de Santa María
Facultad de Ciencias e Ingenierías Físicas y
Formales
Escuela Profesional de Ingeniería de Sistemas



**MEJORA DEL PROCESO DE PRODUCCIÓN DE UNA EMPRESA PRODUCTORA
DE HARINAS APLICANDO INTELIGENCIA DE NEGOCIOS CON
HERRAMIENTAS OPEN SOURCE**

Tesis presentada por el Bachiller:

Montánchez Montesinos, Arturo Renato

Para optar el Título Profesional de

**Ingeniero de Sistemas: Especialidad en Sistemas de
Información**

Asesor:

Mg. Zúñiga Carnero, Manuel Mariano

Arequipa- Perú

2018

ACTA TITULO PROFESIONAL

ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS



En Arequipa, a los 30 días del mes de OCTUBRE del 2018, siendo las 18:30 horas, en el local

de la Universidad Católica de Santa María, se reunió el Jurado

Presidido por: ING. GUILLERMO E. CALDERON RUIZ

Integrantes: ING. HÉCTOR R. VELARDE BEDREGAL

ING. MANUELA M. ZÚÑIGA CARWEN

Actuando este último como Secretario con la finalidad de recibir las previas orales del(os) (la) (s)

Señor(ita) Bachiller(s)

MONTANCHER MONTESINOS ARTURO RENATO

Quien(es) pretende(n) optar el Título Profesional de INGENIERO DE SISTEMAS: ESPECIALIDAD EN SISTEMAS DE INFORMACION

Sustentando: LA TESIS

Titulado: MEJORA DEL PROCESO DE PRODUCCIÓN DE UNA EMPRESA

PRODUCTORA DE HARINAS APLICANDO INTELIGENCIA DE

NEGOCIOS CON HERRAMIENTAS OPEN SOURCE "

El Presidente del Jurado invitó al (los) Titulando(s) a hacer una exposición de su trabajo, conclusiones y recomendaciones, para luego proceder a realizar las preguntas que los Miembros del Jurado consideraron pertinente plantear. Posteriormente, se pasó a deliberar y emitir su voto en la forma establecida por el Reglamento de Grados y Títulos de la Facultad de Ciencias e Ingenierías Físicas y

Formales siendo el resultado el siguiente:

APROBADO POR UNANIMIDAD

Con lo que se dio por terminado el Acto siendo las 21:00 horas, para dar fe firmamos a continuación los Miembros del Jurado y el (los) Titulado(s).

[Signature] PRESIDENTE

[Signature] INTEGRANTE

[Signature] SECRETARIO

[Signature] TITULANDO

[Signature] TITULANDO

FACULTAD DE CIENCIAS E INGENIERIAS FISICAS Y FORMALES
ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS

INFORME DICTAMEN DE BORRADOR TESIS

VISTO

El Borrador de TESIS titulado:

"MEJORA DEL PROCESO DE PRODUCCION DE UNA EMPRESA PRODUCTORA
DE HARINAS APLICANDO INTELIGENCIA DE NEGOCIOS CON
HERRAMIENTAS OPEN SOURCE"

Presentado por (el) (la) (los) Bachiller (es):

MONTANCHEZ MONTESINOS ARTURO RENATO

Nuestro dictamen es:

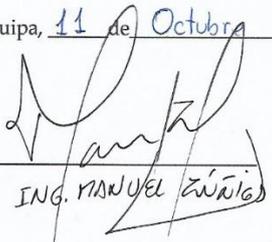
Aprobado.

OBSERVACIONES:

Arequipa, 11 de Octubre de 2018



1742



ING. MANUEL ZÚÑIGA c.

PRESENTACIÓN

Señor Decano de la Facultad de Ciencias e Ingenierías Físicas y Formales.

Señor Director de la Escuela Profesional de Ingeniería de Sistemas.

Señores Miembros del Jurado Dictaminador de la Tesis

De conformidad con las disposiciones del Reglamento de Grados y Títulos de la Escuela Profesional de Ingeniería de Sistemas, pongo a vuestra consideración el presente trabajo de investigación titulado:

MEJORA DEL PROCESO DE PRODUCCIÓN DE UNA EMPRESA PRODUCTORA DE HARINAS APLICANDO INTELIGENCIA DE NEGOCIOS CON HERRAMIENTAS OPEN SOURCE

El trabajo de investigación fue realizado aplicando los conocimientos adquiridos durante mi formación universitaria, el mismo que al ser aprobado me permitirá optar por el Título Profesional de Ingeniero de Sistemas.

DEDICATORIA

Este trabajo está dedicado a mis padres, a mi esposa y a toda mi familia, gracias a los cuales pude realizar este proyecto recibiendo su constante apoyo y motivación durante toda esta fase, siempre impulsándome a mejorar cada día.



INTRODUCCION

Las actividades de las personas, mantienen intervenciones de productos, ya sean bienes y servicios, los mismos que son empleados para la satisfacción de sus necesidades, sin embargo, para lograr estos productos, se debe de encontrar organizaciones que los produzcan y pongan a disposición del mercado. Dentro de este proceso de producción, se plasma información que permite controlar, medir y tomar decisiones que aportarán a las metas u objetivos planteados por cada organización.

Entonces, toda empresa genera una variable cantidad de datos con tan sólo estar en actividad. Convertir esos datos en información, es de muy importante para cualquier organización que desee desarrollarse y ser competitiva.

En la industria moderna, gran parte de máquinas y equipos están implementados con dispositivos electrónicos que permiten un control remoto y el monitoreo de sus estados y variables, permitiendo el acceso a datos que pueden ser de alto valor.

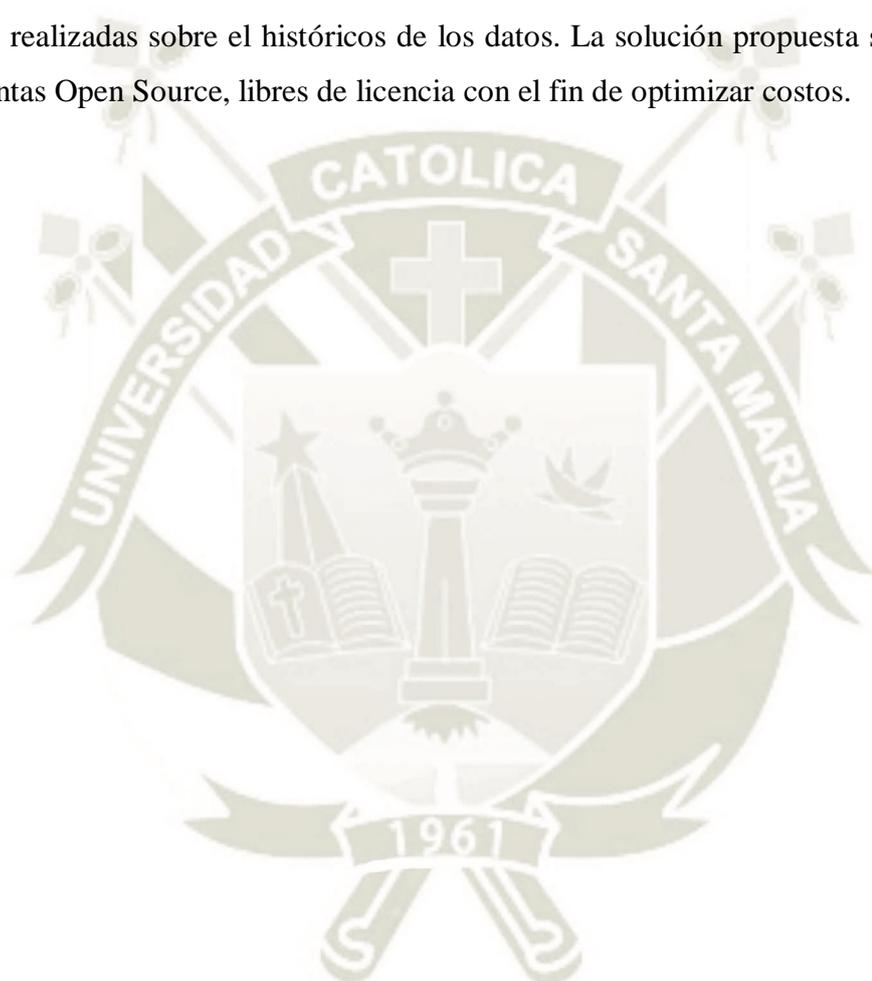
Las organizaciones líderes están en la búsqueda constante de caminos novedosos e inteligentes para mejorar su desempeño en varios aspectos dentro de la misma, haciéndose más eficientes; y a su vez esperan que cada individuo sea capaz de aportar para estos resultados.

La inteligencia de negocios o BI es un factor clave que permite alcanzar esos resultados, ya que permite una toma de decisiones informada en todos los niveles de la organización, permitiendo a cada uno de estos actuar de la manera más adecuada y efectiva en una situación dada. Aun así, para cada empresa representa un desafío encontrar la mejor solución para obtener, visualizar y analizar su información y que a su vez se adecue a sus recursos y presupuesto.

En las empresas de producción, es muy común encontrar implementados sistemas SCADA que les permite, además del control remoto, visualizar el estado actual y funcionamiento de sus equipos, así como captar datos específicos de los equipos que conforman la industria. Sin embargo, la visualización de datos, el acceso a un histórico visualmente amigable o poder

analizar múltiples conjuntos de información no es posible ya que el sistema SCADA no tiene dicho fin.

Es así que identificada la falta de información precisa y confiable en una empresa productora de harinas, se requiere de una solución que pueda brindarles las herramientas necesarias para acceder a la información sobre la producción de sus Molinos, así como poder consultar un histórico, visualizar tendencias y disponer de diferentes tipos de criterios en cuanto a las consultas realizadas sobre el históricos de los datos. La solución propuesta se desarrolló con herramientas Open Source, libres de licencia con el fin de optimizar costos.



RESUMEN

La integración y mejora permanente de las Tecnologías de Información y Comunicaciones (TIC) en las actividades de las empresas, permite siempre la optimización de los procesos que intervienen en el desarrollo de los productos a comercializar. Esta es una de las motivaciones que se tuvo para el desarrollo del presente trabajo de investigación, que logró implementar una solución de Inteligencia de Negocios en una empresa productora de Harinas haciendo uso de herramientas Open Source.

Se logró mejorar su proceso de producción y también se proporcionó el acceso a sus datos con la finalidad de poder analizar los mismos y permitir una oportuna toma de decisiones de ser necesario. La empresa cuenta con un sistema SCADA para el control de sus equipos, el software correspondiente permite la comunicación con bases de datos por lo que se extrae los datos directamente desde los PLC's y se insertan a la base de datos PostgreSQL. Posteriormente se implementó el respectivo proceso ETL que extrae la información de las tablas, y la carga a un datamart, diseñado para el área de producción de Molinos; se definió y se implementó el cubo OLAP para el análisis de datos respectivo; todas estas tareas se llevaron a cabo utilizando software de Pentaho y en el caso del análisis de datos un plugin compatible con el servidor. Todas estas herramientas son open source en sus respectivas ediciones de la comunidad.

La metodología empleada para el desarrollo del presente trabajo de investigación fue personalizada, partiendo del análisis de otras metodologías correspondientes al proyecto a desarrollar como lo son la metodología Kimball, la metodología Moss, y la metodología del desarrollo de sistemas de información; en base a las metodologías mencionadas se elaboró el plan de trabajo para éste proyecto.

Palabras clave:

Proceso de negocio

Inteligencia de Negocios

Open Source

ABSTRACT

The integration and permanent improvement of Information and Communication Technologies (ICT) in the industry activities, always allows the optimization of the processes that intervene in the development of the products to be commercialized. This is one of the motivations for the development of this research work, which managed to apply a Business Intelligence solution in a flour production company using Open Source tools.

Its production process was improved and access to its data was also provided in order to be able to analyze them and allow a timely decision making if necessary. The company has a SCADA system to control its equipment, this software allows communication with databases, so the data is extracted from the PLC's and inserted into the PostgreSQL database. Additionally, the respective ETL process was implemented, which extracted the information from the tables, and then load it to a datamart designed for the Mills production area; the OLAP cube was defined and implemented for the respective data analysis; all these tasks were carried out using Pentaho software, and in the case of data analysis, a plugin compatible with the server. All these tools are open source in their respective community editions.

The methodology used for the development of this research work was personalized, based on the analysis of other methodologies corresponding to the project to be developed, such as the Kimball methodology, the Moss methodology and the information systems development methodology; this project's work plan was elaborated based on the mentioned methodologies.

Keywords:

Business process

Business intelligence

Open Source

INDICE GENERAL

DEDICATORIA	iii
INTRODUCCION	iv
RESUMEN	vi
ABSTRACT	vii
INDICE GENERAL.....	viii
INDICE DE IMAGENES.....	xiii
NOMENCLATURA	xv
1 CAPÍTULO I PLANTEAMIENTO TEORICO	1
1.1 PLANTEAMIENTO DEL PROBLEMA.....	1
1.1.1 Descripción.....	1
1.1.2 Problema.....	2
1.2 OBJETIVOS DEL PROYECTO.....	2
1.2.1 Objetivo General.....	2
1.2.2 Objetivos Específicos.-	3
1.3 JUSTIFICACIÓN	3
1.3.1 Académica	3
1.3.2 Técnica	4
1.3.3 Económica y Social.....	4
1.4 IMPACTO	4
1.4.1 Local.....	4
1.4.2 Regional y Nacional.....	4
1.4.3 Internacional	5
1.5 METODOLOGIA	5
1.5.1 Diseño	6
1.5.2 Nivel.....	6

1.5.3	Tipo	6
1.5.4	Alcances y Limitaciones	7
2	CAPÍTULO II MARCO TEORICO	8
2.1	ANTECEDENTES	8
2.1.1	Históricos.-	8
2.1.2	Estado del Arte.-	9
2.2	MARCO TEORICO.....	10
2.2.1	Inteligencia de Negocios (o Business Intelligence, BI).-.....	10
2.2.2	Data warehouse (DW).-	11
2.2.3	Datamart.-.....	11
2.2.4	OLAP (On-line analytical processing).-	11
2.2.5	Juicio de Expertos.-.....	12
2.2.6	Business Process Management (BPM).-.....	12
2.3	MARCO CONCEPTUAL.....	13
3	CAPÍTULO III GESTIÓN DEL PROCESO DE NEGOCIO	15
3.1	GENERALIDADES	15
3.2	FASE DE DESCUBRIMIENTO.....	15
3.2.1	Descripción de la situación.....	15
3.2.2	Sipoc.....	17
3.2.3	Reci	18
3.3	FASE DE MODELADO	18
3.3.1	Modelo del proceso actual.....	18
3.3.2	Flujo de la recolección de datos durante el proceso.	20
3.3.3	Problemas del proceso	21
3.4	FASE DE EJECUCIÓN	21
4	CAPÍTULO IV INGENIERIA DEL PROYECTO.....	23
4.1	GENERALIDADES	23

4.2	ETAPA DE ANALISIS	23
4.2.1	Identificación de problemas y oportunidades	23
4.2.2	Determinación de los requerimientos de información	26
4.2.3	Análisis de las necesidades del sistema	30
4.2.4	Análisis de recursos disponibles	31
4.2.5	Herramientas de trabajo	31
4.2.6	Diagrama de Casos de Uso.....	37
4.2.7	Diagrama de Secuencia	39
4.3	ETAPA DE DISEÑO.....	40
4.3.1	Base de Datos	40
4.3.2	Solución de monitoreo de Indicadores.....	49
4.3.3	Análisis de Datos	51
4.3.4	Reportes.....	54
5	CAPÍTULO V DESARROLLO E IMPLEMENTACION	56
5.1	GENERALIDADES.	56
5.2	METODOLOGÍA A EMPLEAR.	56
5.2.1	Metodología de Kimball	56
5.2.2	Metodología de Inmon	57
5.2.3	Metodología de Moss.....	57
5.3	IMPLEMENTACIÓN DE BASE DE DATOS	57
5.3.1	Tabla de turnos	57
5.3.2	Tabla de totalizadores	58
5.3.3	Tabla de producción.....	58
5.3.4	Tabla Bits de Estado	59
5.3.5	Tabla de Motor	59
5.3.6	Tabla de Balanza.....	60
5.3.7	Tabla de Motor Detenido	60

5.3.8	Tabla de parada de Emergencia	61
5.3.9	Tabla de Log.....	61
5.4	CONFIGURACIÓN EN SCADA	62
5.5	MOTOR DE APLICACIONES TOMCAT	68
5.6	SERVIDOR JASPER REPORT SERVER	71
5.7	IMPLEMENTACIÓN DE REPORTES	74
5.8	PROCESO DE INTEGRACIÓN DE DATOS	75
5.9	IMPLEMENTACIÓN DE ESQUEMA	75
5.10	SERVIDOR PENTAHO BUSINESS ANALITYCS	76
5.10.1	Preparación del entorno de instalación	76
5.10.2	Instalación	77
6	CAPÍTULO VI RESULTADOS	85
6.1	GENERALIDADES	85
6.2	PROCESAMIENTO DE DATA	87
6.2.1	Administrativos.....	89
6.2.2	Producción.....	89
6.2.3	Técnicos	90
6.2.4	Económicos	90
6.3	PROYECCIONES	91
7	CONCLUSIONES	92
8	RECOMENDACIONES	94
	BIBLIOGRAFÍA	95
	ANEXOS	98
	Anexo 1 Entrevista semi estructurada PREVIA A IMPLEMENTACIÓN	98
	Anexo 2 Entrevista semi estructurada AL FINAL DE LA IMPLEMENTACIÓN	101
	Anexo 3 Plan de TESIS	102

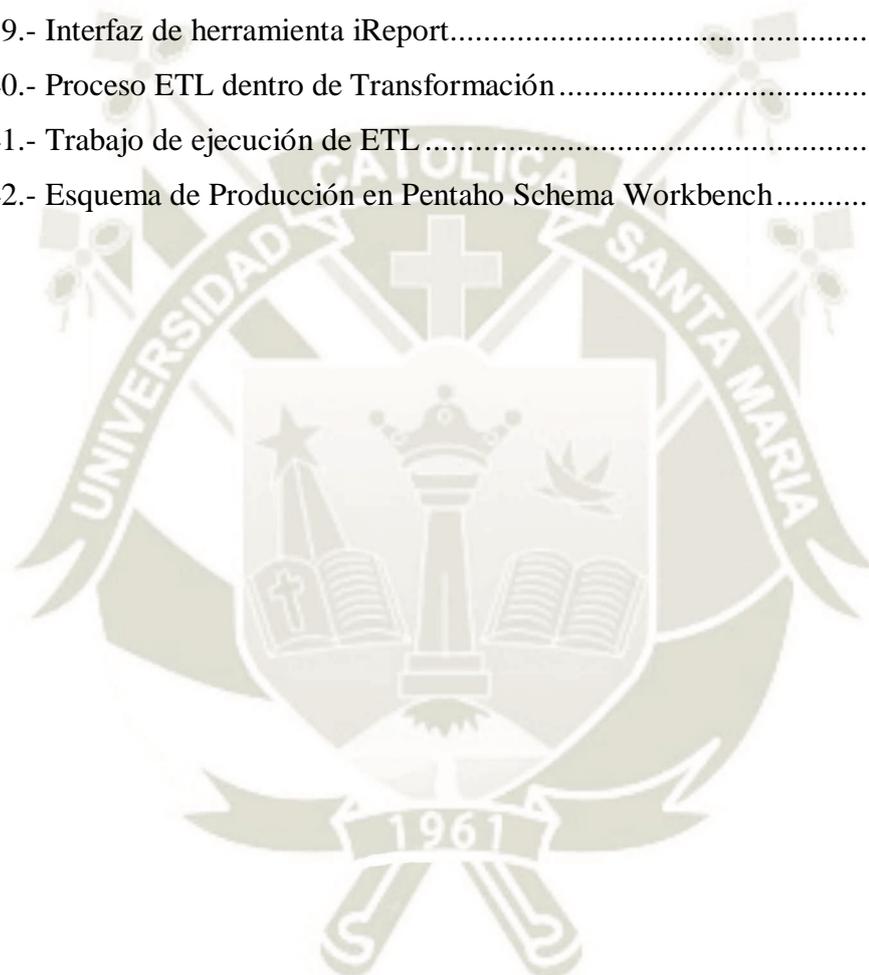
INDICE DE TABLAS

Tabla 1.- Tabla de Tags requeridos.....	30
Tabla 2.- Tabla de Turnos de Producción	30
Tabla 3.- Comparativo de Bases de Datos	34
Tabla 4.- Definición de Tabla de Totalizadores	41
Tabla 5.- Definición de Tabla de Turnos	41
Tabla 6.- Definición de Tabla de Producción	43
Tabla 7.- Definición de Tabla de Bits de Estado	43
Tabla 8.- Definición de Tabla de Motor Activo	44
Tabla 9.- Definición de Tabla de Molino	45
Tabla 10.- Definición de Tabla de paradas Emergencia	46
Tabla 11.- Dimensión de horas	52
Tabla 12.- Dimensión de fechas	52
Tabla 13.- Tabla hechos de producción	53
Tabla 14.- Comportamiento de Indicadores	87
Tabla 15.- Comparativo de Sistema en dos tiempos	88
Tabla 16.- Referencial del cálculo	90

INDICE DE IMAGENES

Imagen 1.- Diagrama SIPOC.....	17
Imagen 2.- Diagrama RECI.....	18
Imagen 3.- Proceso de producción.....	19
Imagen 4.- Toma de datos en proceso de producción.....	20
Imagen 5.- Proceso con implementación de Sistema.....	22
Imagen 6.- Interfaz de Control de Balanzas.....	24
Imagen 7.- Jerarquía del área de Producción.....	26
Imagen 8.- Relación de tags.....	28
Imagen 9.- Intouch Window Maker.....	32
Imagen 10.- Logo Play Framework.....	33
Imagen 11.- Logo TIBCO Jaspersoft.....	36
Imagen 12.- Logo Hitachi Pentaho.....	36
Imagen 13.- Logo de Apache Tomcat.....	37
Imagen 14.- Caso de Uso de Software de Análisis de Datos.....	38
Imagen 15.- Caso de Uso de Software de Monitoreo.....	38
Imagen 16.- Caso de Uso Tratamiento de Datos.....	39
Imagen 17.- Diagrama de Secuencia.....	40
Imagen 18.- Flujo de función sobre tabla totalizadores previo al registro.....	47
Imagen 19.- Flujo de función sobre tabla totalizadores tras registro.....	48
Imagen 20.- Flujo de función sobre tabla bits de estado.....	49
Imagen 21.- Diseño de Interfaz de Monitoreo.....	51
Imagen 22.- Configuración en 32 bit de ODBC.....	62
Imagen 23.- Bind List para Tabla Totalizadores.....	63
Imagen 24.- Bind List para Tabla de bits de Estado.....	63
Imagen 25.- Condición de Scada: On Startup.....	65
Imagen 26.- Condición de Scada: While Running.....	65
Imagen 27.- Scripts de condición.....	66
Imagen 28.- Script de cambio de dato.....	67
Imagen 29.- Script de condición: Balanza Flujo 0.....	67
Imagen 30.- Script de condición: Balanza Flujo mayor a 0.....	68
Imagen 31.- Creación de variable Java Home.....	69

Imagen 32.- Variable de Sistema PATH	69
Imagen 33.- Entrada para Java Home dentro de Path	70
Imagen 34.- Pantalla de Inicio de Tomcat.....	71
Imagen 35.- Wizard de instalación para Jaspersoft Report Server	72
Imagen 36.- Pantalla de Inicio de Jasper Report Server	72
Imagen 37.- Configuración Fuente de Datos en Jasper Server.....	73
Imagen 38.- Fuente de datos Creada.....	73
Imagen 39.- Interfaz de herramienta iReport.....	74
Imagen 40.- Proceso ETL dentro de Transformación.....	75
Imagen 41.- Trabajo de ejecución de ETL	75
Imagen 42.- Esquema de Producción en Pentaho Schema Workbench.....	76



NOMENCLATURA

BI	Business Intelligence
CTE	Common Table Expression
CRUD	Create, Read, Update, Delete
DM	Data Mart
DW	Data Warehouse
ETL	Extract, Transform & Load
ERP	Enterprise Resource Planning
FK	Foreign Key
IN	Inteligencia de Negocios
JDK	Java Development Kit
JRE	Java Runtime Environment
OLAP	On-Line Analytical Processing
OLTP	On-Line Transactional Processing
PK	Primary Key
PBA	Pentaho Business Analytics
PDI	Pentaho Data Integration
PSW	Pentaho Schema Workbench
SCADA	Supervisory Control and Data Acquisition
SQL	Structured Query Language

CAPÍTULO I

PLANTEAMIENTO TEORICO

1.1 PLANTEAMIENTO DEL PROBLEMA

1.1.1 Descripción

En la actualidad, la tecnología ha alcanzado a las empresas de todo tamaño a lo largo de nuestro país. En primera instancia el alza de uso de sistemas de gestión empresarial como son los Enterprise Resource Planning (ERP), los cuales cuentan con un amplio margen de manejo sobre diversas áreas de una institución, logrando integrar sus funciones. Son miles de empresas que desde los años 90 han implementado estos sistemas integrados, debido a que apoyan en la totalidad de los procesos empresariales (Lorenzo, 2003). Las innovaciones, han tenido como influencia más importante las TIC, las mismas que en los ERP han implementado cambios en las empresas, favoreciendo las actividades de éstas (Ramirez, 2004).

Por otro lado, y más enfocado en las áreas operativas, se encuentran los sistemas Supervisory Control and Data Acquisition / Supervisión, Control y Adquisición de Datos (SCADA) a través de los cuales se toma mayor control de la infraestructura operativa y se automatiza varios procesos. Estas aplicaciones controlan y supervisan procesos a distancias. Adquieren datos de procesos remotos para comunicarlos a diferentes dispositivos de campo (controladores autónomos, autómatas programables, etc.) y controlan el proceso de forma automática desde una PC y en tiempo real (Montalvo Jaramillo, Hernandez Cevallos, Ledesma Marcalla, & Santillán Gallegos, 2010). Es a través de estos sistemas que también puede recolectarse datos que llegan a ser críticos para la eficiente operación de la empresa, por lo que deben ser explotados en la mejor medida posible. Sin embargo, los datos presentados en estos sistemas no siempre son claros, o no son presentados de una manera sencilla y fácil de entender por lo que se

requiere la elaboración de reportes dedicados o el uso de otras herramientas para poder explotar los datos y transformarlos en información valiosa para una correcta toma de decisiones.

Ambos sistemas, aportan soluciones que pueden desarrollarse e implementarse en las necesidades y características de diferentes empresa, para alcanzar la eficiencia en los procesos que intervienen en la producción de harinas.

La actual empresa productora de harinas carece de un sistema automatizado para la recolección de datos de producción dentro del área de Molinos, y consecuentemente, tiene la difícil tarea de poder realizar un efectivo análisis de sus datos para la toma de decisiones, tomando un tiempo prolongado el poder elaborar un reporte básico sobre la producción, complicando aún más las cosas cuando se requiere una comparación con datos históricos. Del mismo modo, y a pesar de poseer datos básicos en su sistema SCADA, no logra un control efectivo sobre su producción y mantenerla a un índice de efectividad ideal de la planta.

1.1.2 Problema

En vista de las carencias tecnológicas y la posibilidad de generar aporte a través de las mismas, es posible hacernos la pregunta, ¿Cuál sería la mejora del proceso de Producción de una empresa productora de harinas que se obtendría aplicando Inteligencia de Negocios con herramientas Open Source?

1.2 OBJETIVOS DEL PROYECTO

Dentro de los objetivos del proyecto se tienen los siguientes:

1.2.1 Objetivo General

Mejorar el proceso de producción de una empresa productora de harinas aplicando inteligencia de negocios con herramientas Open Source.

1.2.2 Objetivos Específicos.-

- Automatizar y optimizar el trabajo de recolección de datos acerca de estados y cantidades de producción.
- Implementar un repositorio de datos óptimo para poder almacenar los datos.
- Implementar una solución basada en herramientas de inteligencia de negocios ágiles y escalables que permitan consultas de datos históricos del área de producción.
- Diseñar reportes automáticos para los niveles tácticos o estratégicos de acuerdo a sus requerimientos.
- Proveer de una solución que permita la visualización y monitoreo constante de indicadores establecidos.

1.3 JUSTIFICACIÓN

Como toda compañía, para asegurar el buen desempeño o rendimiento de sus instalaciones así como mantener la efectividad de su producción, es importante una toma de decisiones oportuna, basada en información real y disponible para el nivel táctico y estratégico de la empresa, que les permita generar conocimientos y obtener ventajas. Debido a lo señalado, es fundamental que la empresa disponga de los medios necesarios que le permitan realizar un rápido diagnóstico de su situación actual y tomar medidas adecuadas; estos medios pueden traducirse en herramientas tecnológicas que permitan visualizar la información crítica para la empresa de manera intuitiva y sencilla, facilitando a los jefes y gerentes acceso oportuno a la información sin tener que solicitar la misma a personal especializado.

1.3.1 Académica

El desarrollo del trabajo de investigación aportará el uso de Open Source en SCADA, siendo este un aporte para ser tomado en cuenta, por quienes consideren, una base o modelo a desarrollar en industrias o empresas que mantengan procesos de producción.

1.3.2 Técnica

Las consideraciones de implementación, asumidas en el desarrollo del presente trabajo de investigación, permiten plasmar una nueva forma metodológica, modificando algunas de las ya existentes, logrando los objetivos de la investigación.

1.3.3 Económica y Social

Las empresas ejecutan aportes al Estado, por medio de los pagos de impuestos y tributos, los mismos que son empleados en brindar el bienestar a todas las personas del territorio; al hacer más eficiente los procesos de las empresas, estos otorgan, de alguna manera, mayores ingresos a las mismas, por lo que se refleja en el pago de impuestos. Además de esta justificación, se considera que se mantiene una relación con lo social, ya sean por las políticas de gobierno como por la posibilidad de contratación de nuevo personal en la empresa, motivados por las mejoras que se implementan.

1.4 IMPACTO

1.4.1 Local

La empresa productora de harinas logrará tener información oportuna y confiable para tomar decisiones que impacte en su proceso de producción, logrando una mejora en su rendimiento y productividad. Adicionalmente se adopta nuevas tecnologías y se va eliminando barreras para la adopción de las nuevas técnicas de tratamiento de datos.

1.4.2 Regional y Nacional

Lograr mejoras en una industria, puede motivar en otras del mismo sector o similares, implementaciones similares, por lo que, al hacer pública los resultados del presente trabajo de investigación, proporciona al mercado regional y nacional alternativas de mejoras en sus procesos.

1.4.3 Internacional

El impacto internacional, es que al hacer uso de Open Source, se hace aporte a estos, debido a que la comunidad considera siempre que el uso y propuestas de mejoras son muy importantes, sea en la aplicación que sea.

1.5 METODOLOGIA

Si bien es cierto en el presente punto se debe establecer lo que corresponde a metodología del proceso de investigación, al ser un trabajo dentro de lo que se considera implementación de tecnología, es necesario indicar cuál ha sido la que se ha seleccionado y modificado para alcanzar los objetivos plasmados.

La mayoría de metodologías que se encuentran en la literatura son adecuadas para el desarrollo de proyectos de TI, orientados a sistemas de información o desarrollos de software; sin embargo para la finalidad de este proyecto se seguirá una metodología que se basa en un ciclo de los sistemas de inteligencia de negocios. Este ciclo consiste de cinco etapas, análisis, diseño, desarrollo, implementación y evolución siendo este último un paso muy importante ya que abarca el crecimiento del sistema de BI, el cual una vez adaptado e implementado va en crecimiento generando nuevos requerimientos y necesidades de información. Así se vuelve a iniciar el ciclo con el paso del análisis de los requerimientos y necesidades del sistema.

Cabe mencionar que la implementación de inteligencia de negocios, no se trata de una técnica metodológica exacta, o una plantilla ya preparada que permita guiar su desarrollo paso a paso, ya que en el mundo del BI, cada proyecto debe ser adaptado a cada organización tomando en cuenta sus requerimientos de información, su historia, sus procesos, el ambiente en el que se desenvuelve, sus políticas y normas internas, ya que todo esto influye en el proceso del BI y en la solución propuesta.

Es así, que en base a dicha metodología se realizó el análisis de los requerimientos del sistema mediante unas entrevistas con el personal responsable del proceso de negocio, definiendo inicialmente la recolección de datos, el procesamiento sobre los datos y los resultados que se buscaban obtener. Se definió los usuarios que estarían directamente involucrados y serían

responsables de la validación de los datos y por último se definió el alcance del proyecto. Posteriormente se realizó el diseño de todos los requerimientos que se lograron recoger desde la parte del cliente, validando cada parte con lo que se había logrado conversar.

Para la correcta implementación, se consideró el uso de un sistema gestor de base de datos con una interfaz amigable permitiendo realizar todas las tareas necesarias para un correcto desempeño de la base de datos, así como para asegurar la recolección de los datos.

Considerando que se debe realizar un desarrollo personalizado, se implementó el uso de un servidor web de aplicaciones, donde se colgará una aplicación de uso continuo.

Finalmente se consideró las respectivas herramientas que permitieron el análisis de datos, considerando de esta manera una o más herramientas que lograron los siguientes procesos: proceso ETL, elaboración del esquema o cubo OLAP, proceso de análisis de datos. Así mismo se considera el uso de un servidor de reportes.

1.5.1 Diseño

El presente trabajo de investigación mantuvo un diseño, cuantitativo, por el fundamento epistemológico del positivismo, además, se considera dentro de lo tecnológico, al aplicar tecnologías vigentes de información para poder dar solución al problema presentado.

1.5.2 Nivel

El nivel de este proyecto de tesis es descriptivo y relacional.

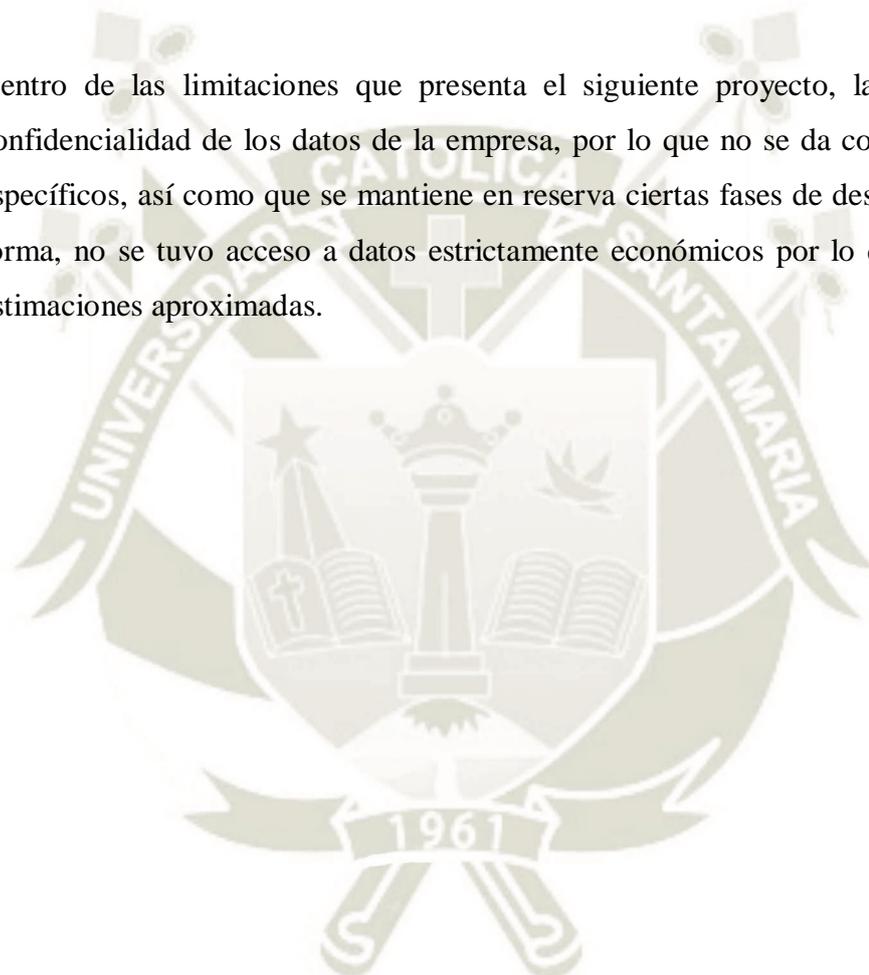
1.5.3 Tipo

Por reglamento de calificación y registro de Investigadores en Ciencia y Tecnología del SINACYT, es de tipo Aplicada, así también es experimental, ya que se comprobará la hipótesis planteada a través de la implementación del proyecto de inteligencia de negocio y posterior evaluación de las variables.

1.5.4 Alcances y Limitaciones

El proyecto se desarrollará en una empresa dentro de la ciudad de Arequipa, tomando en cuenta sólo su proceso de producción de harina respectivo a su única línea de producción. Dentro de los alcances, se considera la entrega de una solución de automatización de recolección de datos, un sistema de monitoreo de indicadores, reportes automatizados y una solución para el análisis de los datos.

Dentro de las limitaciones que presenta el siguiente proyecto, la principal es la confidencialidad de los datos de la empresa, por lo que no se da conocer datos muy específicos, así como que se mantiene en reserva ciertas fases de desarrollo. De igual forma, no se tuvo acceso a datos estrictamente económicos por lo que se realizarán estimaciones aproximadas.



CAPÍTULO II

MARCO TEORICO

2.1 ANTECEDENTES

2.1.1 Históricos.-

La industria en la ciudad de Arequipa presentó un gran impulso pasados los terremotos de 1958 y 1960, razón por la cual se presentaron incentivos tributarios para alcanzar un parque industrial, planificado en cinco etapas. La Junta de Rehabilitación de Arequipa crea un parque con una distribución de 80 lotes medianos y grandes distribuidos en 19 manzanas en un área total de 664,429.32 m². Se procedió a su funcionamiento en el año 1966. En 1978, los empresarios fundan la Asociación de Empresas del Parque Industrial de Arequipa (ADEPIA), para establecer estrategias sobre cómo enfrentar la desaparición de los beneficios tributarios, sin embargo en la apertura de mercados de los años 90, por políticas del estado, es que la mayoría de los parques industriales del Perú, se les comenzó a titular “cementerio de empresas” con la competencia de productos importados y la desleal que hacían presencia mayor en el mercado, sin embargo, luego de ver esta apertura como una amenaza, la transformaron en una oportunidad, por el cual se lograría la incursión en el comercio exterior (Asociación del Parque Industrial de Arequipa, 2018).

El día de hoy, en el parque industrial se puede ver empresas productoras de bienes, denominados de pan llevar, bebidas, papel, entre otros, es allí donde se ubica la empresa donde se implementó la mejora.

2.1.2 Estado del Arte.-

El desarrollo de una plataforma de Business Intelligence (BI) que permita a una institución de educación en Ecuador, un análisis de datos para poder tener la información de las competencias generales de formación establecidas en sus egresados (Universidad Técnica del Norte – UTN). Empleó la metodología Kimball para el desarrollo del datawarehouse. Primero se estableció los conceptos y competencias generales de formación de BI, luego establecieron la arquitectura de la plataforma, acondicionaron el entorno tecnológico donde se instalaría y configuraría el BI, desarrollaron el Data Warehouse, por medio de filtros y carga de datos, para al final implementar las herramientas de reporte y cuadros de mando, con ello se apreció el fortalecimiento en el proceso de seguimiento a egresados (Guevara Vega, 2015).

Comentario.- Este trabajo de tesis resulta muy interesante ya que contiene la opinión de algunos expertos en el área de BI a momento de realizar la elección de la herramienta y tecnologías a usar, indicando su criterio con respecto a las opciones y posibles herramientas a elegir. De tal forma, provee un enfoque a ser tomado en cuenta al momento de realizar nuestra elección.

El éxito de las empresas está en función al uso de sus recursos intangibles, dados en los datos e información que se controle, para alcanzar un proceso de toma de decisiones más asertivo, con el incremento de datos almacenados en los procesos empresariales es que se mantienen las necesidades de aplicar BI, que logra capturar, almacenar, procesar, analizar y mostrar de manera eficiente, los datos generados. Para ello se empleó la metodología HEFESTO en el desarrollo de almacenes de datos y la suite comunitaria de Pentaho, en su versión 4.8.0 (Vanegas Lago & Guerra Cantero, 2013).

Comentario.- Este trabajo es muy claro al reconocer la importancia del manejo de los datos para lograr la máxima eficiencia a la hora de tomar decisiones. Debido al alcance de este trabajo, la metodología mencionada es apropiada y sirvió de guía para afinar los criterios de elección de metodología.

En la comparación de las herramientas Open Source aplicadas en BI, para ver los beneficios que podrían tener en una empresa de Lácteos de Ecuador, muestra en la

personalización de reportes una eficiencia en el tratamiento de los datos, se considera junto a BI la aplicación de la Suite de Pentaho Community Edition (CE), con la cual se dan reportes que permiten mejores tomas de decisiones de la gerencia. Además, la data warehouse se dio con la metodología HEFESTO, que incluye el análisis de requerimientos, análisis de los sistemas transaccionales, modelo lógico y la integración de datos (Carrión Albán & Torres Sosa, 2016).

Comentario.- Este trabajo de tesis, principalmente me otorgó una clara visión de las herramientas Open Source más relevantes en el mercado a la fecha, a través de la comparación de las tres más populares. De amplio valor a la hora de tener que escoger entre, prácticamente, las mismas herramientas.

2.2 MARCO TEORICO

2.2.1 Inteligencia de Negocios (o Business Intelligence, BI).-

Para hablar de inteligencia de negocios tenemos que remontarnos al año 1958, en el cual Hans Peter Luhn hace la primera referencia al término definiendo la inteligencia de negocios como “la habilidad de encontrar las relaciones que se presentan en un conjunto de hechos, de tal manera que permita encaminar acciones hacia un objetivo deseado”

Se considera su concepto referido a la inteligencia como aquella información apreciada por su valor y relevancia. Es información experta, conocimiento y tecnologías eficientes en el manejo de negocios organizacionales e individuales. Por lo tanto, la inteligencia de negocios es una amplia categoría de aplicaciones y tecnologías para recolectar, proveer acceso a, y analizar data con el propósito de ayudar a que los usuarios empresariales puedan tomar mejores decisiones sobre negocios. (Ranjan, 2009)

El concepto de BI puede ser dividido en tres partes (Khan & Quadri, 2012):

- Captura o Adquisición de datos.- Que puede ser de fuentes de datos internas o externas.

- Almacenamiento de datos.- Normalmente se almacena en una Almacén de datos, datamart o Data Warehouse, tras haber pasado por un proceso de ETL (Extracción, Transformación y carga).
- Acceso y análisis de datos.- Una vez se tienen los datos listos en un repositorio de datos, estos pueden ser analizados para extraer conocimiento y lograr una toma de decisiones.

2.2.2 Data warehouse (DW).-

Es el componente significativo de la inteligencia de negocios. Es orientado a temas, integrado. El data warehouse soporta la propagación física de data manejando los numerosos registros empresariales para tareas de integración, limpieza, agregación y consulta. También puede contener la data operativa, la cual puede ser definida como un conjunto integrado de data actualizable usada para una amplia toma de decisiones tácticas de un área en particular. Contiene data viva, no fotografías, y retienen un mínimo de historia. Las fuentes de datos pueden ser bases de datos operativas, data histórica, data externa como por ejemplo de compañías de investigación de mercado o de Internet, o información del ambiente de un data warehouse que ya exista. (Ranjan, 2009)

2.2.3 Datamart.-

Según lo descrito por Inmon (1999), es una colección de campos de información organizados para el soporte a la toma de decisiones basado en la necesidad de un departamento en específico. Cada departamento tiene su propia interpretación de cómo debería verse su datamart, y cada datamart de cada departamento es peculiar y específico a sus propias necesidades.

2.2.4 OLAP (On-line analytical processing).-

De acuerdo a Se refiere a la manera en la que los usuarios del negocio pueden cortar y arreglar su camino a través de los datos utilizando herramientas sofisticadas que permiten la navegación de dimensiones tales como el tiempo o las jerarquías.

Procesamiento analítico en línea u OLAP proporciona vistas multidimensionales y resumidas de datos comerciales y se utiliza para generar informes, análisis, modelado y planificación para optimizar el negocio. Las técnicas y herramientas OLAP se pueden usar para trabajar con almacenes de datos o datamarts diseñados para sistemas sofisticados de inteligencia empresarial. Estos sistemas procesan consultas necesarias para descubrir tendencias y analizar factores críticos. El software de generación de informes genera vistas agregadas de datos para mantener a la administración informada sobre el estado de su negocio.

2.2.5 Juicio de Expertos.-

El Juicio de expertos es un medio o estrategia de evaluación de resultados o de contenidos. Esta estrategia consiste en solicitar a personas que dominen un tema o tópico en específico, compartir su opinión o criterio acerca de un tema en concreto, ya sea acerca de una materia en cuestión, el resultado de una investigación o en el caso de las TIC's una herramienta desarrollada. En este medio los expertos son seleccionados con mucho cuidado a través de procedimientos diversos, y los mismos deberán ser personajes relacionados al tema que se quiere evaluar, que tengan bases sólidas de conocimiento y experiencia. (Cabero Almenara & Llorente Cejudo, 2013)

2.2.6 Business Process Management (BPM).-

La gestión de procesos de negocio o business process management (BPM), es una metodología de tendencia que se encarga de administrar y mejorar los procesos. Este concepto se centra en la administración de los procesos del negocio desde que comienzan hasta su final, analizando cada etapa dentro del mismo buscando mejorar la eficiencia y la eficacia de todo el proceso a través una gestión adecuada de cada etapa al interior del proceso, orientando los esfuerzos para la mejora y optimización de los mismos. Cabe señalar que de acuerdo a Díaz, la filosofía BPM *“se ve como un sistema completo de información y comunicación, a través de un marco documental que permite publicar, almacenar, crear, modificar y gestionar procesos, así como acceder a ellos en cualquier momento y lugar”*. (Díaz Piraquive, 2008)

2.3 MARCO CONCEPTUAL

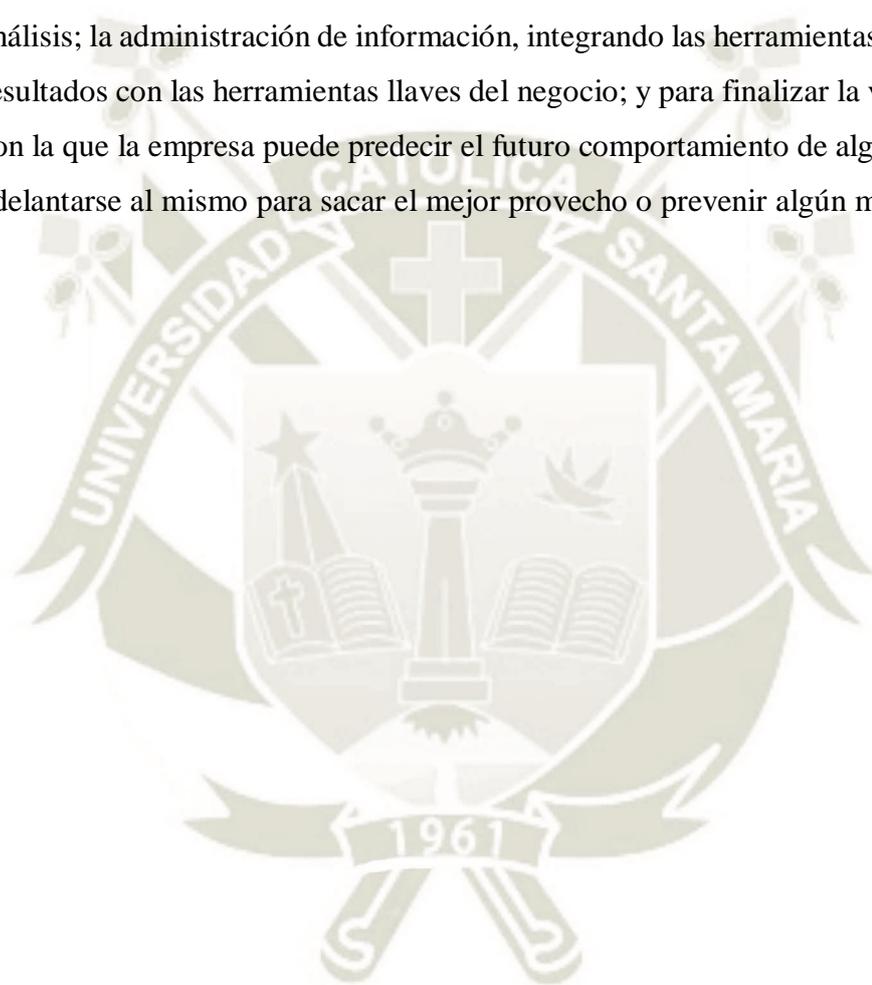
La inteligencia de negocios es una práctica que se vuelve cada vez más común en la actualidad, sobre todo en empresas que aspiran a un efectivo y eficaz desarrollo; para lo cual éstas buscan la mejor manera de poder recolectar todos sus datos, poder sintetizarlos y analizarlos con herramientas que les permitan tomar decisiones de importancia o criticidad para la compañía. De acuerdo a mi experiencia toda empresa u organización que maneja un negocio genera una enorme cantidad de datos a lo largo de su vida desde su creación. Y para poder ser efectivo en lo que se hace y lograr un crecimiento en el negocio es necesario saber utilizar y sacar provecho de esos datos, pero esto no es tan sencillo como puede sonar; no todos los datos generados son de utilidad o representan algo de interés para los usuarios por lo que dichos datos deben pasar por un proceso de selección y filtro, seleccionando aquellos que brinden un verdadero aporte a las operaciones o que sean de interés a los usuarios y posteriormente cargar estos datos a un repositorio donde puedan ser accesibles y procesables. Con estos datos una vez agrupados podemos llegar al nivel de poseer información que puede servir para optimizar procesos o procedimientos del negocio. Pero no podemos detenernos allí, ya que para poder dar una ventaja al negocio es necesario disponer del conocimiento preciso para una toma efectiva de decisiones que puedan dar el giro correcto al negocio y encaminarlo al éxito. Para poder obtener dicho conocimiento es necesario aplicar Inteligencia de Negocios.

Hoy en día, el comportamiento de una organización se ve reflejado en sus datos históricos, es allí donde la inteligencia de negocios entra al juego, permitiendo a los niveles tácticos y estratégicos identificar patrones o tendencias dentro de su organización, o de entidades externas para así lograr una toma de decisiones efectiva, respaldada no en simples datos, sino en información que ha sido debidamente procesada y tratada para convertirse en conocimiento.

“La inteligencia de negocios consiste en extraer y analizar la información de una organización con el fin de obtener su conocimiento para aplicarlo en la organización, de manera que pueda orientar adecuadamente la toma de decisiones y permita un

mejoramiento incremental de los procesos para la obtención de mejores resultados” (Flórez Fernández, 2012).

Es una práctica común de hoy el hacer uso de herramientas de inteligencia de negocios que permitan a las empresas evolucionar y pasar de ser reactivas a ser proactivas, esta evolución está definida en 03 etapas (IBM, 2012) que son la empresa centrada en TI, donde se enfoca en la mejora de recolección de datos y selección de herramientas de análisis; la administración de información, integrando las herramientas de análisis y sus resultados con las herramientas llaves del negocio; y para finalizar la visión predictiva, con la que la empresa puede predecir el futuro comportamiento de algo en específico y adelantarse al mismo para sacar el mejor provecho o prevenir algún mal porvenir.



CAPÍTULO III

GESTIÓN DEL PROCESO DE NEGOCIO

3.1 GENERALIDADES

En el siguiente capítulo se describe el procedimiento realizado en el proceso de producción de harinas correspondiente al BPM, donde se pudo definir que existían problemas que podían ser solucionados a través de la aplicación de tecnologías específicas.

3.2 FASE DE DESCUBRIMIENTO

Para lograr una identificación eficiente del proceso objetivo, se aplicaron unas técnicas que permiten el correcto descubrimiento del proceso que se quiere considerar, definiendo así su alcance dentro del proyecto además de aclarar ciertos puntos específicos. Estas técnicas fueron las siguientes:

3.2.1 Descripción de la situación

Actualmente el proceso de producción de harina en la empresa inicia con un requerimiento de producto al área de molinos, que puede provenir del área comercial o de otras líneas de producción. Al recibir esta solicitud, el jefe de producción es el encargado de generar un plan de producción para el molino; éste plan es enviado al supervisor de planta para que éste pueda realizar las gestiones y coordinaciones necesarias para cumplir con dicho plan. En la primera etapa de producción se realiza el abastecimiento de molino, para esto es preciso contar con la orden de producción que indica qué producto es el que se va a elaborar y, en base a dicho documento, se procede a abastecer el molino con el grano o la combinación de granos de trigo respectivos al producto a elaborar, éstos son llevados desde el o los silos de almacenamiento hacia el molino a través de fajas o paletas para ingresar a la posterior etapa de limpieza; cabe tener en cuenta que existe uno o más silos por cada tipo de grano. Luego se da la puesta en marcha del molino desde la aplicación SCADA, encendiendo los motores respectivos y las máquinas a utilizar haciendo uso del control automatizado que ofrece la ventana de control de SCADA, dando inicio al procedimiento de conversión del

grano de trigo en harina de trigo. En la etapa de limpieza, a través de diversos mecanismos se eliminan impurezas de la materia prima como pueden ser ramas, piedras, pasto, entre otros elementos ajenos a los granos de trigo, así mismo dentro de la etapa de limpieza, se ejecuta un procedimiento de selección de granos óptimos para el proceso de producción; se realiza una clasificación de los granos descartando aquellos que puedan no estar en las mejores condiciones.

Finalizada ésta etapa, se procede a la etapa de acondicionamiento en la cual, dependiendo del grano, se realiza un ajuste de porcentaje de humedad aplicando agua para llegar a los niveles óptimos del estado físico del grano el cual se almacena, tras haber reposado, para el proceso de molienda.

Al ya disponer de un grano óptimo este es llevado hacia los mecanismos de molienda. Dentro de éste se ejecutan dos o más etapas de rotura dependiendo del producto a obtener, donde los granos pasan a través de unos cilindros estriados para quebrar el grano separando de ésta manera el grano en endospermo y salvado de trigo; a medida que va avanzando, los cilindros presentan mayor cantidad de estrías para poder obtener una molienda más fina y llegar así a la harina de trigo. Generalmente tras la segunda rotura ya es posible realizar el cernido para separar la harina del salvado, manteniendo sólo la harina, que es llevada hacia los siguientes procedimientos con un fluido constante a través de presión de aire. Posteriormente se purifica la harina extrayendo los últimos restos de salvado que puedan quedar presentes a través de una combinación de cernido y aspiración; a todo lo mencionado desde la limpieza de los granos hasta su purificación final se le engloba como procedimiento de molienda para fines de este trabajo. Finaliza con el balanceado del producto y su respectivo embolsado para ser llevado a almacén y posteriormente despachado.

A lo largo de todo el proceso mencionado se realiza la supervisión del mismo y de las máquinas, para asegurar que todo marche de manera correcta; siendo éstas dos últimas tareas normalmente trabajos del operador de planta. De igual forma, el supervisor realiza un control de la producción y una supervisión llevada más hacia el cumplimiento del plan de producción, registrando la producción al final de cada turno, y en ciertas ocasiones registra los datos de producción durante al momento en que se le solicita dicha tarea, sobre todo para llevar un control. Todo evento se registra en un libro llamado parte de producción diaria. En este se lleva 2 caras por cada día de producción, donde se registra la cantidad producida al final de cada turno (el valor del totalizador

no borrrable), la cantidad de sacos de producto y se registran todos los eventos que puedan suceder desde paradas programadas hasta paradas súbitas y cualquier otro evento del personal de planta.

3.2.2 Sipoc

Nos permite identificar cada pilar clave sobre el proceso. En cuanto a este proyecto, nuestro objetivo es el proceso de producción de harina en los molinos, así que al tener claro el diagrama SIPOC se puede identificar de manera precisa los proveedores, los ingresos, salidas y clientes además de las actividades más relevantes del proceso.

Imagen 1.- Diagrama SIPOC

Procedimiento de conversión del grano de trigo en harina de trigo

S Proveedor	I Entrada	P Proceso	O Salida	C Cliente
Jefe de producción	Trigo sucio	Ordenar producción	Harina	Silo de Harina
Silos de Trigo	Orden de abastecimiento	Abastecimiento	Subproductos	Almacén
Supervisor de producción	Orden de producción	Encendido	Bolsas de harina	Distribuidor
		Procedimiento de Molienda		Cientes internos
		Supervisión		
		Embolsado		
		Despacho		

Fuente: Elaboración propia

3.2.3 Reci

Una vez aplicado SIPOC procedemos a elaborar el diagrama RECI con el cual podemos identificar los responsables de cada parte del proceso así como los roles involucrados en el mismo. Para nuestro proceso de producción de harina tenemos en cuenta todos los participantes del proceso desde el abastecimiento hasta el despacho.

Imagen 2.- Diagrama RECI

Actividades / Roles	Gerente de Producción	Jefe de Producción	Analista de Producción	Supervisor de producción	Operadores	Logística	Comercial	Mantenimiento
Planificar producción	I	R/E	I	I		C	C	
Abastecimiento		R/C	C	E	E	I		
Encendido	I	R		E	E			C/I
Procedimiento de Molienda		R/I	C/I	C	E			
Supervisión y verificación		R/C	C/I	E	E			
Embolsado	I	R		C	E	I		
Despacho	I	I				R/E	I	

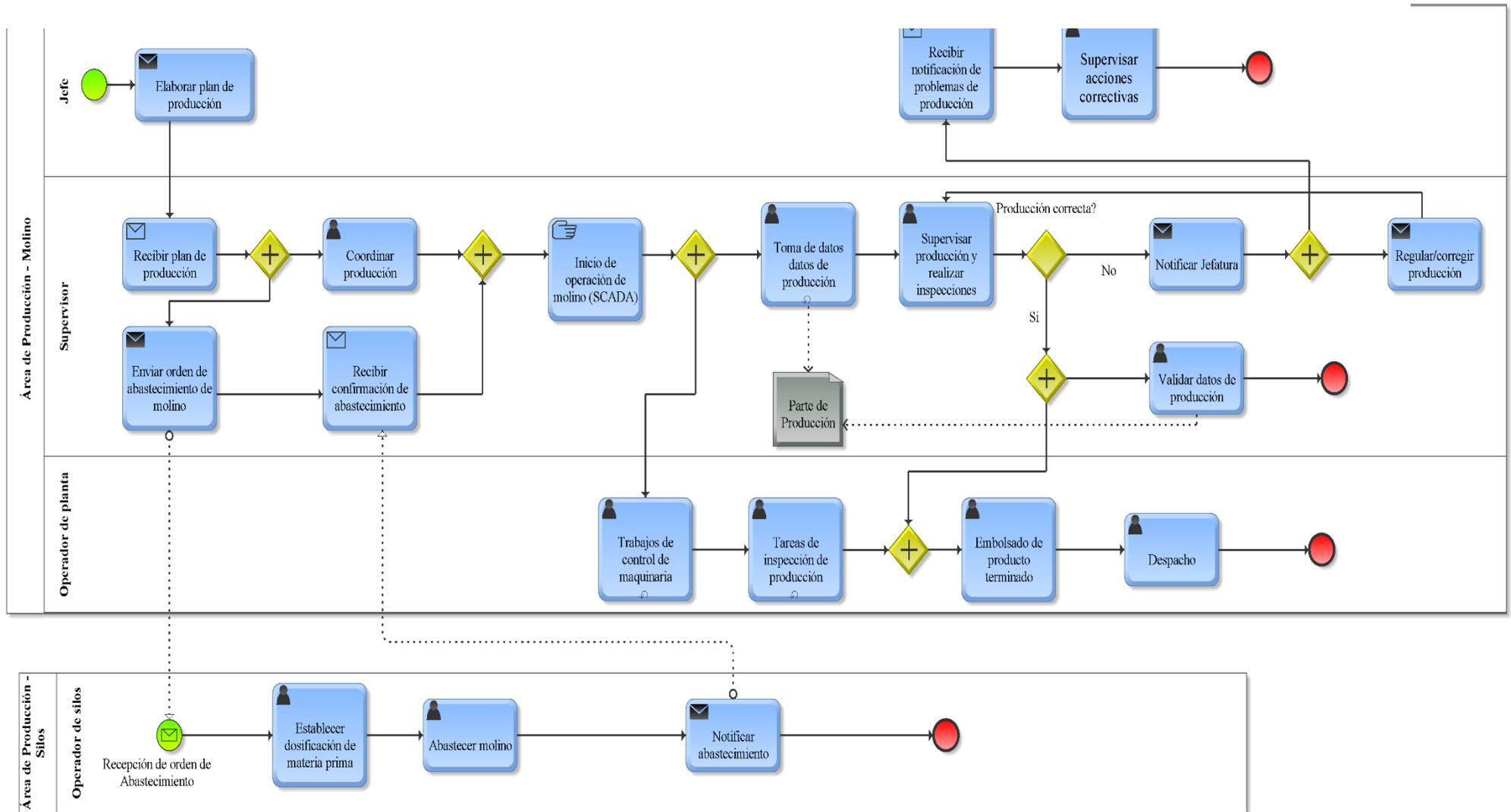
Fuente: Elaboración propia

3.3 FASE DE MODELADO

3.3.1 Modelo del proceso actual

Haciendo uso de la notación BPMN, modelamos el proceso actual que se daba en la empresa productora de harina, enfocándonos en su proceso productivo de harina. Consideramos todo el proceso sin detallar el procedimiento de molienda.

Imagen 3.- Proceso de producción

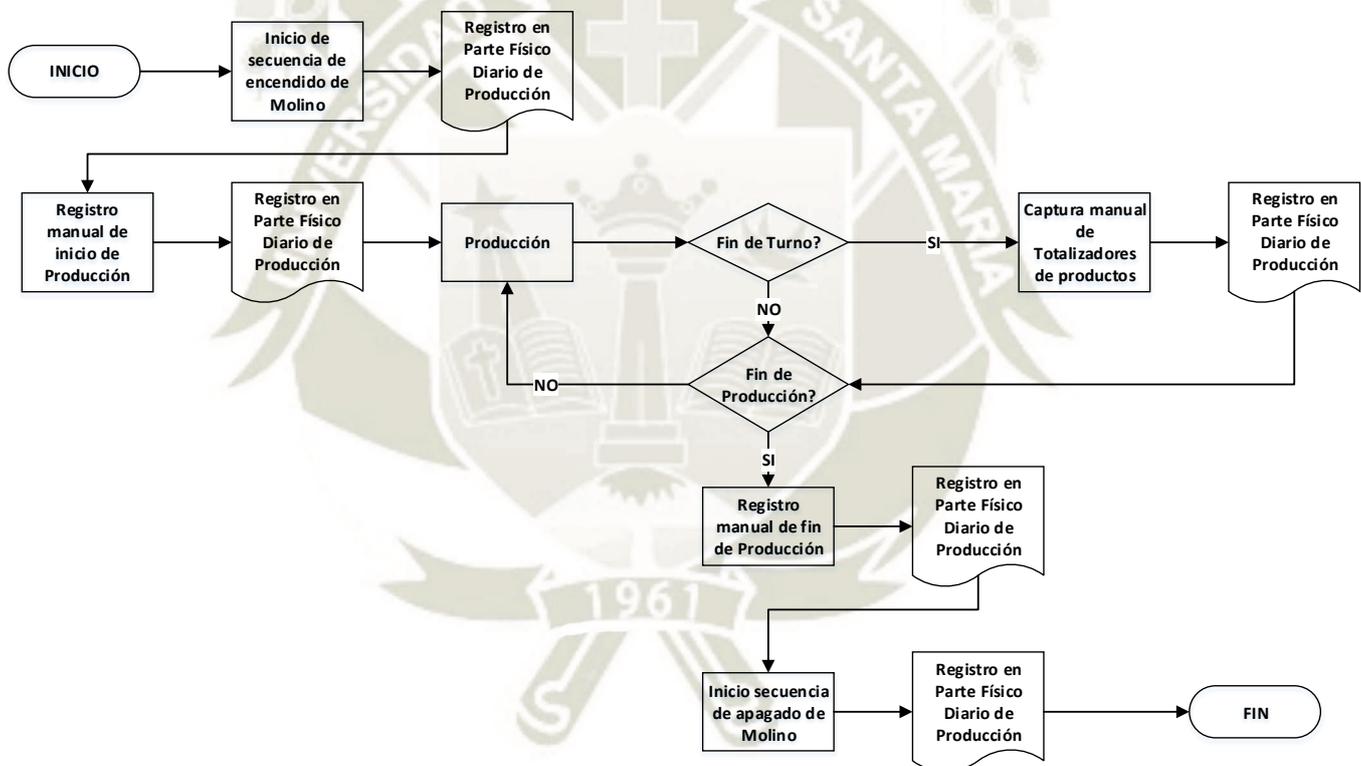


Fuente: Elaboración propia junto a personal de operaciones de la empresa

3.3.2 Flujo de la recolección de datos durante el proceso.

En el proceso actual, los datos de producción son recopilados de manera periódica de cada balanza dentro del proceso. La frecuencia de esta toma de datos es, usualmente, a final de cada turno; sin embargo, existen ocasiones frecuentes en las que el supervisor es solicitado para tomar estos datos de producción durante el turno en curso, para llevar a cabo un seguimiento de la producción. Esta toma de datos se puede graficar en el flujo siguiente:

Imagen 4.- Toma de datos en proceso de producción



Fuente: Elaboración propia

Se observa en el diagrama anterior, que parte del proceso es el registro de la hora de inicio de funcionamiento del molino, así como la hora de fin de funcionamiento.

3.3.3 Problemas del proceso

En el proceso actual, tras haber analizado sus componentes, comportamiento y flujo de información, es posible determinar que la carencia de un sistema que brinde la información oportuna y adecuada es el principal problema del proceso. Esto centrado sobre todo en la recolección de datos para el debido proceso de los mismos.

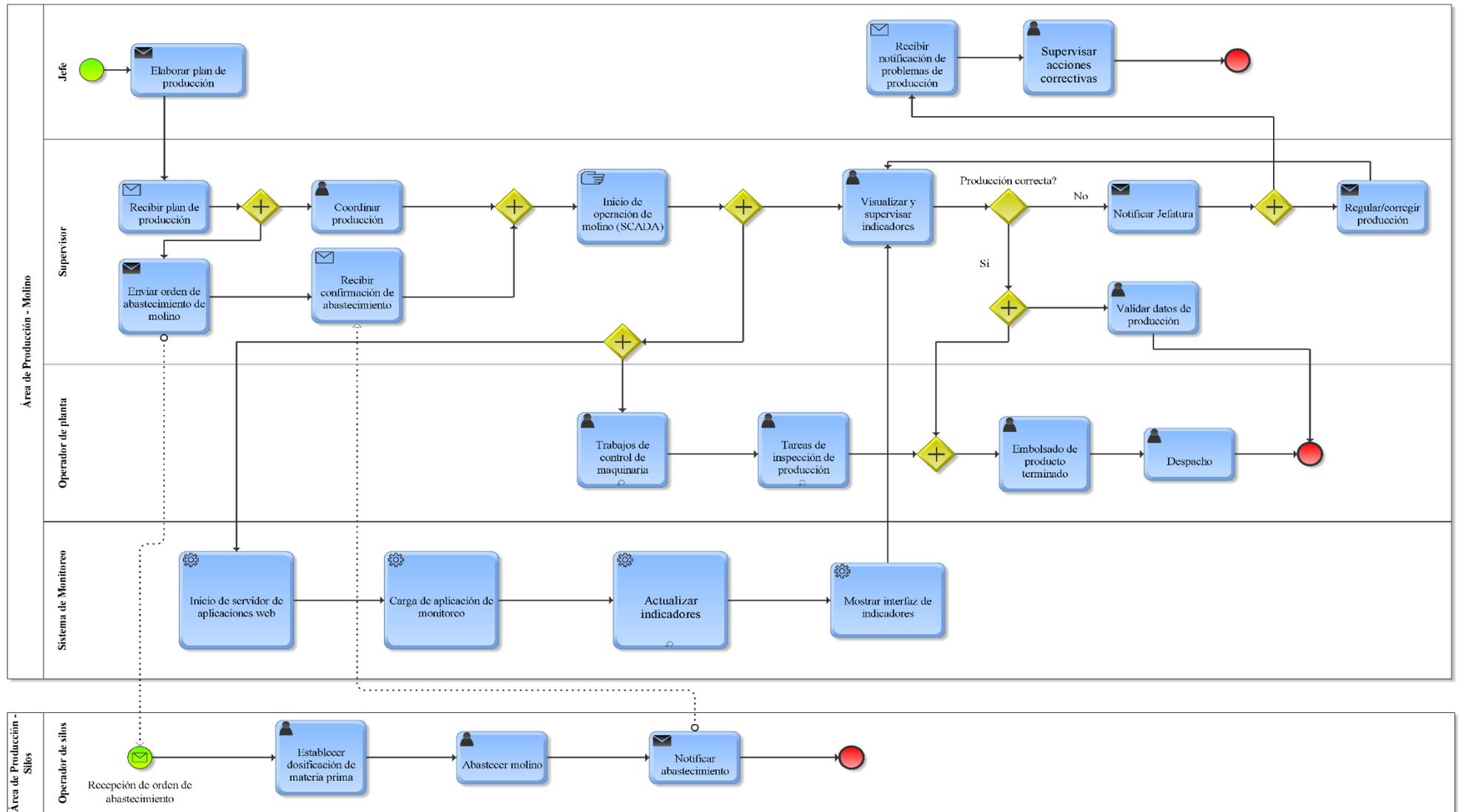
El primer problema que suele presentarse, es el de los atoros en el flujo de transporte de harina. Los ductos que llevan la harina hacia la etapa de balanceado tienen disposición a bloquearse o atorarse con el mismo producto. Esto se evidenciaba con el sobre flujo de harina y el consecuente desborde, causando pérdida de producto en las líneas de transporte.

Por otro lado, una mala regulación del molino ocasiona una producción defectuosa y e insuficiente, siendo percibido esto tras largos períodos de tiempo o al ya estar finalizando el turno en curso. A veces la toma de datos en el transcurso del turno ayuda a detectar este tipo de inconvenientes, pero la información no es oportuna y requiere tiempo de recopilación de datos y de procesamiento.

3.4 FASE DE EJECUCIÓN

Tras haber identificado la falencia principal para mejorar el proceso, la cual es la falta de información, podemos diseñar un proceso en el cual un sistema automatizado sea el responsable de tomar los datos de manera constante, procesarlos y mostrarlos oportunamente para que las decisiones sean tomadas de manera oportuna y correcta, generando así mayor valor al proceso. Por otro lado, se puede añadir la capacidad de analizar la información histórica para poder descubrir patrones de comportamiento o tendencias que no son visibles al analizar información instantánea. La ingeniería, el proceso de análisis y diseño del sistema se plasma en los siguientes capítulos. Para finalizar modelaremos el proceso buscado donde se hace uso del sistema desarrollado.

Imagen 5.- Proceso con implementación de Sistema



Fuente: Elaboración propia

CAPÍTULO IV

INGENIERIA DEL PROYECTO

4.1 GENERALIDADES

En el presente capítulo, se detallará el proceso de análisis de problemas y de requerimientos dados para dar solución al problema que enfrentaba la empresa. Cabe mencionar que para recabar información de las necesidades de la empresa se realizó una entrevista a los interesados. Dicha entrevista se adjunta en el Anexo N°1.

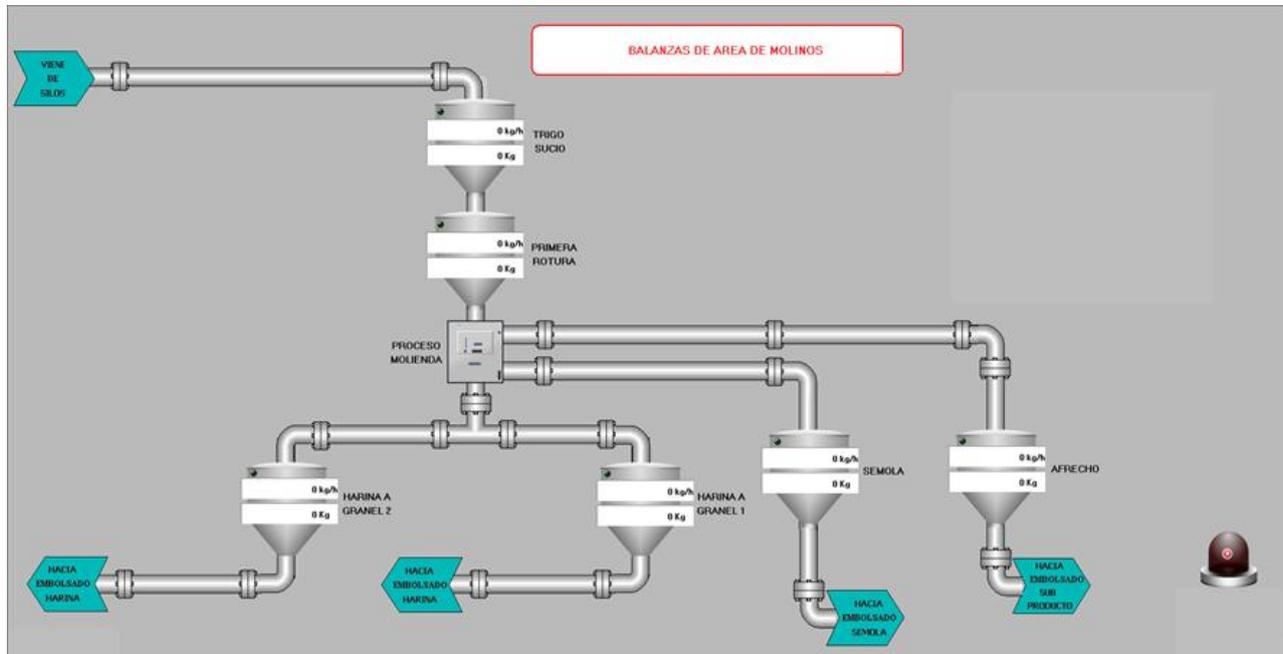
4.2 ETAPA DE ANALISIS

4.2.1 Identificación de problemas y oportunidades

La empresa cuenta con una línea de producción de harina en su planta ubicada en Arequipa, y cuenta con un sistema de Supervisión, Control y Adquisición de Datos (SCADA) que controla su infraestructura de producción que comprende los motores, balanzas, tolvas, control de flujos, etc.

En el SCADA se representan, a través de una pantalla, los valores actuales de los flujos (Kg/h) que identifican cada balanza, el valor del totalizado no borrrable de cada balanza, así como el estado de operación de cada elemento del sistema.

Imagen 6.- Interfaz de Control de Balanzas



Dada esta herramienta, actualmente parte del trabajo de los supervisores de cada turno es tomar nota del totalizado no borrable al final de su turno, y posteriormente se registra en un libro llamado Parte Diario de Producción. En este libro se lleva un registro por día principalmente de los totalizadores, y los eventos que suceden por cada turno.

El primer problema radica en la disponibilidad de información, ya que no se tiene un dato que pueda ser consultado a demanda sobre la producción en curso, considerando el estado actual de cada producto o las cantidades producidas al momento de la consulta. Para tener estos datos, como se había mencionado anteriormente, se espera al final del turno para la adquisición de datos manual, su tratamiento y cálculos respectivos que por otro lado, pueden ser errados. Existen casos específicos en los que se solicita directamente la toma de datos.

Señalado ese punto, nos enfrentamos al siguiente problema, la veracidad de los datos, ya que nunca se tiene una hora exacta del inicio o parada del molino, y en el parte diario se registra la hora aproximada según criterio del supervisor. Esto podría generar inconvenientes si se contrasta con los datos que mantenimiento podría tomar. De igual

manera, como ya se había mencionado, el proceso manual el cálculo de kilos producidos puede derivar en un error involuntario del supervisor, dando así datos erróneos o irreales.

El último problema se basa en el desconocimiento del funcionamiento del molino, es decir, si está funcionando de la manera correcta, si el proceso de producción está desarrollándose de la manera adecuada, si no existen atoros y si el flujo es el adecuado. Todo va de la mano con el desconocimiento de los valores actuales de producción.

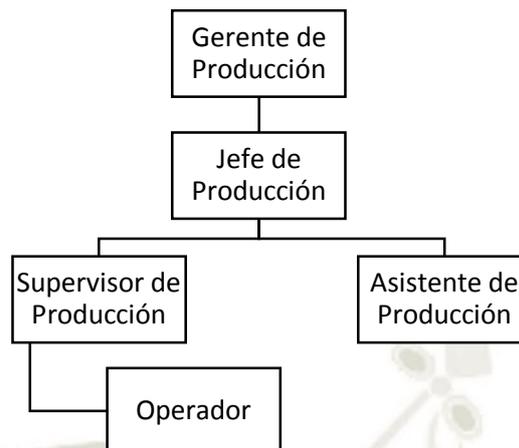
Cabe resaltar, que al culminar el día de producción, se sintetizan todos los datos registrados en el parte, y estos son ingresados a una plantilla en Excel. Esta plantilla con los datos básicos es enviada luego al jefe de Producción para su corroboración y posterior difusión a los interesados.

La falta de este manejo de datos puede ser ocasionada debido a la falta de presencia del departamento de TI o de un responsable de informática en el lugar; dicho departamento se encuentra ubicado en la sede principal en Lima y su principal relación con la sede de Arequipa es la de provisión de hardware para los usuarios finales, así como el manejo de contratos de arrendamientos de equipos de networking, una mesa de ayuda y contratos de soporte de equipos y redes. Todo tipo de gestión de información, datos, usuarios, equipos, permisos se centraliza en la sede corporativa.

Por último, cabe resaltar que cualquier desarrollo adicional sobre el SCADA que actualmente se encuentra en funcionamiento requiere de un presupuesto adicional para adquirir la licencia sobre un nuevo grupo de TAGS que permitan dicho desarrollo.

Los principales problemas fueron identificados de la mano del Jefe de Producción. Con él se conversó también quienes serían los interesados y se armó una pequeña jerarquía como se puede apreciar a continuación:

Imagen 7.- Jerarquía del área de Producción



Fuente: Elaboración propia

En el área de producción sólo se cuenta con un asistente de producción, luego existen tres supervisores desempeñándose uno por cada turno de producción. Por último se encuentran los operadores, los encargados del manejo directo de la infraestructura de operaciones.

Son los supervisores quienes desempeñan el papel más importante ya que son los encargados de realizar la captura de los datos de producción como se había mencionado anteriormente. Es en base a dichos datos que se arma el resto de información que llega a la jefatura.

4.2.2 Determinación de los requerimientos de información

Para poder construir un sistema eficiente, es necesario determinar en primera instancia los datos que se van a trabajar. Se tiene como datos base la cantidad de productos tales como harina, sémola, afrecho, trigo sucio y la primera rotura. Tras haber entrevistado al jefe de producción y al gerente de producción, se determinó con los mismos que un objetivo era llevar el control de tres indicadores planteados.

El primer indicador llamado “KPI-1” que da a conocer el índice del producto aprovechado o extraído del trigo. El segundo indicador es llamado “KPI-2” donde se toma en consideración el afrecho y por último el indicador de “KPI-3”

que puede entenderse como un índice del rendimiento real de la planta de producción.

A continuación definiremos las fórmulas que se precisaron para el cálculo de cada indicador.

• **KPI-1**

$$KPI - 1 = \frac{100 * (Harinas + Semola)}{Primera Rotura}$$

• **KPI-2**

$$KPI - 2 = \frac{100 * (Harinas + Semola + Afrecho - Primera Rotura)}{Primera Rotura}$$

• **KPI-3**

$$KPI - 3 = \frac{\text{Toneladas de Harina producida (Tn)}}{\text{Tiempo de funcionamiento del motor (h)}} \\ 17.7 \frac{Tn}{h} * (\text{Valor de Extracción})$$

Por lo tanto, es imperativo obtener los datos de las siguientes balanzas:

- Balanza de harina 1
- Balanza de harina 2
- Balanza de sémola
- Balanza de afrecho
- Balanza de primera rotura

Adicionalmente, se requiere capturar las jornadas de funcionamiento del molino, por lo que se tomará en cuenta el tiempo de funcionamiento del motor de transmisión 1, de acuerdo a lo conversado con el jefe de producción y gerente de producción.

El software de control que actualmente se encuentra funcionando en el molino es un SCADA que pertenece a la compañía Wonderware, propiamente es el software InTouch. El mismo tiene una serie de bondades que pueden ser trabajadas, pero para los fines buscados se requiere de una adquisición de nuevas licencias, lo que no encaja en el presupuesto disponible.

Dentro de las facultades de este software se encuentra la conexión a bases de datos SQL a través de un conector ODBC, la cual fue considerada oportuna de explotar para beneficio del proyecto. Para tal fin fue necesario identificar los TAGS que contienen la información que buscamos recolectar. Dentro de la relación de tags usados en el SCADA se pudo identificar el nombre de cada tag que contiene el valor del totalizador de cada balanza, lo cual se puede observar en la siguiente imagen:

Imagen 8.- Relación de tags

Tagname	Tag Type
B9_5	I/O Discrete
BALANZA_NE_DATA_SHOW	Memory Disc...
BALANZA_TREND_ACTUAL_RATE	Indirect Analog
BALANZA_TREND_OPERACION	Indirect Discrete
BALANZA_TREND_READY	Indirect Discrete
BALANZA_TREND_TITULO	Indirect Mess...
BALANZA_TREND_TOTAL_WEIGHT	Indirect Analog
BALANZA_TREND_TOTAL_WEIGHT_NE	Indirect Analog
BALANZA1_ACTUAL_RATE	I/O Real
BALANZA1_BIT	Memory Disc...
BALANZA1_OPERACION	I/O Discrete
BALANZA1_READY	I/O Discrete
BALANZA1_TITULO	Memory Mes...
BALANZA1_TOTAL_WEIGHT	I/O Real
BALANZA1_TOTAL_WEIGHT_NE	I/O Real
BALANZA2_ACTUAL_RATE	I/O Real
BALANZA2_OPERACION	I/O Discrete
BALANZA2_READY	I/O Discrete
BALANZA2_TITULO	Memory Mes...
BALANZA2_TOTAL_WEIGHT	I/O Real
BALANZA2_TOTAL_WEIGHT_NE	I/O Real
BALANZA3_ACTUAL_RATE	I/O Real
BALANZA3_OPERACION	I/O Discrete
BALANZA3_READY	I/O Discrete
BALANZA3_TITULO	Memory Mes...
BALANZA3_TOTAL_WEIGHT	I/O Real
BALANZA3_TOTAL_WEIGHT_NE	I/O Real
BALANZA4_ACTUAL_RATE	I/O Real
BALANZA4_OPERACION	I/O Discrete
BALANZA4_READY	I/O Discrete
BALANZA4_TITULO	Memory Mes...
BALANZA4_TOTAL_WEIGHT	I/O Real
BALANZA4_TOTAL_WEIGHT_NE	I/O Real
BALANZA5_ACTUAL_RATE	I/O Real
BALANZA5_OPERACION	I/O Discrete
BALANZA5_READY	I/O Discrete
BALANZA5_TITULO	Memory Mes...
BALANZA5_TOTAL_WEIGHT	I/O Real
BALANZA5_TOTAL_WEIGHT_NE	I/O Real
BALANZA6_ACTUAL_RATE	I/O Real
BALANZA6_OPERACION	I/O Discrete
BALANZA6_READY	I/O Discrete
BALANZA6_TITULO	Memory Mes...
BALANZA6_TOTAL_WEIGHT	I/O Real
BALANZA6_TOTAL_WEIGHT_NE	I/O Real
bb	Memory Disc...

Adicionalmente se debe considerar identificar el tag correspondientes al funcionamiento del motor de transmisión 1.

Se toma en consideración también posibles requerimientos futuros, como el control del funcionamiento de la balanza principal de harina, tiempos de parada de motor o posibles tiempos en que ocurre una parada de emergencia. De igual manera, se considerará entre las balanzas la balanza de trigo sucio.

Tras haber revisado la lista disponible se logró identificar el nombre de los tags requeridos, determinando que se recolectará los datos de los siguientes tags:



Tabla 1.- Tabla de Tags requeridos

Nombre del TAG	Descripción
BALANZA1_TOTAL_WEIGHT_NE	Balanza de Harina 1
BALANZA2_TOTAL_WEIGHT_NE	Balanza de Harina 2
BALANZA3_TOTAL_WEIGHT_NE	Balanza de Trigo Sucio
BALANZA4_TOTAL_WEIGHT_NE	Balanza de primera rotura
BALANZA5_TOTAL_WEIGHT_NE	Balanza de sémola
BALANZA6_TOTAL_WEIGHT_NE	Balanza de afrecho.
B374_SS	Bit de estado del motor de Transmisión 1.
BALANZA1_BIT	Bit de estado de la balanza de harina 1
PAHEG_SS	Bit de estado de parada de emergencia

Fuente: Elaboración propia

Para finalizar, se obtuvo información de los turnos que se aplican en producción, los cuales se desempeñan de lunes a domingo y son:

Tabla 2.- Tabla de Turnos de Producción

TURNO	HORARIO
A	05:00 a 13:00
B	13:00 a 21:00
C	21:00 a 05:00

Fuente: Elaboración propia

4.2.3 Análisis de las necesidades del sistema

Una vez se ha identificado los problemas y las necesidades de información, se puede determinar que el sistema:

- Deberá ser capaz de recolectar de manera automática los datos desde el SCADA.
- El sistema hará uso de un conector ODBC para lograr la conexión del SCADA a una base de datos SQL

- Se recolectarán datos cada 15 segundos desde el SCADA y se creará un registro en la base de datos, específicamente para los datos provenientes de las balanzas.
- Deberá ser posible calcular la producción real de cada elemento.
- Para los datos de funcionamiento del motor de transmisión 1, se deberá basar en eventos, es decir, por cada evento de cambio de estado del bit del motor se deberá generar un registro en la base de datos.
- Se deberá poder explotar los datos de manera oportuna y poder armar reportes en base a los mismos.
- Se dispondrá de una interfaz de monitoreo en tiempo real de los indicadores de producción.
- Las aplicaciones implementadas o desarrolladas deberán ser accesibles desde cualquier punto dentro de la red local de trabajo, por lo que se opta por aplicaciones web.

4.2.4 Análisis de recursos disponibles

Con fin de lograr un óptimo desarrollo, se toma en cuenta las características del hardware de la estación de trabajo sobre la cual se implementará la mayoría de la funcionalidad de la solución:

- Procesador: Intel Core i5 6400T 2.20 GHz (4 CPUs).
- Memoria RAM: 8.00 GB.
- Disco Duro: 500 GB, dos particiones.
- Comunicación: vía Ethernet, IP estática.
- Sistema Operativo: Windows 10 de 64 bits.

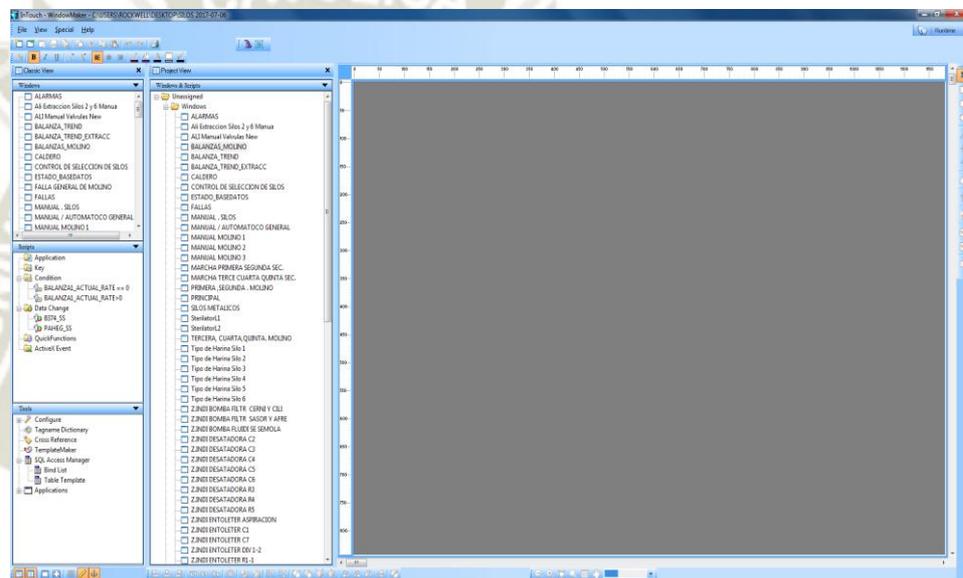
4.2.5 Herramientas de trabajo

Es importante definir con que software hemos de trabajar. En vista del presupuesto, y para lograr un ahorro de costos en licenciamientos, optamos por trabajar con herramientas y software de licencia libre de tipo Open Source siempre que sea posible. Por lo tanto:

- **SCADA.- Intouch WindowMaker**

Para lograr la programación de la funcionalidad en el sistema SCADA, se usará el software Intouch – WindowMaker, propiedad de Wonderware, cuyo software cuenta con licencias para su funcionamiento. En este se puede realizar la configuración y desarrollo de nuevas ventanas, asignación de TAGS, conexión ODBC, etc. Este software será proporcionado por la empresa ya que cuentan con la licencia actualmente activa.

Imagen 9.- Intouch Window Maker



- **Lenguaje de Programación y herramienta de desarrollo.**

Uno de los lenguajes más populares, a la fecha de la elaboración del proyecto, es el lenguaje JAVA. Este lenguaje es muy utilizado para aplicaciones web por lo que se considera el ideal para la tarea por cumplir, además de ser de conocimiento previo del autor de este tema. Tras haber definido el lenguaje, se considera que el uso de frameworks o entornos de trabajo facilitan el desarrollo de soluciones. Siguiendo la arquitectura Modelo – Vista – Controlador, se selecciona Play Framework v 2.2.4, una herramienta Open Source para el desarrollo de aplicaciones web, también de amplio conocimiento del autor, que cumple ciertas características relevantes para un acelerado desarrollo:

- Es una herramienta sumamente ligera. Sólo se requiere un editor de texto para codificar y un navegador web para observar resultados, ya que el trabajo se ejecuta desde la consola del sistema operativo.
- Es una herramienta totalmente reactiva, cumpliendo con el estándar público especificado en: (<https://www.reactivemanifesto.org/>):
 - ✓ *Responsive*, es decir, que permite que los problemas sean identificados rápidamente y corregidos efectivamente. De gran ayuda para el manejo de errores.
 - ✓ *Resilient*, El sistema se mantiene respondiendo aún al enfrentar una falla.
 - ✓ *Elastic*, El sistema se mantiene funcionando bajo una carga de trabajo variable.
- Amigable para el desarrollador, con amplia documentación y alta colaboración de la comunidad.

Imagen 10.- Logo Play Framework



- **Base de datos**

En cuanto al motor de base de datos, y debido a la inexistencia de acceso a uno para los requerimientos del área de Producción, se optó por buscar un motor de base de datos de licencia libre, que sea robusto, con altas capacidades de procesamiento de datos.

Es así como nos encontramos entre dos lo de los motores de licencia libre más utilizados en la industria, MySQL y PostgreSQL. Para llevar a cabo la elección sobre uno de ellos es necesario analizar sus características principales:

Tabla 3.- Comparativo de Bases de Datos

	 PostgreSQL v9.6	 MySQL v5.7
Open Source	Completamente Open Source.	Open Source, aunque de propiedad de Oracle.
Modelo ACID	Soportado.	Soportado con configuración específica en InnoDB.
Estándar SQL	Altamente compatible.	Parcialmente compatible.
Concurrencia	MVCC implementado.	MVCC únicamente en InnoDB.
Desempeño	Mejor utilizado en sistemas que requieren la ejecución de consultas complejas. Buen desempeño en aplicaciones de BI pero más apropiado para análisis de datos y DW.	Más orientado a proyectos basados en web. InnoDB provee de buena velocidad en escenarios OLTP. Funciona muy bien con aplicaciones de BI por requerir de alta velocidad de lectura.
Soporte NoSQL/JSON	Múltiples características de soporte.	Solo JSON.
Métodos de Acceso	Soporta todos los estándares.	Soporta todos los estándares.
Seguridad	Basado en SSL.	Basado en ACL.
Vistas Materializadas	Soportado	No Soportado
Tablas temporales	Soportado	Soportado
CTEs y Funciones Window	Soportado	No Soportado.

Fuente: <https://www.2ndquadrant.com/en/postgresql/postgresql-vs-mysql/>

Para los fines que se plante utilizar los dos cumplen casi todo lo requerido, a excepción de la capacidad de hacer uso de tablas temporales, el uso de Funciones Window y las Expresiones de Tabla Común (Common Table Expressions, CTE), siendo estas últimas características muy importantes al realizar análisis de datos.

Adicionalmente se señalan unos puntos más específicos:

- MySQL no soporta Full Outer Joins.
- MySQL no soporta uso de Arrays.
- PostgreSQL tiene mejor interface al trabajar con fechas y horas.

De acuerdo a lo señalado anteriormente es que se opta por el uso del software PostgreSQL en su versión 9.6.

Se debe considerar que tras la culminación de este proyecto, se dio a conocer oficialmente la versión 8.0 de MySQL con fecha 19 de Abril del 2018, la cual recibió la implementación de poder manejar CTE's y funciones tipo Window.

- **Reportes y servidor de reportes**

De acuerdo a los requerimientos de los usuarios finales, los reportes deberán ejecutarse de manera automática y ser enviados con una frecuencia diaria a las 07:00 am a los correos corporativos de usuarios específicos por cada reporte. Además estos reportes deben ser almacenados en la unidad disco dura a medida de respaldo.

Dado los requerimientos, para el desarrollo se opta por la herramienta Jasper Report Server, de propiedad de Jaspersoft; esta herramienta es muy versátil en su edición de la comunidad, la cual se distribuye bajo licencia libre y Open Source, funcionando sobre un servidor web como por ejemplo Apache Tomcat, con el que es altamente compatible. Para el diseño de los reportes se usará la herramienta iReport, la cual no requiere instalación alguna, más que descomprimir la descargar y hacer uso del ejecutable. La elección de esta herramienta es completamente por el dominio sobre la misma por parte del autor y la experiencia previa, como herramienta libre y Open Source.

Otra opción sobre la que se realizó la evaluación, es la herramienta Pentaho. Al momento del desarrollo, dicha herramienta, en específico Pentaho Server, presentaba algunos bugs y errores en la ejecución de reportes automatizados y enviados automáticamente vía correo, por lo que no se consideró oportuno su

uso. Aun así, a opinión personal del autor, la herramienta de diseño de reportes Pentaho Report Designer, puede resultar más amigable para el usuario y más fácil en el uso para el diseño de reportes.

Cabe resaltar que una herramienta que cumple también todo lo necesario es BIRT, software diseñado para trabajar con la plataforma de Eclipse. BIRT posee también un diseñador de reportes que puede instalarse como plugin a Eclipse, o descargar el SDK para el diseño de reportes. Para acceder a una interfaz gráfica amigable y poder hacer uso de características semi avanzadas es necesario instalar la herramienta Birt iHub F-Type.

Imagen 11.- Logo TIBCO Jaspersoft



- **Software para análisis de datos**

En el punto de análisis de datos, se optó por revisar las características disponibles por parte del software Jaspersoft, encontrando limitaciones por parte de este software en su versión de la comunidad, sobre todo en su herramienta de análisis de datos, la cual al momento del desarrollo, en versión de la comunidad no tiene implementado la funcionalidad de crear reportes interactivos de tipo cubo OLAP, que permita la navegación y elaboración del reporte a criterio del usuario final.

Tras haber revisado este punto, se opta por la herramienta de Pentaho.

Imagen 12.- Logo Hitachi Pentaho



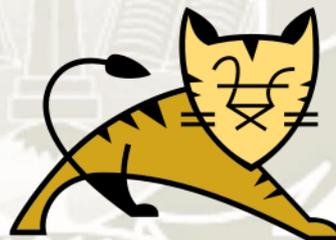
- **Servidor de Aplicaciones Web**

Entre las aplicaciones web, se hará uso de Apache Tomcat versión 8.5, por las siguientes características:

- Experiencia del autor en el uso y entendimiento de la herramienta.
- Herramienta altamente popular como servidor web de soluciones JAVA.
- Open Source, con licencia libre.
- Amplia documentación y apoyo de la comunidad desarrolladora.
- Alta compatibilidad con las herramientas seleccionadas de reportes y análisis de datos.

Se debe tener en cuenta que para el uso de esta herramienta es indispensable tener instalado en el equipo el software JDK (java Development Kit) y JRE (Java Runtime Environment).

Imagen 13.- Logo de Apache Tomcat

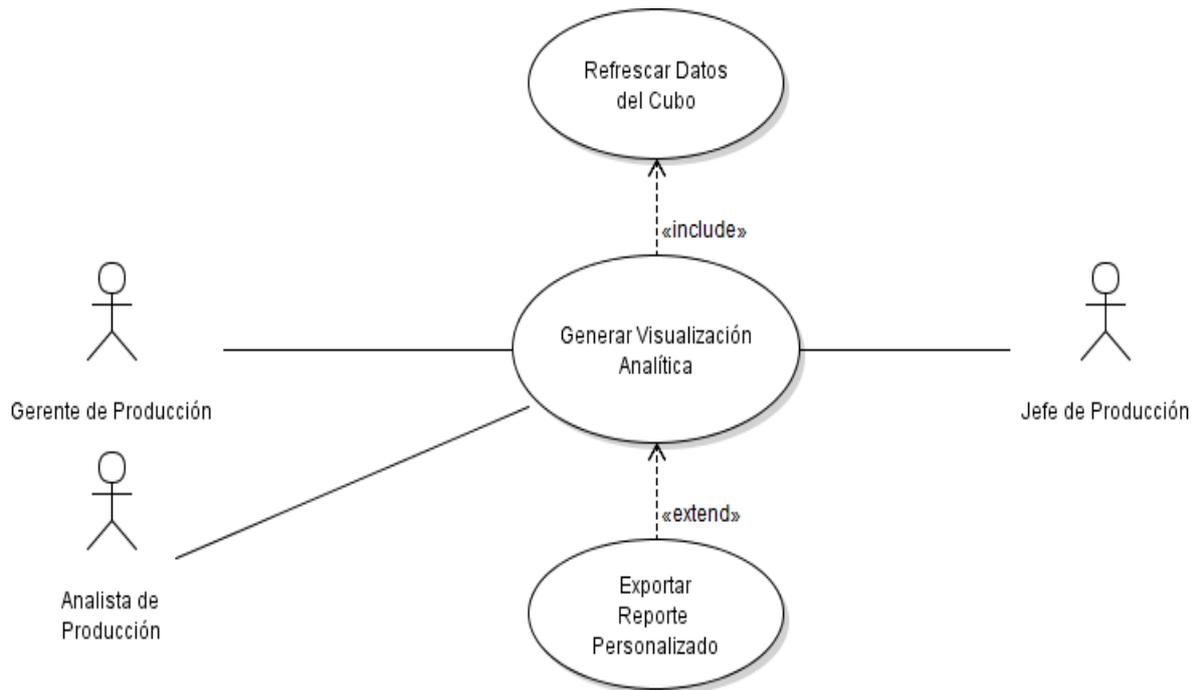


Cabe mencionar que para los fines perseguidos, esta herramienta cumple totalmente lo que se requiere.

4.2.6 Diagrama de Casos de Uso

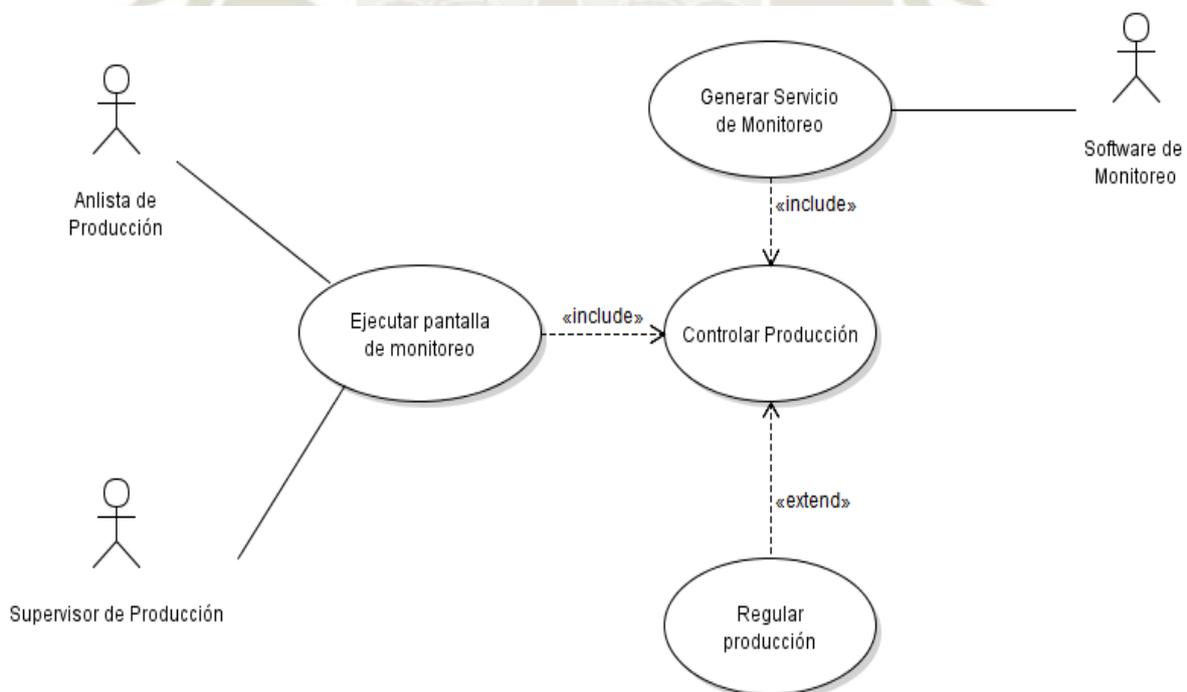
De acuerdo al análisis de requerimientos y necesidades anteriormente descritas, se elaboraron los siguientes casos de uso.

Imagen 14.- Caso de Uso de Software de Análisis de Datos



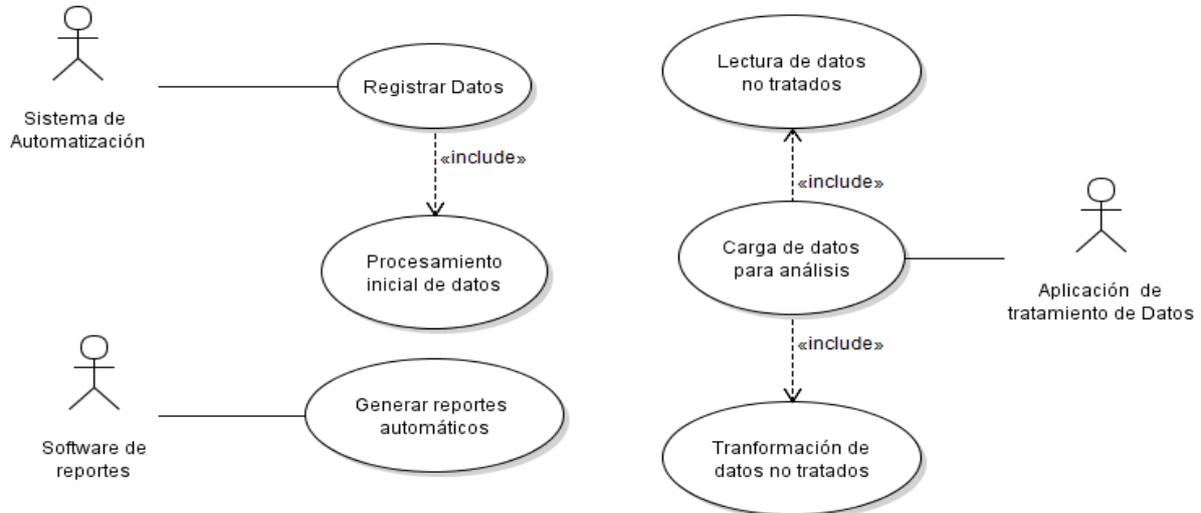
Fuente: Elaboración propia

Imagen 15.- Caso de Uso de Software de Monitoreo



Fuente: Elaboración propia

Imagen 16.- Caso de Uso Tratamiento de Datos

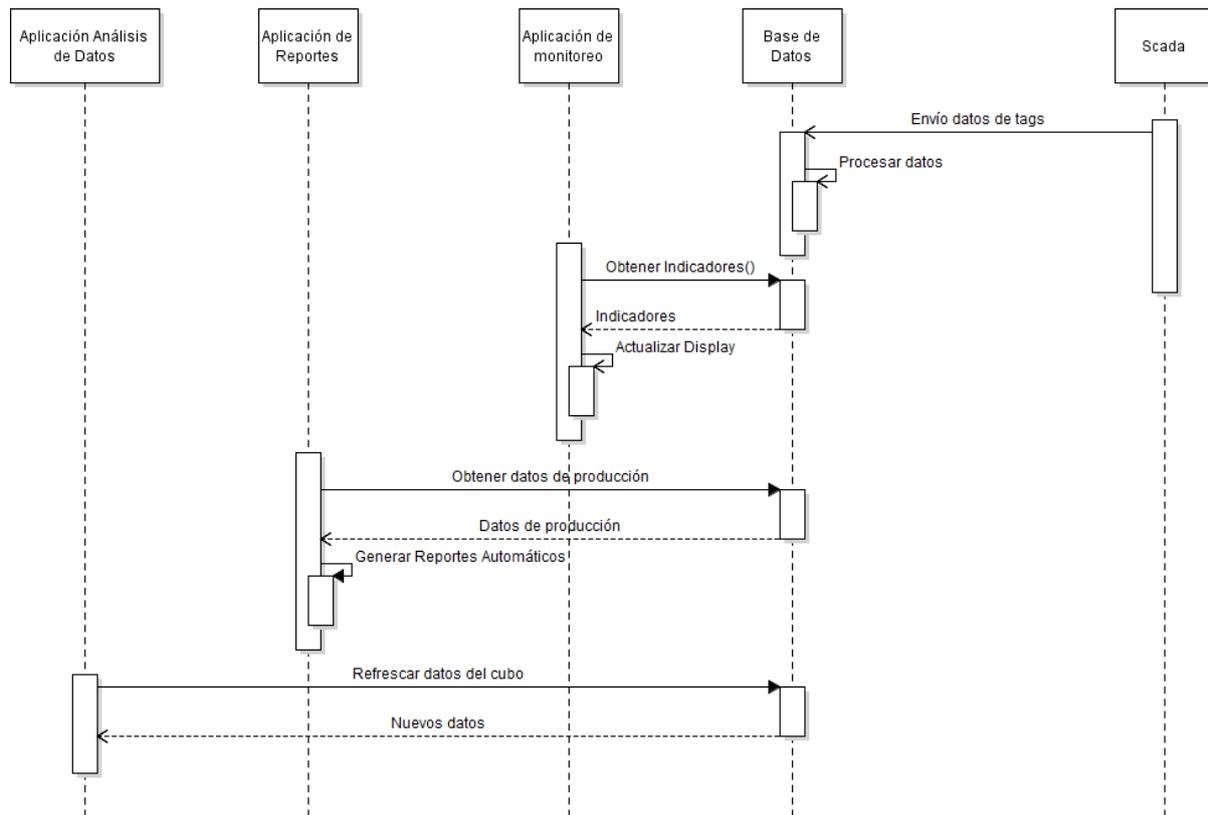


Fuente: Elaboración propia

4.2.7 Diagrama de Secuencia

Para finalizar la etapa de análisis se debe considerar el siguiente diagrama de secuencia, que determina el uso de los servicios o sistemas. En el mismo se consideran como actores el sistema SCADA, la base de datos, la aplicación de monitoreo desarrollada, y las dos aplicaciones tipo Open Source utilizadas tanto para la gestión y automatización de reportes como la que permite realizar análisis de datos.

Imagen 17.- Diagrama de Secuencia



Fuente: Elaboración propia

4.3 ETAPA DE DISEÑO

4.3.1 Base de Datos

La base de datos que se ha de definir debe ser capaz de contener los datos de los totalizadores acompañados de la respectiva hora y fecha de registro. Se debe tener en consideración que el registro debe corresponder a un turno específico. Adicionalmente se debe registrar el funcionamiento del motor, el cual se identifica por un bit de funcionamiento. Las tablas que constituirán la base de datos quedan entonces definidas como a continuación:

- **Tabla de Totalizadores.-**

La tabla que recibirá los datos directamente desde el SCADA, acerca de los valores del totalizador de cada balanza

Tabla 4.- Definición de Tabla de Totalizadores

Nombre	Tipo de Dato	Descripción
Id_totalizador	Bigserial	Identificador de cada registro en la tabla
Harina_g1	Entero	Valor del totalizador de la balanza de harina a granel 1
Harina_g2	Entero	Valor del totalizador de la balanza de harina a granel 2
Trigo_Sucio	Entero	Valor del totalizador de la balanza de trigo sucio
Primera_rotura	Entero	Valor del totalizador de la balanza de primera rotura
Semola	Entero	Valor del totalizador de la balanza de sémola
Afrecho	Entero	Valor del totalizador de la balanza de afrecho
Fecha	Date	Fecha a la que pertenece la producción.
Hora	Time	Hora de registro. Formato HH:mm:SS
Fecha_real	Date	Fecha real de registro en base de datos.

Fuente: Elaboración propia

• **Tabla de turnos.-**

La tabla turnos contendrá el registro de los turnos que se desempeñan dentro del área de producción. Debido a que un turno (turno C) pasa de un día hacia el otro, es necesario considerar que la tabla registre dichos intervalos como múltiples campos en un solo registro, ya que un intervalo de tiempo que pasa de un día hacia otro no es reconocido como tal en una sentencia SQL. Por lo tanto se crea cuatro columnas en vez de dos para los intervalos de tiempo de cada turno. Se considera que tanto para los turnos A y B, los valores dentro de las columnas para el intervalo 2 serán nulos.

Tabla 5.- Definición de Tabla de Turnos

Nombre	Tipo de Dato	Descripción
Id_turno	Entero	Identificador de cada registro en la tabla
nombre	Character (10)	Descripción o nombre del turno registrado
Hora_inicio_1	Entero	Hora inicio de intervalo 1

Hora_fin_1	Entero	Hora de fin de intervalo 1
Hora_inicio_2	Entero	Hora inicio de intervalo 2
Hora_fin_2	Entero	Hora de fin de intervalo 2

Fuente: Elaboración propia

- **Tabla de Producción.-**

Esta tabla es la que contendrá el valor real producido de cada flujo que ha pasado por las balanzas. Es decir, esta tabla contendrá los valores de kilogramos que fueron pesados realmente por las balanzas.

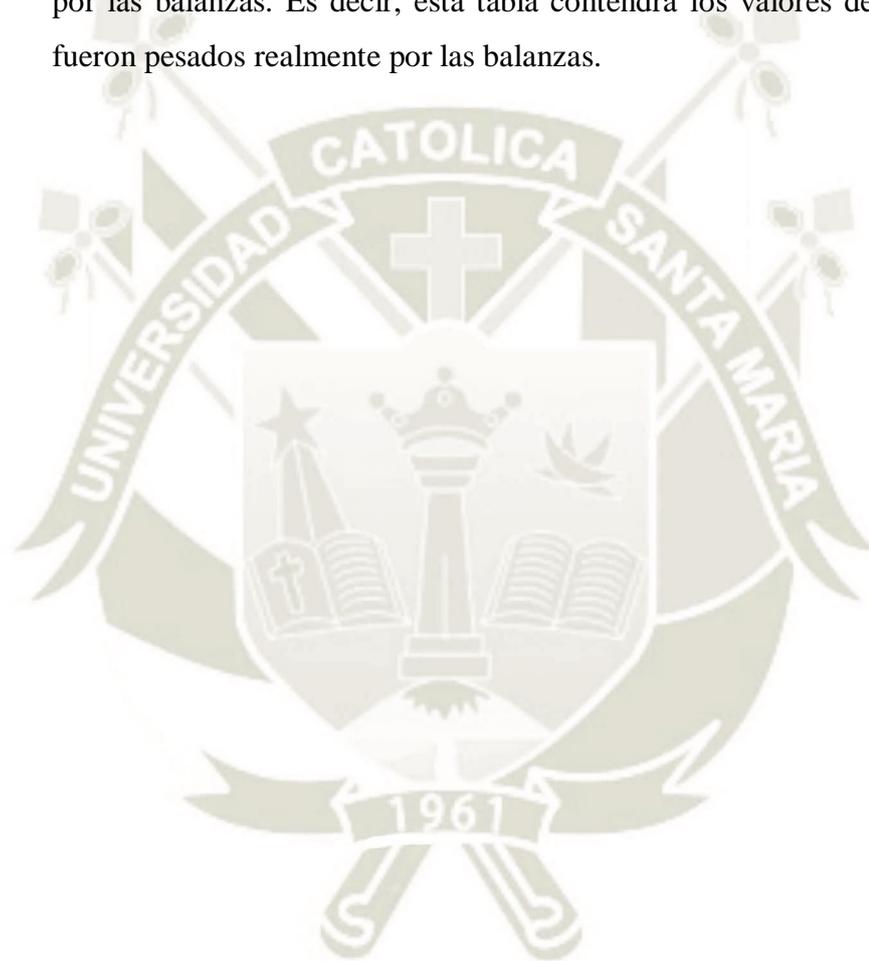


Tabla 6.- Definición de Tabla de Producción

Nombre	Tipo de Dato	Descripción
Id_produccion_turno	Bigserial	Identificador de cada registro en la tabla
Fecha_produccion	Character (10)	Descripción o nombre del turno registrado
Hora	Integer	Hora inicio de intervalo 1
Id_turno	Integer	Hora de fin de intervalo 1
Harina_g1	Integer	Valor de kilogramos producidos según la balanza de harina a granel 1
Harina_g2	Integer	Valor de kilogramos producidos según la balanza de harina a granel 2
Trigo_Sucio	Integer	Valor de kilogramos producidos según la balanza de trigo sucio
Primera_rotura	Integer	Valor de kilogramos producidos según la balanza de primera rotura
Semola	Integer	Valor de kilogramos producidos según la balanza de sémola
Afrecho	Integer	Valor de kilogramos producidos según la balanza de afrecho

Fuente: Elaboración propia

• **Tabla de Bits de Estado.-**

En esta tabla se registrarán los bits de estado tanto del motor de transmisión 1, de la balanza de harina y del circuito de parada de emergencia. Cabe resaltar que se recibirá eventos de cambio de estado de cada bit, es decir, el cambio de uno solo de los bits generará un registro.

Tabla 7.- Definición de Tabla de Bits de Estado

Nombre	Tipo de Dato	Descripción
Id	Integer	Identificador de cada registro en la tabla
Bitmotormolino	Smallint	Bit de funcionamiento de molino
Bitbalanzamolino	Integer	Bit de funcionamiento de balanza

Bitpemergenciamolino	Integer	Bit de funcionamiento de circuito de parada de emergencia
Fecha	Date	Fecha de registro de evento
hora	Time	Hora de registro de evento

Fuente: Elaboración propia

• **Tabla de Motor Activo.-**

Contendrá los lapsos de tiempo de funcionamiento del motor.

Tabla 8.- Definición de Tabla de Motor Activo

Nombre	Tipo de Dato	Descripción
Id	Integer	Identificador de cada registro en la tabla
Tipofuncionamiento	Text	Descripción de funcionamiento
Horainicio	Time	Hora de inicio del intervalo
Fechainicio	Date	Fecha de inicio del intervalo
Horafin	Time	Hora de fin del intervalo
Fechafin	Date	Fecha de fin del intervalo
Idbitinicio	Integer	Identificador del registro en tabla de bits que corresponde al inicio del intervalo
idbitfin	Integer	Identificador del registro en tabla de bits que corresponde al fin del intervalo.
Totalfuncionamiento-segundos	Integer	Tiempo total en segundos del intervalo de funcionamiento.
Activo	Smallint	Flag de registro activo

Fuente: Elaboración propia

• **Tabla de Molino.-**

Contendrá los lapsos de tiempo de funcionamiento de la balanza.

Tabla 9.- Definición de Tabla de Molino

Nombre	Tipo de Dato	Descripción
Id	Integer	Identificador de cada registro en la tabla
Tipofuncionamiento	Text	Descripción de funcionamiento
Horainicio	Time	Hora de inicio del intervalo
Fechainicio	Date	Fecha de inicio del intervalo
Horafin	Time	Hora de fin del intervalo
Fechafin	Date	Fecha de fin del intervalo
Idbitinicio	Integer	Identificador del registro en tabla de bits que corresponde al inicio del intervalo
idbitfin	Integer	Identificador del registro en tabla de bits que corresponde al fin del intervalo.
Totalfuncionamiento-segundos	Integer	Tiempo total en segundos del intervalo de funcionamiento.
Activo	Smallint	Flag de registro activo

Fuente: Elaboración propia

• **Tabla de Paradas de emergencia.-**

Contendrá los intervalos en que ocurre una parada de emergencia.

Tabla 10.- Definición de Tabla de paradas Emergencia

Nombre	Tipo de Dato	Descripción
Id	Integer	Identificador de cada registro en la tabla
Tipofuncionamiento	Text	Descripción de funcionamiento
Horainicio	Time	Hora de inicio del intervalo
Fechainicio	Date	Fecha de inicio del intervalo
Horafin	Time	Hora de fin del intervalo
Fechafin	Date	Fecha de fin del intervalo
Idbitinicio	Integer	Identificador del registro en tabla de bits que corresponde al inicio del intervalo
idbitfin	Integer	Identificador del registro en tabla de bits que corresponde al fin del intervalo.
Totalfuncionamiento-segundos	Integer	Tiempo total en segundos del intervalo de funcionamiento.
Activo	Smallint	Flag de registro activo

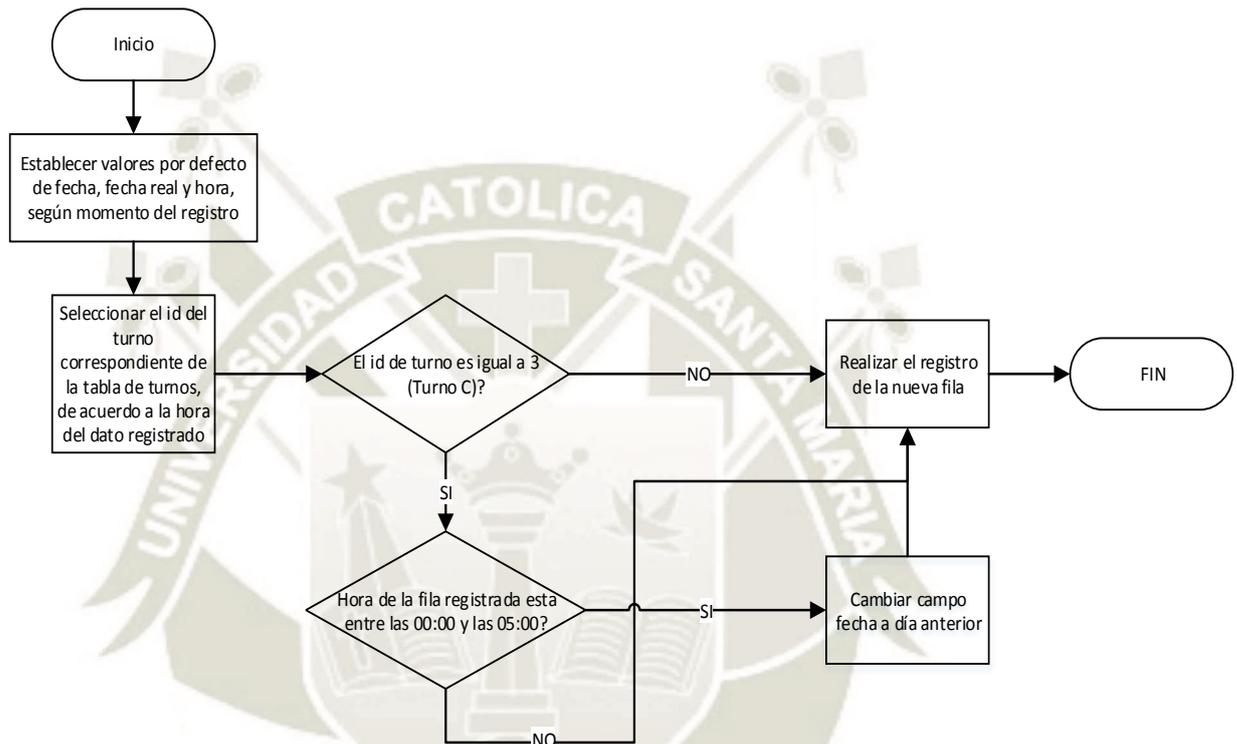
Fuente: Elaboración propia

Para que el llenado sea automatizado, se programarán funciones almacenadas que serán ejecutadas por disparadores, de manera que cada tabla pueda completar sus registros de manera automática. A continuación se presenta el flujo de la lógica de cada función.

• **Función sobre Totalizadores previo a registro.-**

La función se ejecuta previo al registro de una nueva fila en la tabla de totalizadores. Su principal función es de establecer los campos de fecha, fecha_real, hora y turno.

Imagen 18.- Flujo de función sobre tabla totalizadores previo al registro

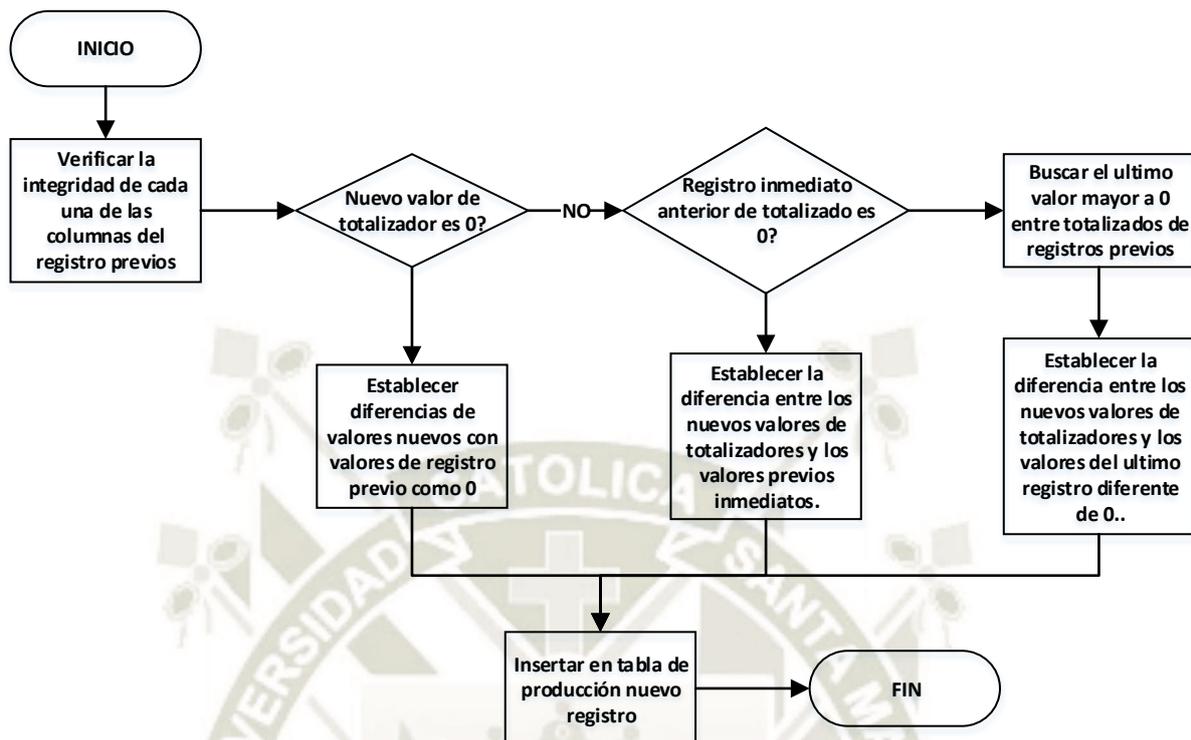


Fuente: Elaboración propia

• **Función sobre Totalizadores después del registro**

El objetivo de esta función es de enviar los datos de la producción real en kilogramos a la tabla de producción real.

Imagen 19.- Flujo de función sobre tabla totalizadores tras registro.

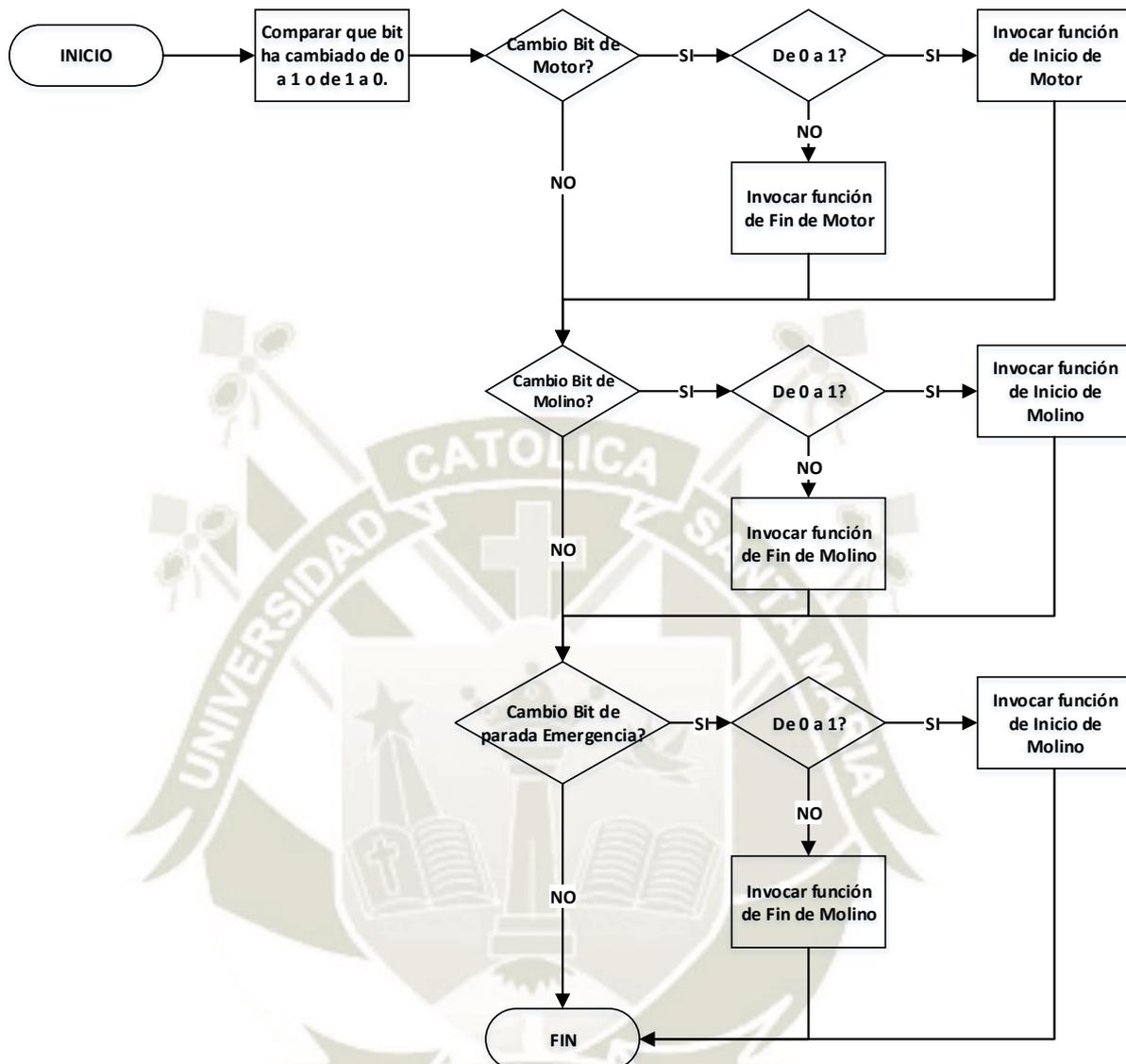


Fuente: Elaboración propia

- **Función sobre tabla de bits de estado**

Esta función se encargará de realizar el registro adecuado de funcionamiento tanto de motor de transmisión 1, la balanza de harina y las paradas de emergencia. Esta función se ejecuta sobre la tabla de bits de estado que recibe un registro cada vez que existe una variación en el bit de estado de algún PLC de los tres monitoreados.

Imagen 20.- Flujo de función sobre tabla bits de estado.



Fuente: Elaboración propia

4.3.2 Solución de monitoreo de Indicadores

Se recabó la necesidad del usuario, el cual plantea que quiere monitorear tres de sus indicadores de manera constante, a través de una interfaz que pueda ser observada desde cualquier punto de su red interna. Para tal fin se decide realizar un desarrollo completamente web, que pueda ser ejecutado desde el servidor elegido previamente, el cual es Apache Tomcat. La aplicación mostrará tres indicadores kpi-1, kpi-2 y kpi-3 con métodos de visualización diferentes agrupados cada indicador en una columna. El requerimiento indica que sobre cada columna estará una etiqueta que indicará el nombre del indicador al que

pertenece; luego se mostrará el valor presente del indicador en un recuadro calculado sobre el turno de producción en curso, seguido de la lista de valores del indicador de los tres turnos previos al actual.

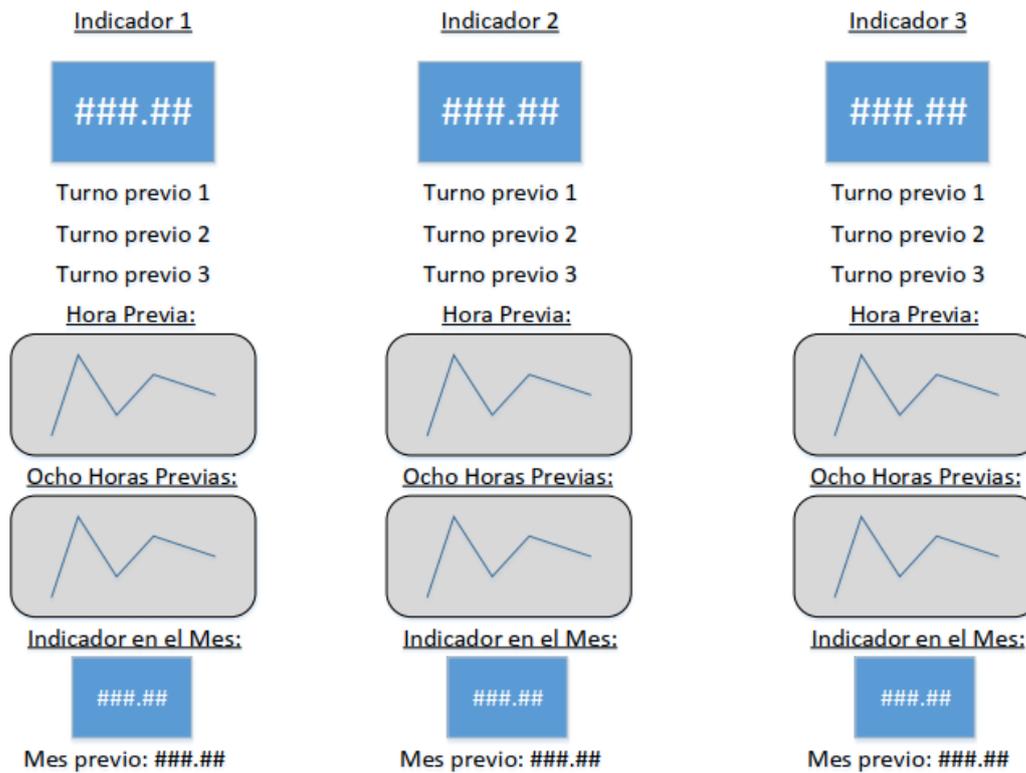
A continuación, se visualizará un gráfico lineal de la tendencia de dicho indicador en la última hora, graficando un punto por cada minuto con un total de 60 puntos. Posteriormente se empleará otro gráfico lineal que dibujará la tendencia del indicador en las últimas ocho horas, graficando un punto por cada minuto.

Para finalizar se mostrará otro recuadro con el valor actual del indicador calculado sobre el mes en curso, más visible que la etiquetas que la seguirá que representará el valor final del indicador del mes previo.

Esta interfaz deberá actualizarse cada cierto intervalo de tiempo, el cual se establece inicialmente como quince segundos, siendo capaz de recibir un parámetro que modifique dicho intervalo.

Cada nuevo valor que se represente en los gráficos añadirá un nuevo punto en la gráfica y borrará el punto del valor más antiguo en la gráfica, manteniendo el mismo número de puntos durante toda su visualización.

Imagen 21.- Diseño de Interfaz de Monitoreo



Fuente: Elaboración propia

4.3.3 Análisis de Datos

En el análisis de datos, se estableció que se requerirá tener las siguientes medidas listas:

- Harina
- Sémola
- Afrecho
- Primera Rotura
- Trigo Sucio
- Kpi-1
- Kpi-2
- Kpi-3
- Tiempo de funcionamiento de motor.

Una vez establecidos los valores que buscan ser medidos, debemos definir las dimensiones sobre los cuales se aplicarán los criterios de análisis. La dimensión básica incluye las fechas que comprende el año, el mes y el día, y adicionalmente se incluye la dimensión de tiempo considerando la hora y minuto; para finalizar se incluye como dimensión el turno de producción

correspondiente. Como resultado obtendremos las siguientes tablas que conformarán la topología estrella:

- **Dimensión de horas**

Tabla 11.- Dimensión de horas

Nombre	Tipo de Dato	Descripción
Id	Integer	Identificador de cada registro en la tabla
Horadeldía	Time	Hora completa que contiene hora, minuto y segundos.
Hora	Integer	Hora del día
Cuartohora	Texto	Texto que marca a que cuarto de hora pertenece (1ro,2do,3ro,4to)
Minutodia	Entero	Número de minuto en el día.
Segundo	Entero	Segundo del día
Jornada	Texto	Jornada del día (madrugada, día, tarde, noche)
Dianoche	Texto	Si pertenece al día o la noche (cada 12 horas)
Turno	Texto	
Minuto	Entero	Minuto del día

Fuente: Elaboración propia

- **Dimensión de fechas**

Tabla 12.- Dimensión de fechas

Nombre	Tipo de Dato	Descripción
Id	Integer	Identificador de cada registro en la tabla
Fecha	Fecha	Fecha completa. Contiene día, mes y año.
Anho	Entero	Año de la fecha
Mes	Entero	Mes de la fecha

Nombremes	Texto	Nombre del mes.
Dia	Entero	Dia de la fecha.
Diadelanho	Entero	Número del día en el año.
Nombredia	Texto	Nombre del día.
Semanacalendario	Entero	Número de semana en el año.
Cuarto	Texto	A que cuarto en el año pertenece.
Trimestre	Texto	A que trimestre pertenece.
Semestre	Texto	A que semestre pertenece.
Anhocuarto	Texto	Número del año y del cuarto correspondiente.
Anhomes	Texto	Número del año y del mes.
Finde semana	Texto	Si es fin de semana.
Feriadofijoperu	Texto	Si corresponde a feriado en Perú.
Iniciosemana	Fecha	La fecha de inicio de la semana de la fecha actual
Finsemana	Fecha	La fecha de fin de la semana de la fecha actual
Iniciomes	Fecha	La fecha de inicio del mes de la fecha actual
Finmes	Fecha	La fecha de fin del mes de la fecha actual

Fuente: Elaboración propia

• **Tabla de hechos de producción.-**

Tabla 13.- Tabla hechos de producción

Nombre	Tipo de Dato	Descripción
Id	Entero	Identificador de cada registro en la tabla
idfecha	Entero	Id de la fecha correspondiente en la dimensión de fechas.
Idhora	Entero	Id de la hora correspondiente en la dimensión de horas.
Fecha	Fecha	Fecha de datos.
Idturno	Entero	Id del turno correspondiente.
Harina_g1	Entero	Producción en kg de harina.

Harina_g2	Entero	Producción en kg de harina.
Sémola	Entero	Producción en kg de sémola.
Afrecho	Entero	Producción en kg de afrecho.
PrimeraRotura	Entero	Producción en kg de primera rotura.
TrigoSucio	Entero	Producción en kg de trigo sucio.
Tiempo_on	Entero	Tiempo en segundos de motor activo
Tiempo_off	Entero	Tiempo en segundos de motor inactivo.

Fuente: Elaboración propia

4.3.4 Reportes

Se identifican tres reportes que son de importancia del usuario, así se tomará en cuenta para su diseño:

- **Reporte de Indicadores.-**

Debe considerar una tabla con la producción de cada elemento, agrupados por turno y debe señalar los indicadores calculados por cada turno. Adicionalmente debe considerarse unos gráficos de línea de las tendencias de cada indicador. Por cada indicador se visualizará dos gráficas, una que respecta a la tendencia durante el día de producción y otra el promedio por cada día, de los 30 días previos.

La frecuencia de este reporte para su emisión será diaria y de un acumulado mensual (emitido el primer día de cada mes). Los receptores de este reporte serán: el gerente de producción, el jefe de producción, el jefe de mantenimiento y el analista de producción.

- **Reporte de Producción.-**

Al igual que el reporte de indicadores, considerará la misma tabla de producción de cada elemento, agrupado por turno y con sus respectivos indicadores. Adicionalmente incluirá una gráfica lineal de tendencia por cada elemento de producción, representando los valores totales por turno en los últimos 30 días. Los receptores de este reporte serán: el jefe de

producción, el analista y los supervisores. De igual manera la emisión será diaria y mensual.

○ **Reporte de eventos de Motor de Transmisión 1.-**

Este reporte contendrá los eventos de inicio y parada de motor dentro del mes en curso. Adicionalmente poseerá un gráfico de tipo pie que indicará en porcentaje el tiempo de funcionamiento y de parada de la planta del día emitido. Su frecuencia de emisión será diaria y mensual.



CAPÍTULO V

DESARROLLO E IMPLEMENTACION

5.1 GENERALIDADES.

En el presente capítulo se dará a conocer el proceso de desarrollo e implementación de lo requerido. Así se especifica el código fuente desarrollado para la puesta en marcha de los sistemas que se utilizaron.

5.2 METODOLOGÍA A EMPLEAR.

Dentro de los enfoques o metodologías más utilizadas para el desarrollo de proyectos de inteligencia de negocios, específicamente la elaboración del datamart o datawarehouse, analizaremos las tres más resaltantes, considerando que buscamos la metodología que sea lo suficientemente flexible, ya que el alcance del proyecto es pequeño. Así que al trabajar bajo la metodología escogida, ésta debe permitir modificar algunos pasos u obviar algunos si es necesario, proporcionando confianza y consistencia en la culminación de los pasos o tareas que se consideren dentro de la metodología:

5.2.1 Metodología de Kimball

Más conocida como la metodología de enfoque Bottom Up (de la base hacia arriba), esta metodología o enfoque se orienta a la construcción de repositorios de datos pequeños llamados datamarts, cada cual es específico a un área de la organización conteniendo todos los datos de interés de dicha área o si se quiere ver de otra forma, se construye un datamart por cada proceso principal del negocio.

5.2.2 Metodología de Inmon

Esta metodología es la opuesta a la de Kimball, y es que el enfoque es el inverso; se habla de un enfoque Top Down (del tope hacia abajo) en el cual se designan todos los esfuerzos en construir el data warehouse de manera óptima y eficiente, y luego a partir de este extraer los datos para cada área específica o proceso de negocio que requiera de la tecnología.

5.2.3 Metodología de Moss

Esta metodología es una de las más flexibles y considera la Justificación del proyecto, la planeación, el análisis de negocio, el diseño de la solución, la construcción de la solución diseñada y por último el despliegue de dicha solución. Dentro de cada etapa existen tareas que se deben realizar para llevar a cabo esta metodología, sin embargo, estos pasos o tareas se pueden obviar dependiendo del caso y magnitud del proyecto, siendo la autora propiamente quien señala los riesgos de saltar cada paso para evitar inconvenientes futuros.

Tras tener en cuenta estos tres enfoques, se decide hacer uso de una combinación de la metodología de Kimball por su orientación a compañías que recién inician con este tipo de proyectos y con la metodología de Moss por la flexibilidad que ofrece y la adaptación.

5.3 IMPLEMENTACIÓN DE BASE DE DATOS

A continuación se detallan los scripts utilizados para cada tabla o funciones de la base de datos:

5.3.1 Tabla de turnos

```
CREATE TABLE tabla_turnos (  
    id_turno integer NOT NULL,  
    nombre character(10),  
    hora_inicio_1 time without time zone,  
    hora_fin_1 time without time zone,  
    hora_inicio_2 time without time zone,  
    hora_fin_2 time without time zone,  
    CONSTRAINT "ID_Turnos_PK" PRIMARY KEY (id_turno)  
);
```

5.3.2 Tabla de totalizadores

```
CREATE TABLE public.tabla_totalizadores (
    id_totalizador bigint NOT NULL
    DEFAULT nextval('tabla_totalizadores_id_totalizador_seq'::regclass),
    harina_g1 integer,
    harina_g2 integer,
    trigo_sucio integer,
    primera_rotura integer,
    semola integer,
    afrecho integer,
    id_turno integer,
    fecha date,
    hora time without time zone,
    fecha_real date,
    CONSTRAINT "ID_PK" PRIMARY KEY (id_totalizador),
    CONSTRAINT fk_id_turno FOREIGN KEY (id_turno)
    REFERENCES tabla_turnos (id_turno) MATCH SIMPLE
    ON UPDATE NO ACTION ON DELETE NO ACTION
);
```

5.3.3 Tabla de producción

```
CREATE TABLE tabla_produccion_turno (
    id_produccion_turno bigint NOT NULL
    DEFAULT nextval('tabla_produccion_turno_id_produccion_turno_seq'::regclass),
    fecha_produccion date,
    hora time without time zone,
    id_turno integer,
    harina_g1 integer,
    harina_g2 integer,
    trigo_sucio integer,
    primera_rotura integer,
    semola integer,
    afrecho integer,
    procesadobi smallint DEFAULT 0,
    CONSTRAINT tabla_produccion_turno_pk PRIMARY KEY (id_produccion_turno),
```

```

CONSTRAINT fk_id_turno FOREIGN KEY (id_turno)
REFERENCES tabla_turnos (id_turno) MATCH SIMPLE
ON UPDATE NO ACTION ON DELETE NO ACTION
);
    
```

5.3.4 Tabla Bits de Estado

```

CREATE TABLE tabla_mvic_bitsestado (
id bigint NOT NULL DEFAULT nextval('tabla_mvic_bitsestado_id_seq'::regclass),
bitmotormolino smallint,
bitbalanzamolino smallint,
bitpemergiamolino smallint,
fecha date,
hora time without time zone,
CONSTRAINT tabla_mvic_bitsestado_pk PRIMARY KEY (id)
);
    
```

5.3.5 Tabla de Motor

```

CREATE TABLE tabla_mvic_motor (
id bigint NOT NULL DEFAULT nextval('tabla_mvic_motor_id_seq'::regclass),
tipofuncionamiento text,
horainicio time without time zone,
horafin time without time zone,
fechainicio date,
fechafin date,
idbitinicio integer,
idbitfin integer,
totalfuncionamientosegundos integer,
activo smallint,
CONSTRAINT tabla_mvic_motor_pk PRIMARY KEY (id),
CONSTRAINT tabla_mvic_motor_fk_bitfin FOREIGN KEY (idbitfin)
REFERENCES tabla_mvic_bitsestado (id) MATCH SIMPLE
ON UPDATE NO ACTION ON DELETE NO ACTION,
CONSTRAINT tabla_mvic_motor_fk_bitinicio FOREIGN KEY (idbitinicio)
REFERENCES tabla_mvic_bitsestado (id) MATCH SIMPLE
ON UPDATE NO ACTION ON DELETE NO ACTION
);
    
```

5.3.6 Tabla de Balanza

```
CREATE TABLE tabla_mvic_balanza (  
    id bigint NOT NULL DEFAULT nextval('tabla_mvic_balanza_id_seq'::regclass),  
    tipofuncionamiento text,  
    horainicio time without time zone,  
    horafin time without time zone,  
    fechainicio date,  
    fechafin date,  
    idbitinicio integer,  
    idbitfin integer,  
    totalfuncionamientosegundos integer,  
    activo smallint,  
    CONSTRAINT tabla_mvic_motor_fk_bitfin FOREIGN KEY (idbitfin)  
        REFERENCES tabla_mvic_bitsestado (id) MATCH SIMPLE  
        ON UPDATE NO ACTION ON DELETE NO ACTION,  
    CONSTRAINT tabla_mvic_motor_fk_bitinicio FOREIGN KEY (idbitinicio)  
        REFERENCES tabla_mvic_bitsestado (id) MATCH SIMPLE  
        ON UPDATE NO ACTION ON DELETE NO ACTION  
);
```

5.3.7 Tabla de Motor Detenido

```
CREATE TABLE tabla_mvic_motor_detenido (  
    id integer NOT NULL  
    DEFAULT nextval('tabla_mvic_motor_detenido_id_seq'::regclass),  
    fechainicio date,  
    horainicio time without time zone,  
    fechafin date,  
    horafin time without time zone,  
    tiempoaparada integer,  
    estado smallint DEFAULT 0,  
    idbitinicio integer,  
    idbitfin integer,  
    CONSTRAINT tabla_mvic_motor_detenido_pk PRIMARY KEY (id)  
);
```

5.3.8 Tabla de parada de Emergencia

```
CREATE TABLE tabla_mvic_paradasemergencia (
    id bigint NOT NULL DEFAULT nextval('tabla_mvic_paradasemergencia_id_seq'::regclass),
    tipofuncionamiento text,
    horainicio time without time zone,
    horafin time without time zone,
    fechainicio date,
    fechafin date,
    idbitinicio integer,
    idbitfin integer,
    totalfuncionamientosegundos integer,
    activo smallint,
    CONSTRAINT tabla_mvic_paradasemergencia_pk PRIMARY KEY (id),
    CONSTRAINT tabla_mvic_paradasemergencia_fk_bitfin FOREIGN KEY (idbitfin)
        REFERENCES tabla_mvic_bitsestado (id) MATCH SIMPLE
        ON UPDATE NO ACTION ON DELETE NO ACTION,
    CONSTRAINT tabla_mvic_paradasemergencia_fk_bitinicio FOREIGN KEY (idbitinicio)
        REFERENCES tabla_mvic_bitsestado (id) MATCH SIMPLE
        ON UPDATE NO ACTION ON DELETE NO ACTION
);
```

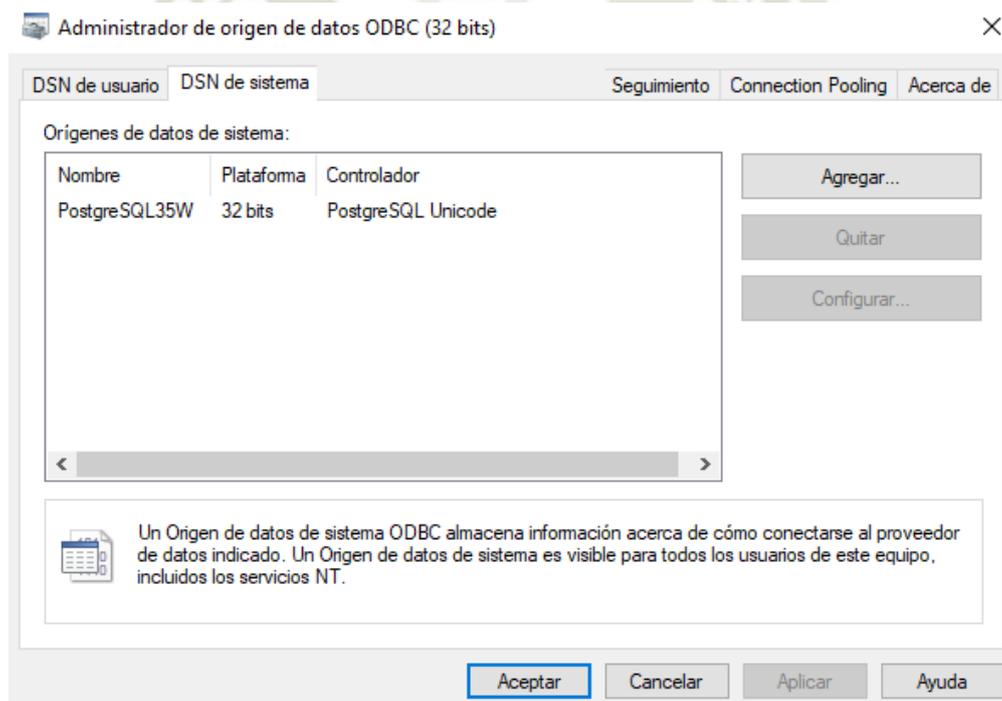
5.3.9 Tabla de Log

```
CREATE TABLE tabla_mvic_logs (
    id integer NOT NULL
    DEFAULT nextval('tabla_mvic_logs_id_seq'::regclass),
    tipo text,
    fuente text,
    descripcion text,
    fecha date,
    hora time without time zone
);
```

5.4 CONFIGURACIÓN EN SCADA

Con el fin de automatizar la recolección de datos, es necesario realizar una configuración específica sobre el software SCADA Intouch Wonderware. Esta configuración es posible debido a una aplicación propia del sistema que es el SQL Data Access, el cual se encarga de realizar la conexión a la base de datos a través de un controlador ODBC y posteriormente realizará las operaciones CRUD. El primer paso es instalar el controlador ODBC específico para la base de datos PostgreSQL. Dado que el sistema SCADA trabaja en base a una arquitectura de 32 bits es necesario que el ODBC que se configure sea también en la misma arquitectura, lo cual se puede lograr en el S.O. Windows a través de la ventana de configuración de orígenes de datos ODBC para 32 bits.

Imagen 22.- Configuración en 32 bit de ODBC



Con el controlador ODBC operativo, procederemos a la configuración de los llamados *Bind List* dentro de la aplicación del SCADA. Esta lista no es más que el mapeo respectivo entre el tag a ser almacenado y la columna correspondiente de la tabla en la base de datos. Cabe precisar que es necesario crear una bind list por cada tabla sobre la cual queramos.

Imagen 23.- Bind List para Tabla Totalizadores

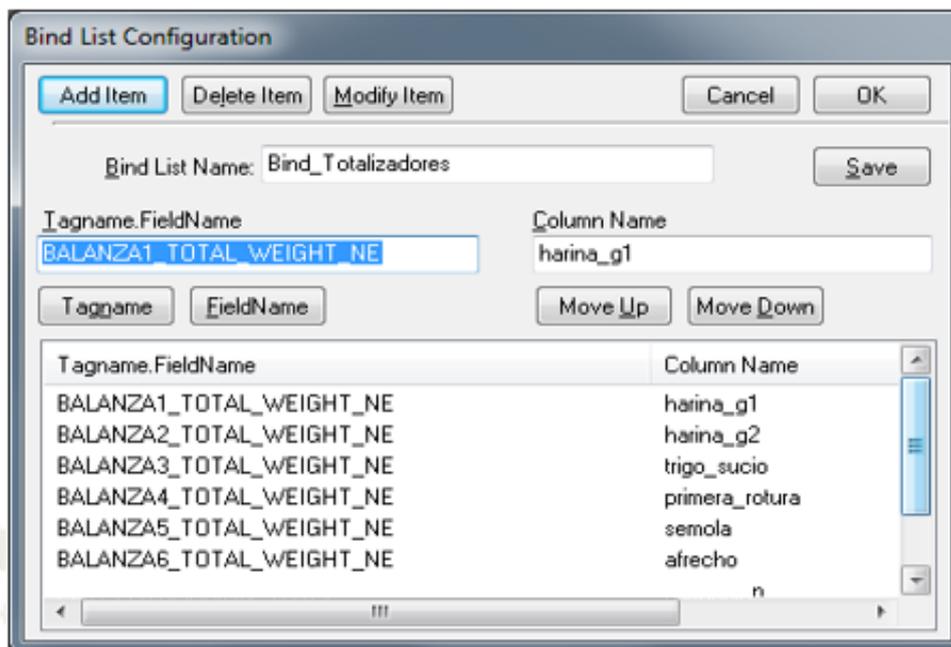
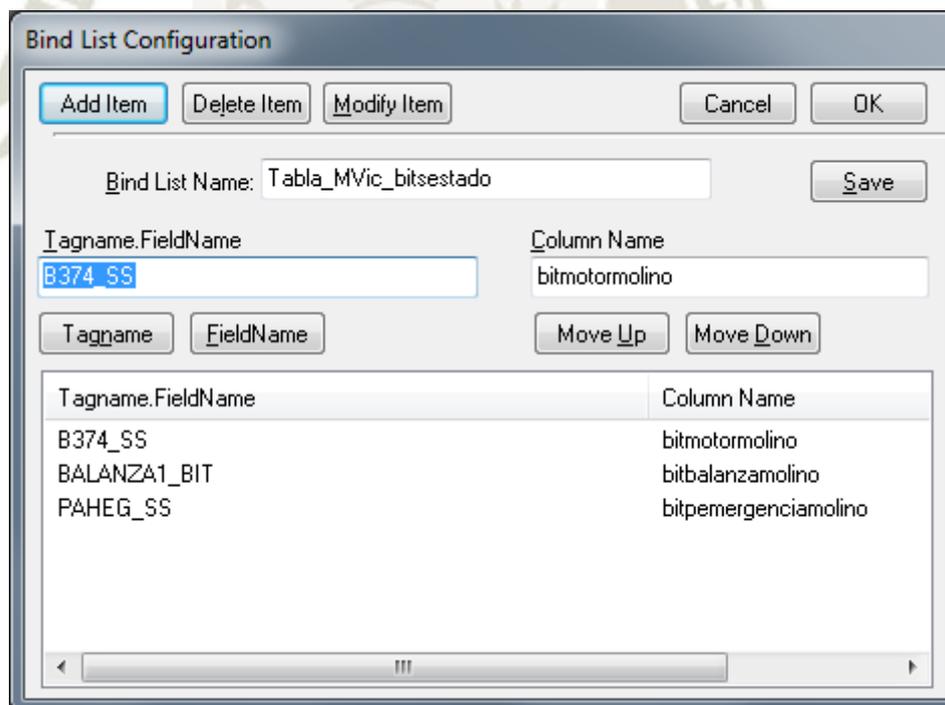


Imagen 24.- Bind List para Tabla de bits de Estado



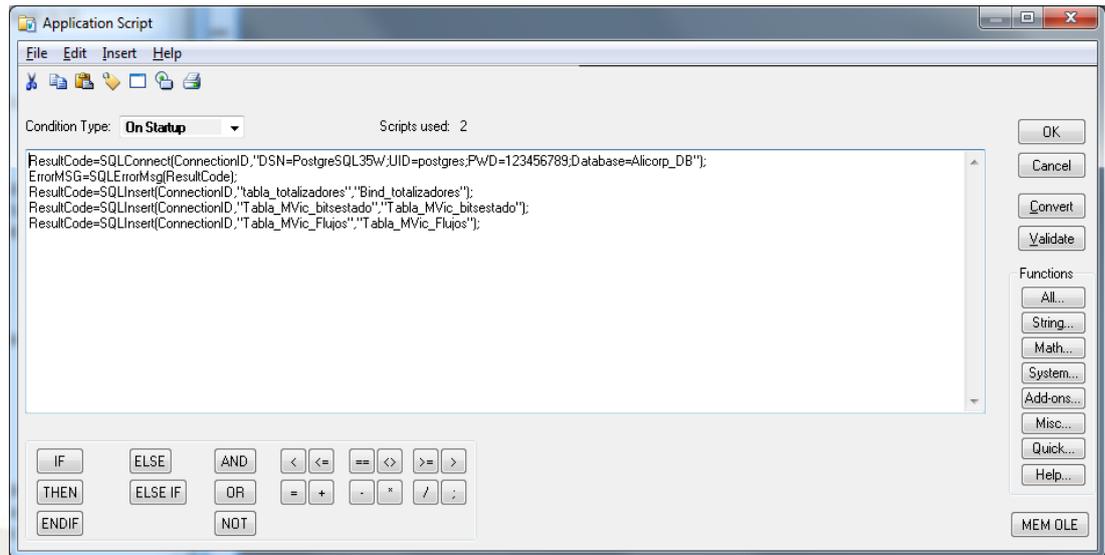
Una vez se ha logrado realizar el mapeo de todos los valores que deseamos extraer, es necesario programar cuándo o cómo lo hará el sistema SCADA, haciendo uso del driver y conexión ODBC. Para el fin del cómo se dispone de sentencias SQL que pueden ser ejecutadas en unos scripts dentro de la aplicación; Existen una variedad de sentencias que pueden ser ejecutadas como la lectura de tablas, creación de tablas, u otras siendo para el

interés de este proyecto la sentencia de crear la conexión insertar registros *SQLInsert*; esta sentencia recibe normalmente 3 parámetros que son: el ID de conexión, nombre de la tabla objetivo en la base de datos y por último el nombre de la Bind List creada que contiene el mapeo de tags contra columnas de la base de datos, que tienen que corresponder a la tabla enviada como parámetro. Esta línea permitirá ingresar los datos, lo siguiente es definir el cuándo.

La aplicación permite establecer criterios o condiciones para la ejecución de scripts, siendo de nuestro interés tres específicamente:

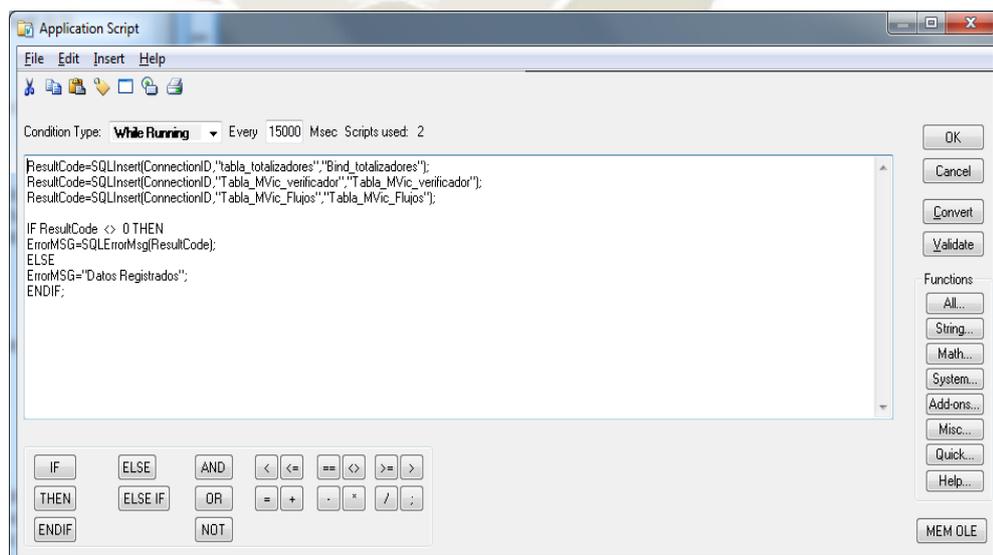
- **On Startup.-** Al iniciar el software SCADA, de manera que cada vez que la aplicación se inicie se ejecutará el script respectivo. Es imperativo que en este script se especifique la línea que realice la conexión a la Base de datos a través de la sentencia *SQLConnect* que recibe dos parámetros, el primero es el ID de la conexión y el segundo es una cadena de texto que contiene cuatro argumentos que son:
 - DSN (Data Source Name).- Que será el nombre de la conexión configurada en los orígenes de datos ODBC.
 - UID (User ID).- El nombre de usuario con privilegios de lectura y escritura sobre la base de datos a utilizar.
 - PWD (Password).- La contraseña respectiva al usuario.
 - Database.- El nombre de la base de datos que contiene las tablas que se intentaran manipular.

Imagen 25.- Condición de Scada: On Startup



- **While Running.-** La segunda condición es mientras la aplicación se ejecuta; esta opción permite efectuar las tareas programadas mientras se ejecute la aplicación en un criterio de tiempo, efectuando un bucle de lo programado en base a una cantidad definida en milisegundos. En este proyecto se definió cada 15000 Msec, es decir, cada 15 segundos.

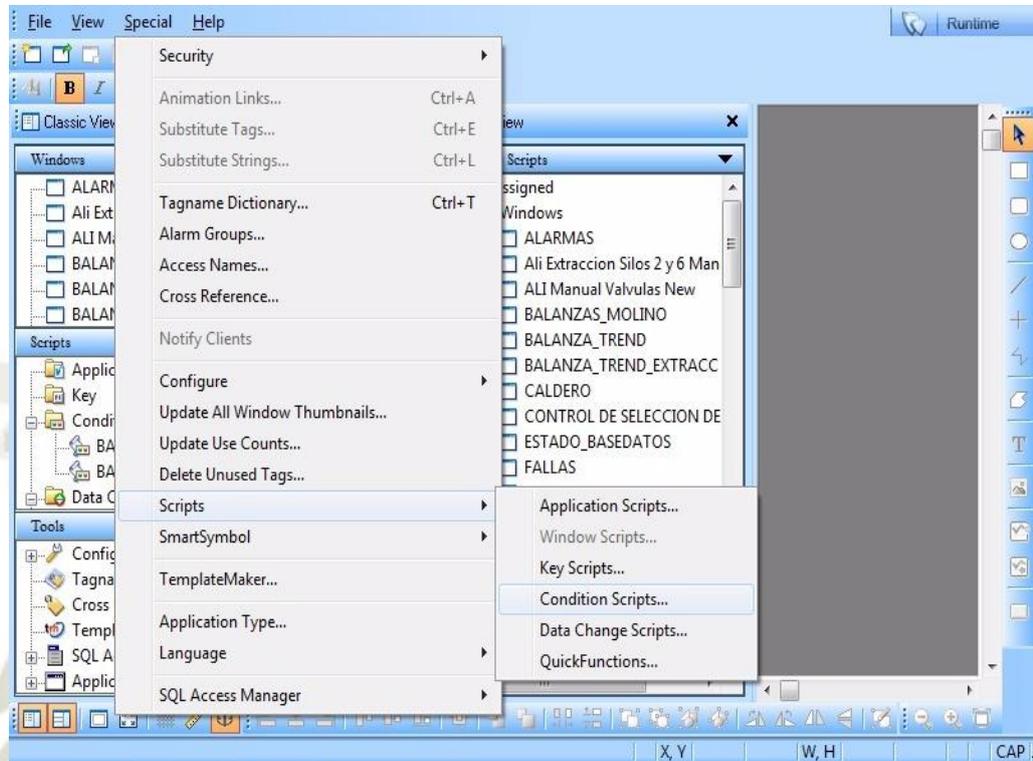
Imagen 26.- Condición de Scada: While Running



- **On condition.-** Esta opción es muy importante, ya que de este script dependerá cuándo se recibirá en base de datos el cambio del bit de estado de lo que queremos monitorear. La manera de definir estas condiciones es

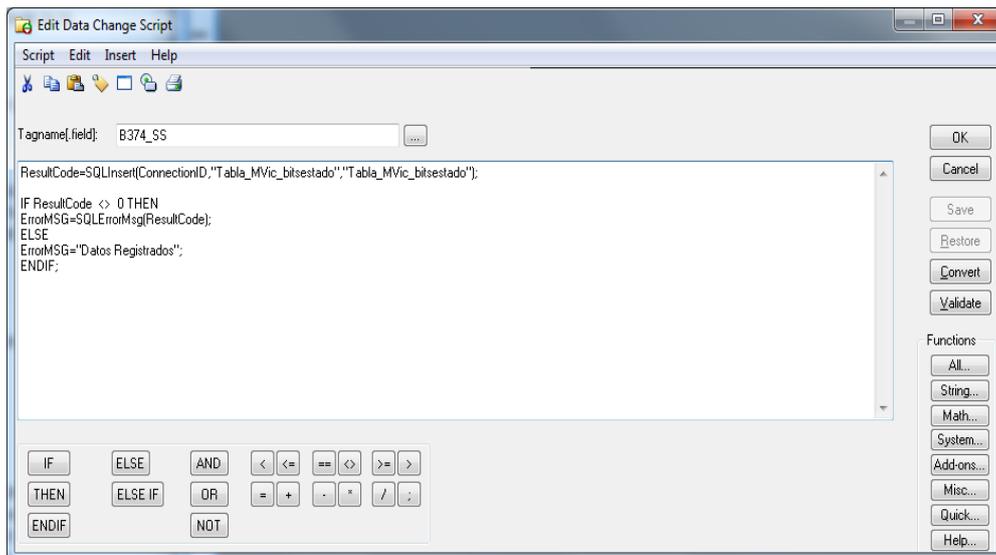
distinta a las dos mostradas anteriormente. Para esta tarea ubicaremos la opción de Scripts de condición o de cambio de data dentro del menú especial.

Imagen 27.- Scripts de condición



Para capturar el cambio del funcionamiento del motor de transmisión 1 de encendido a apagado, debemos configurar el monitoreo del tag correspondiente a dicho motor. Para esto añadiremos un script de cambio de dato y en el tag a seguir se seleccionará el tag correspondiente al Motor. Entonces cuando la aplicación detecte un cambio, ejecutará el script implementado.

Imagen 28.- Script de cambio de dato



Para el caso de la balanza se implementará un script de condición, debido a que no se dispone de un tag directamente relacionado con el encendido pero sí uno que mide el flujo. En este tipo de scripts se debe colocar la condición a evaluar y el comportamiento de ser verdadero o falso. En este escenario se entenderá un flujo de cero como una balanza inactiva, y un flujo mayor a cero como una balanza activa.

Imagen 29.- Script de condición: Balanza Flujo 0

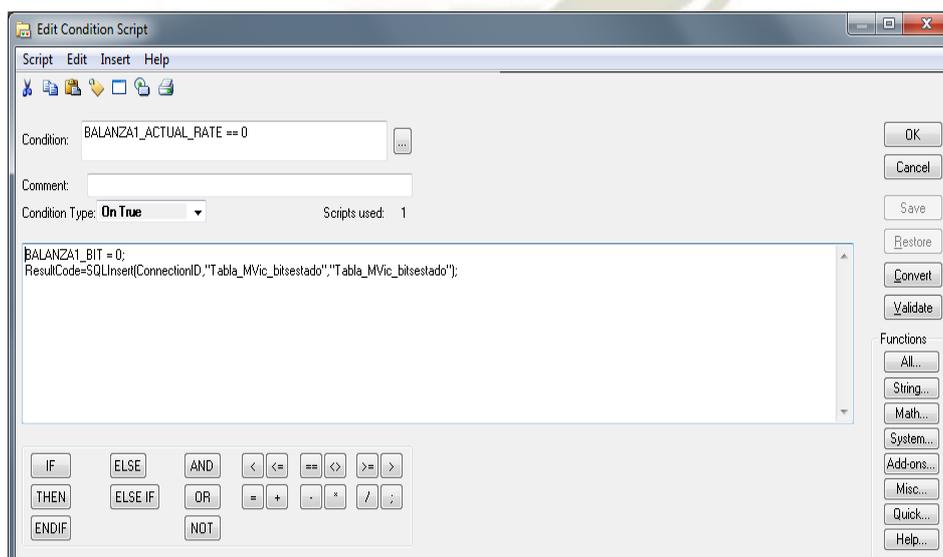
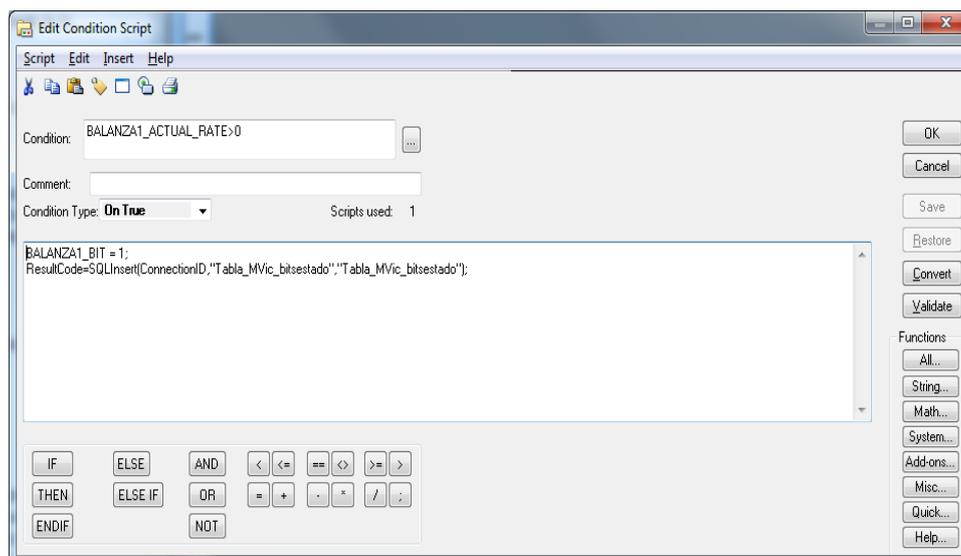


Imagen 30.- Script de condición: Balanza Flujo mayor a 0



Finalizando estos scripts, ya se ha configurado todo el comportamiento necesario dentro de la aplicación para el envío de datos a la base de datos.

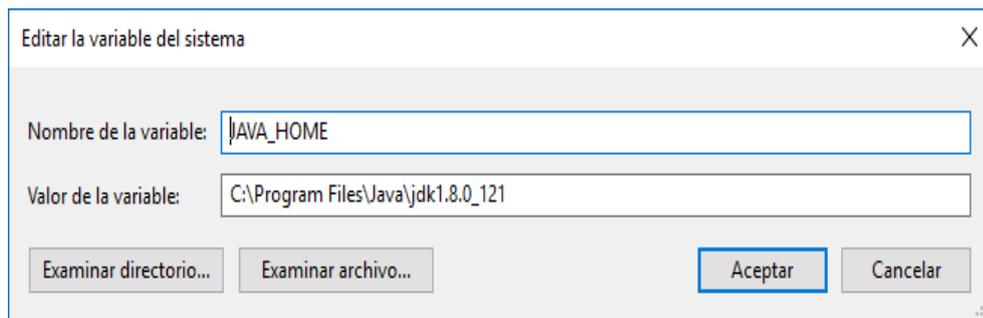
5.5 MOTOR DE APLICACIONES TOMCAT

Para el correcto funcionamiento del software de aplicaciones Tomcat, es necesario instalar los componentes de software JDK y JRE de JAVA; además se requiere configurar las variables de entorno de sistema pertinentes a estas soluciones para el correcto funcionamiento del servidor web Tomcat. Primero es comprobar si ya se tiene establecida la variable de entorno en el sistema, para lo cual se puede abrir una ventana de comando y ejecutar lo siguiente:

```
javac -version
```

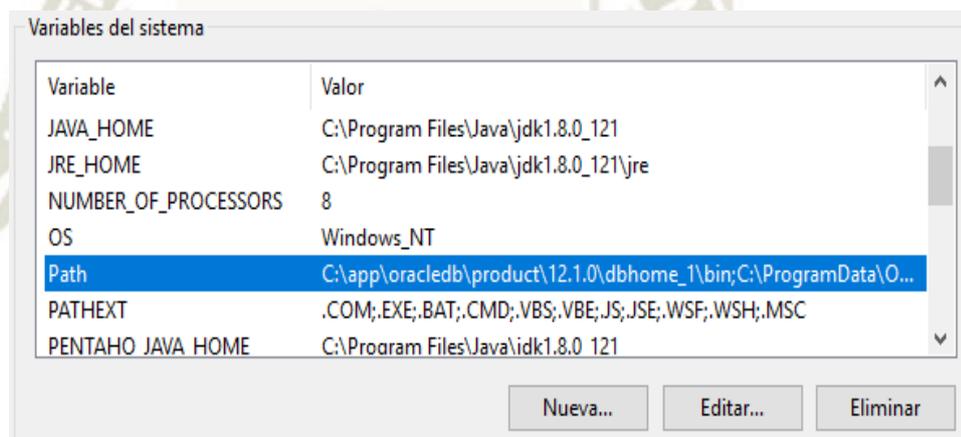
Si ya se tiene configurada la variables de entorno se obtendrá una respuesta con el texto *javac ##* donde *##* es el número de la versión de la máquina virtual de java que se ha instalado. En caso no se tenga configurada la variables el sistema responderá con un mensaje indicando que se desconoce el comando *javac*; de ser este el caso hay que configurar la variable primero creando una nueva variable de entorno del sistema llamada *JAVA_HOME* y en el valor se indicará la ruta donde se instaló la herramienta JDK.

Imagen 31.- Creación de variable Java Home



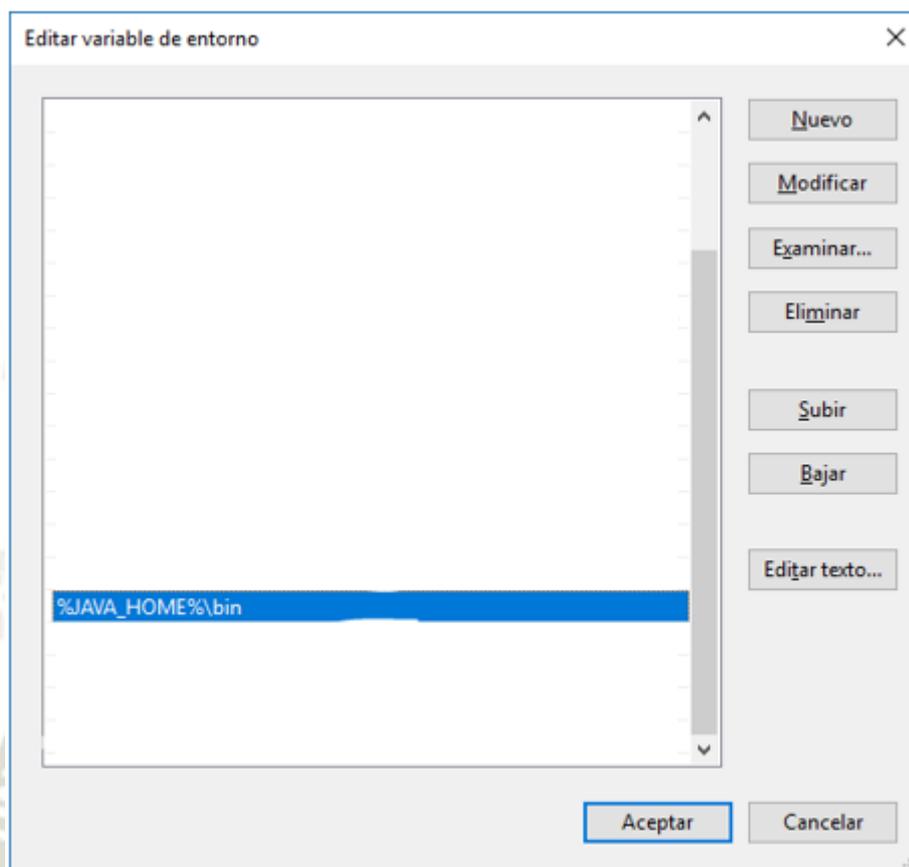
Posteriormente hay que añadir una línea a la variable de sistema “Path” agregando la variable de Java_Home, en específico a la carpeta *bin* dentro de la ruta de instalación de Java.

Imagen 32.- Variable de Sistema PATH



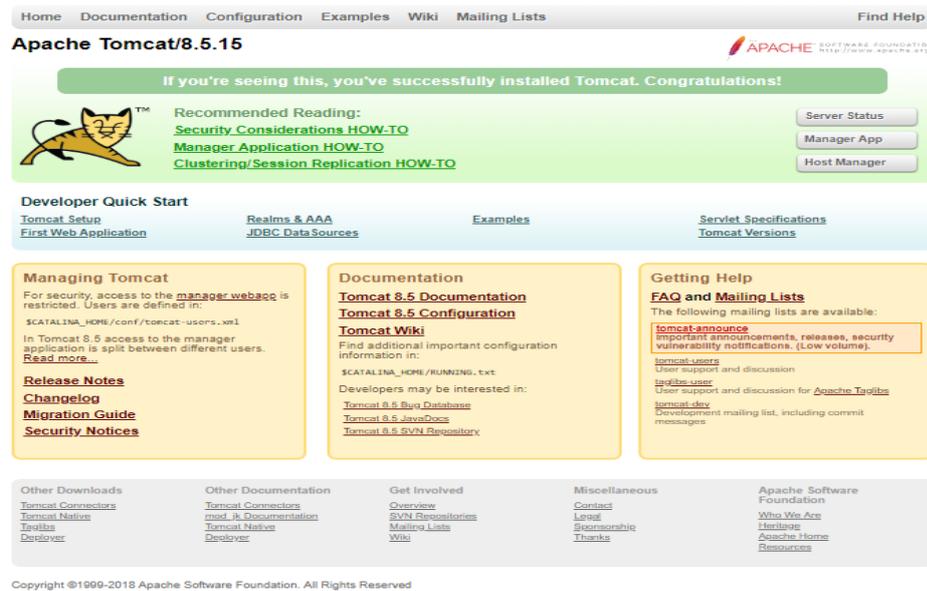
Al finalizar esta configuración se debe probar nuevamente ejecutar el comando anteriormente indicado, obteniendo el resultado mencionado así corroborando que la configuración ha sido la adecuada.

Imagen 33.- Entrada para Java Home dentro de Path



Una vez instalados los aplicativos pertenecientes a java, se puede descomprimir el paquete de Tomcat en el directorio de preferencia y está listo para su uso.

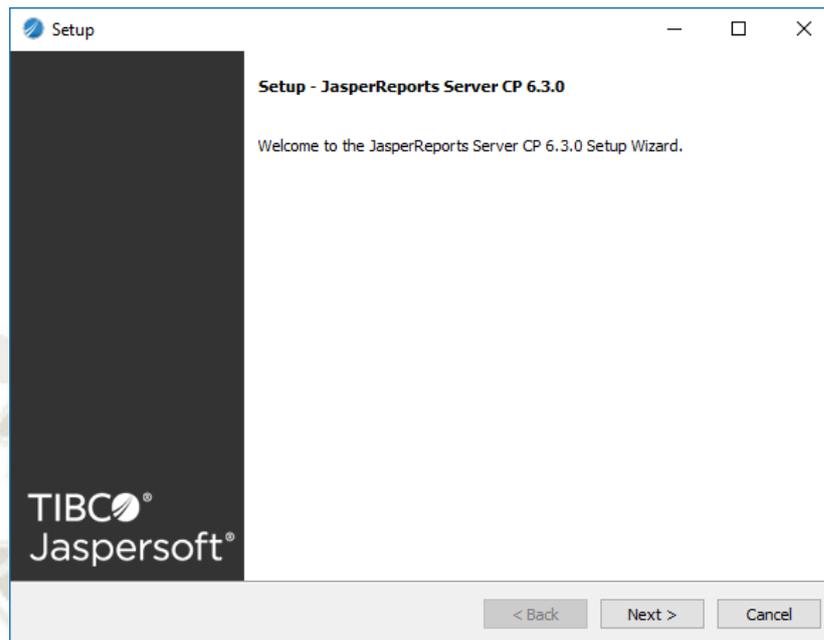
Imagen 34.- Pantalla de Inicio de Tomcat



5.6 SERVIDOR JASPER REPORT SERVER

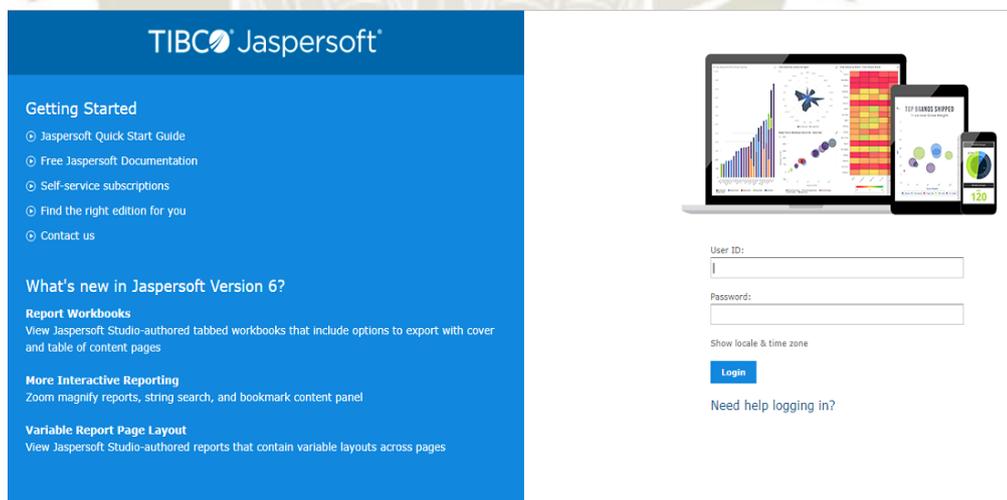
El servidor de reportes de JasperSoft tiene la facilidad de contar con un instalador para Windows, el cual ofrece una guía paso a paso para la instalación del servidor. Tras descargar el instalador de la versión de la comunidad 6.3, se ejecuta el mismo con privilegios administrativos, obteniendo una pantalla de bienvenida en la cual solo presionaremos el botón de siguiente, siguiendo cada paso para las respectivas configuraciones como lo son el directorio de la instalación, la selección del servidor Tomcat instalado en el equipo especificando la configuración y por último la selección del motor de base de datos:

Imagen 35.- Wizard de instalación para Jaspersoft Report Server



Finalizada la instalación se puede acceder al servidor de reportes haciendo uso del servidor Tomcat.

Imagen 36.- Pantalla de Inicio de Jasper Report Server



Ingresamos con las credenciales administrativas. Posteriormente se configurará la fuente de datos, es decir, la conexión a la base de datos desde la cual se extraerán los datos para los reportes; y para tal fin se usa una interfaz de configuración propia de la herramienta, amigable e intuitiva.

Imagen 37.- Configuración Fuente de Datos en Jasper Server

Finalizado el ingreso de los campos requeridos como la ruta o dirección del host de la base de datos, puerto, el nombre de la base de datos y las respectivas credenciales, se procede a guardar la configuración y se creará una entrada para la fuente de datos, lista para ser explotada.

Imagen 38.- Fuente de datos Creada

Name	Description	Type	Created Date	Modified Date
PostgreSQL		JDBC Data Source	12/22/2017	February 6

Es importante precisar que para la versión de comunidad 6.3 de Jasper Report Server existe un error que no permite visualizar los reportes, cuando se hace uso del idioma

español o francés. Este error se puede solucionar cambiando la línea de código correspondiente a dichos idiomas. Para tal fin modificamos dentro del archivo *jasperserver_messages_es.properties* la siguiente línea de texto:

De:

```
jasper.report.view.data.snapshot.message=Datos actualizados el {0,date} a las {0,time}
```

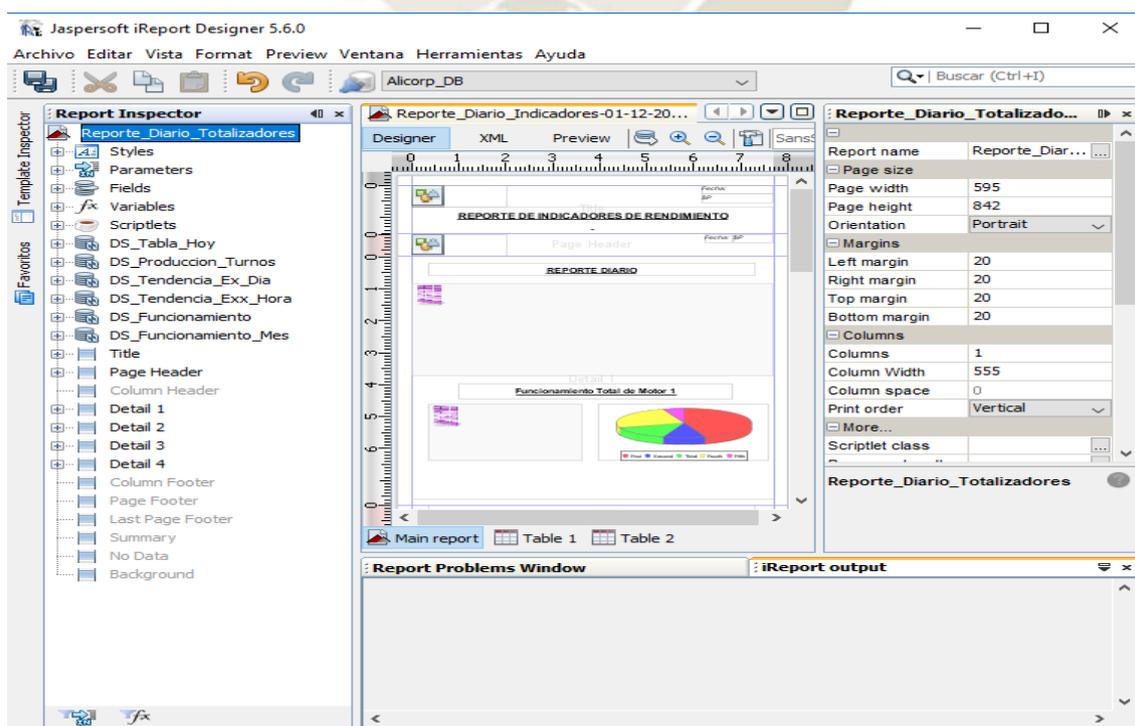
A:

```
jasper.report.view.data.snapshot.message=Datos actualizados el {0} a las {1}
```

5.7 IMPLEMENTACIÓN DE REPORTES

Para la implementación de los reportes buscados utilizamos la herramienta iReport versión 5.6 la cual se ofrece de manera gratuita en la página de jaspersoft. Este software es el ideal para diseñar los reportes que se ejecutarán en el servidor de reportes Jasper Report Server, obteniendo un archivo de extensión *.jrxml* que luego será compilado por el motor de reportes generando el documento deseado.

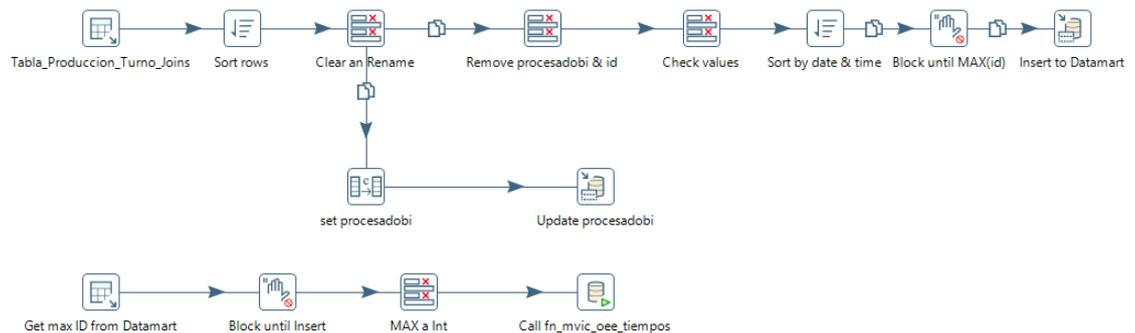
Imagen 39.- Interfaz de herramienta iReport



5.8 PROCESO DE INTEGRACIÓN DE DATOS

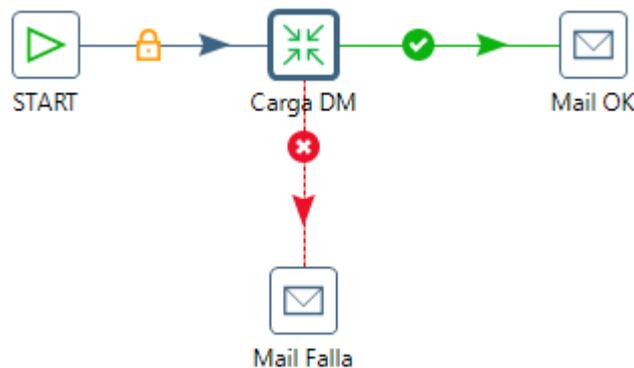
El siguiente paso para llegar al análisis de datos es poder extraer los datos recolectados y transformarlos de manera que puedan ser consultados efectivamente, resultando en información útil y analizable.

Imagen 40.- Proceso ETL dentro de Transformación



Fuente: Elaboración propia

Imagen 41.- Trabajo de ejecución de ETL

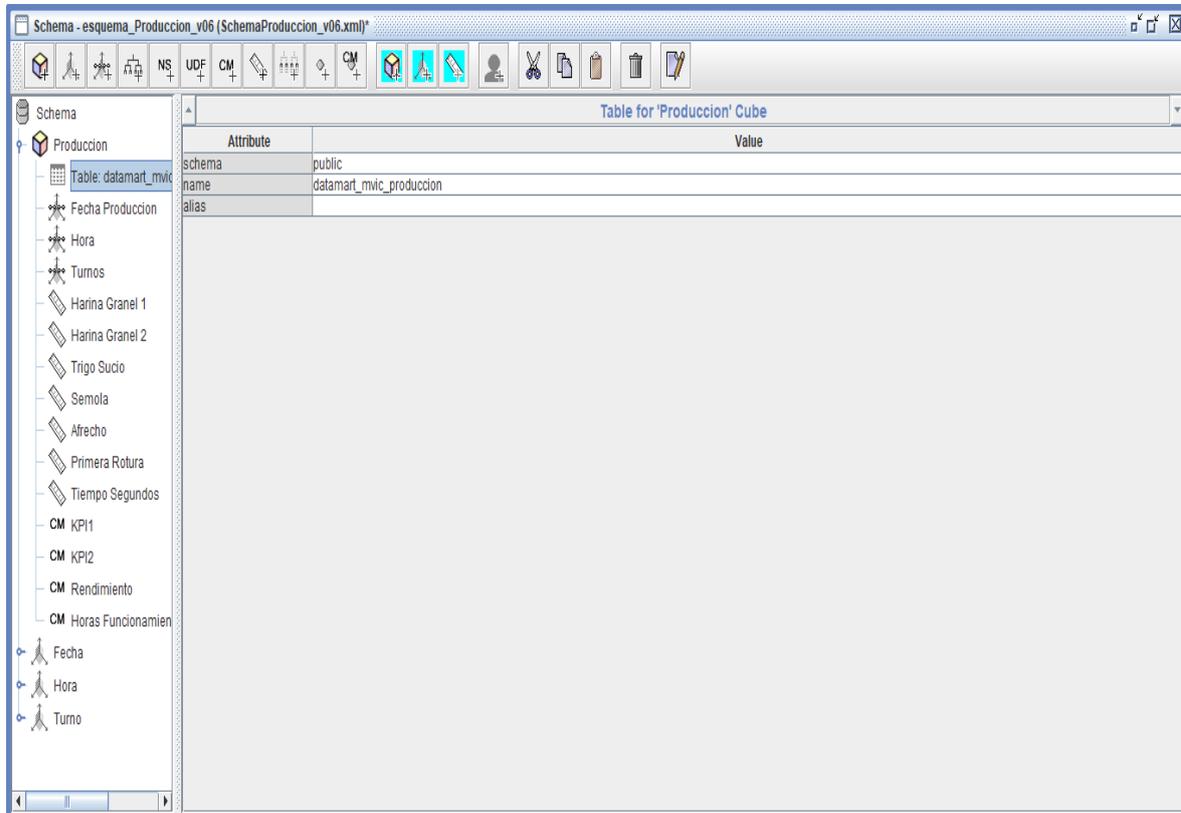


Fuente: Elaboración propia

5.9 IMPLEMENTACIÓN DE ESQUEMA

Dentro de la implementación del esquema se tiene las consideraciones definidas en la etapa de diseño. Para el desarrollo y publicación de este esquema se usa la herramienta Pentaho Schema Workbench, que dispone de una interfaz gráfica amigable, a la vez que la configuración directamente en XML.

Imagen 42.- Esquema de Producción en Pentaho Schema Workbench



5.10 SERVIDOR PENTAHO BUSINESS ANALITYCS

La implementación del servidor de pentaho deberá realizarse de forma manual para que trabaje en conjunto con el servidor de aplicaciones Tomcat y la base de datos ya existente PostgreSQL.

Para lograr dicha implementación se indicará paso a paso las acciones tomadas:

5.10.1 Preparación del entorno de instalación

Es indispensable asegurarnos que el entorno de instalación está listo. Debemos comprobar que tenemos los componentes de Java instalados, es decir el paquete JDK y JRE; adicionalmente validar que se tiene la variable de entorno de java configurada.

5.10.2 Instalación

Para instalar el servidor Pentaho haciendo uso del motor de base de datos PostgreSQL que se encuentra en funcionamiento, ejecutamos los siguientes scripts que crearán diversas bases de datos para el funcionamiento deseado.

5.10.2.1 Base de datos Hibernate

```
drop database if exists hibernate;  
drop user if exists hibuser;  
CREATE USER hibuser PASSWORD 'password';  
CREATE DATABASE hibernate WITH OWNER = hibuser ENCODING =  
'UTF8' TABLESPACE = pg_default;  
GRANT ALL PRIVILEGES ON DATABASE hibernate to hibuser;
```

5.10.2.2 Base de datos Jackrabbit

```
drop database if exists jackrabbit;  
drop user if exists jcr_user;  
CREATE USER jcr_user PASSWORD 'password';  
  
CREATE DATABASE jackrabbit WITH OWNER = jcr_user ENCODING =  
'UTF8' TABLESPACE = pg_default;  
  
GRANT ALL PRIVILEGES ON DATABASE jackrabbit to jcr_user;
```

5.10.2.3 Base de datos Quartz

```
drop database if exists quartz;  
drop user if exists pentaho_user;  
  
CREATE USER pentaho_user PASSWORD 'password';  
CREATE DATABASE quartz WITH OWNER = pentaho_user ENCODING  
= 'UTF8' TABLESPACE = pg_default;  
  
GRANT ALL ON DATABASE quartz to pentaho_user;  
  
\connect quartz pentaho_user
```

```

begin;
drop table if exists qrtz5_job_listeners;
drop table if exists qrtz5_trigger_listeners;
drop table if exists qrtz5_fired_triggers;
drop table if exists qrtz5_paused_trigger_grps;
drop table if exists qrtz5_scheduler_state;
drop table if exists qrtz5_locks;
drop table if exists qrtz5_simple_triggers;
drop table if exists qrtz5_cron_triggers;
drop table if exists qrtz5_blob_triggers;
drop table if exists qrtz5_triggers;
drop table if exists qrtz5_job_details;
drop table if exists qrtz5_calendars;
drop table if exists "QRTZ";

CREATE TABLE "QRTZ"
(
NAME VARCHAR(200) NOT NULL,
PRIMARY KEY (NAME)
);

ALTER TABLE "QRTZ" OWNER TO pentaho_user;

CREATE TABLE qrtz5_job_details (
JOB_NAME VARCHAR(200) NOT NULL,
JOB_GROUP VARCHAR(200) NOT NULL,
DESCRIPTION VARCHAR(250) NULL,
JOB_CLASS_NAME VARCHAR(250) NOT NULL,
IS_DURABLE BOOL NOT NULL,
IS_VOLATILE BOOL NOT NULL,
IS_STATEFUL BOOL NOT NULL,
REQUESTS_RECOVERY BOOL NOT NULL,
JOB_DATA BYTEA NULL,
PRIMARY KEY (JOB_NAME, JOB_GROUP)
);
    
```

```
CREATE TABLE qrtz5_job_listeners (
    JOB_NAME VARCHAR(200) NOT NULL,
    JOB_GROUP VARCHAR(200) NOT NULL,
    JOB_LISTENER VARCHAR(200) NOT NULL,
    PRIMARY KEY (JOB_NAME,JOB_GROUP,JOB_LISTENER),
    FOREIGN KEY (JOB_NAME,JOB_GROUP)
        REFERENCES qrtz5_JOB_DETAILS(JOB_NAME,JOB_GROUP)
);
```

```
CREATE TABLE qrtz5_triggers (
    TRIGGER_NAME VARCHAR(200) NOT NULL,
    TRIGGER_GROUP VARCHAR(200) NOT NULL,
    JOB_NAME VARCHAR(200) NOT NULL,
    JOB_GROUP VARCHAR(200) NOT NULL,
    IS_VOLATILE BOOL NOT NULL,
    DESCRIPTION VARCHAR(250) NULL,
    NEXT_FIRE_TIME BIGINT NULL,
    PREV_FIRE_TIME BIGINT NULL,
    PRIORITY INTEGER NULL,
    TRIGGER_STATE VARCHAR(16) NOT NULL,
    TRIGGER_TYPE VARCHAR(8) NOT NULL,
    START_TIME BIGINT NOT NULL,
    END_TIME BIGINT NULL,
    CALENDAR_NAME VARCHAR(200) NULL,
    MISFIRE_INSTR SMALLINT NULL,
    JOB_DATA BYTEA NULL,
    PRIMARY KEY (TRIGGER_NAME,TRIGGER_GROUP),
    FOREIGN KEY (JOB_NAME,JOB_GROUP)
        REFERENCES qrtz5_JOB_DETAILS(JOB_NAME,JOB_GROUP)
);
```

```
CREATE TABLE qrtz5_simple_triggers (
    TRIGGER_NAME VARCHAR(200) NOT NULL,
    TRIGGER_GROUP VARCHAR(200) NOT NULL,
    REPEAT_COUNT BIGINT NOT NULL,
    REPEAT_INTERVAL BIGINT NOT NULL,
```

```
TIMES_TRIGGERED BIGINT NOT NULL,
PRIMARY KEY (TRIGGER_NAME,TRIGGER_GROUP),
FOREIGN KEY (TRIGGER_NAME,TRIGGER_GROUP)
```

REFERENCES

```
qrtz5_TRIGGERS(TRIGGER_NAME,TRIGGER_GROUP)
);
```

```
CREATE TABLE qrtz5_cron_triggers (
```

```
TRIGGER_NAME VARCHAR(200) NOT NULL,
TRIGGER_GROUP VARCHAR(200) NOT NULL,
CRON_EXPRESSION VARCHAR(120) NOT NULL,
TIME_ZONE_ID VARCHAR(80),
PRIMARY KEY (TRIGGER_NAME,TRIGGER_GROUP),
FOREIGN KEY (TRIGGER_NAME,TRIGGER_GROUP)
```

REFERENCES

```
qrtz5_TRIGGERS(TRIGGER_NAME,TRIGGER_GROUP)
);
```

```
CREATE TABLE qrtz5_blob_triggers (
```

```
TRIGGER_NAME VARCHAR(200) NOT NULL,
TRIGGER_GROUP VARCHAR(200) NOT NULL,
BLOB_DATA BYTEA NULL,
PRIMARY KEY (TRIGGER_NAME,TRIGGER_GROUP),
FOREIGN KEY (TRIGGER_NAME,TRIGGER_GROUP)
```

REFERENCES

```
qrtz5_TRIGGERS(TRIGGER_NAME,TRIGGER_GROUP)
);
```

```
CREATE TABLE qrtz5_trigger_listeners (
```

```
TRIGGER_NAME VARCHAR(200) NOT NULL,
TRIGGER_GROUP VARCHAR(200) NOT NULL,
TRIGGER_LISTENER VARCHAR(200) NOT NULL,
PRIMARY
(TRIGGER_NAME,TRIGGER_GROUP,TRIGGER_LISTENER),
FOREIGN KEY (TRIGGER_NAME,TRIGGER_GROUP)
```

KEY

REFERENCES

```
qrtz5_TRIGGERS(TRIGGER_NAME,TRIGGER_GROUP)
);
```

```
CREATE TABLE qrtz5_calendars (
    CALENDAR_NAME VARCHAR(200) NOT NULL,
    CALENDAR_BYTEA NOT NULL,
    PRIMARY KEY (CALENDAR_NAME)
);
```

```
CREATE TABLE qrtz5_paused_trigger_grps (
    TRIGGER_GROUP VARCHAR(200) NOT NULL,
    PRIMARY KEY (TRIGGER_GROUP)
);
```

```
CREATE TABLE qrtz5_fired_triggers (
    ENTRY_ID VARCHAR(95) NOT NULL,
    TRIGGER_NAME VARCHAR(200) NOT NULL,
    TRIGGER_GROUP VARCHAR(200) NOT NULL,
    IS_VOLATILE BOOL NOT NULL,
    INSTANCE_NAME VARCHAR(200) NOT NULL,
    FIRED_TIME BIGINT NOT NULL,
    PRIORITY INTEGER NOT NULL,
    STATE VARCHAR(16) NOT NULL,
    JOB_NAME VARCHAR(200) NULL,
    JOB_GROUP VARCHAR(200) NULL,
    IS_STATEFUL BOOL NULL,
    REQUESTS_RECOVERY BOOL NULL,
    PRIMARY KEY (ENTRY_ID)
);
```

```
CREATE TABLE qrtz5_scheduler_state (
    INSTANCE_NAME VARCHAR(200) NOT NULL,
    LAST_CHECKIN_TIME BIGINT NOT NULL,
    CHECKIN_INTERVAL BIGINT NOT NULL,
    PRIMARY KEY (INSTANCE_NAME)
```

```
);

CREATE TABLE qrtz5_locks (
    LOCK_NAME VARCHAR(40) NOT NULL,
    PRIMARY KEY (LOCK_NAME)
);

INSERT INTO qrtz5_locks values('TRIGGER_ACCESS');
INSERT INTO qrtz5_locks values('JOB_ACCESS');
INSERT INTO qrtz5_locks values('CALENDAR_ACCESS');
INSERT INTO qrtz5_locks values('STATE_ACCESS');
INSERT INTO qrtz5_locks values('MISFIRE_ACCESS');

ALTER TABLE qrtz5_job_listeners OWNER TO pentaho_user;
ALTER TABLE qrtz5_trigger_listeners OWNER TO pentaho_user;
ALTER TABLE qrtz5_fired_triggers OWNER TO pentaho_user;
ALTER TABLE qrtz5_paused_trigger_grps OWNER TO pentaho_user;
ALTER TABLE qrtz5_scheduler_state OWNER TO pentaho_user;
ALTER TABLE qrtz5_locks OWNER TO pentaho_user;
ALTER TABLE qrtz5_simple_triggers OWNER TO pentaho_user;
ALTER TABLE qrtz5_cron_triggers OWNER TO pentaho_user;
ALTER TABLE qrtz5_blob_triggers OWNER TO pentaho_user;
ALTER TABLE qrtz5_triggers OWNER TO pentaho_user;
ALTER TABLE qrtz5_job_details OWNER TO pentaho_user;
ALTER TABLE qrtz5_calendars OWNER TO pentaho_user;

commit;
```

Para la finalidad de la presentación de este proyecto de tesis, se presenta una contraseña básica (**password**) que debería cambiarse en un entorno real. Finalizados los scripts, seguimos los siguientes pasos:

- Editamos el archivo `\webapps\pentaho\META-INF\context.xml` indicando las nuevas bases de datos en postgres a ser utilizadas por el sistema.

```
<?xml version="1.0" encoding="UTF-8"?>
<Context path="/pentaho" docbase="webapps/pentaho/">
```

```

<Resource          name="jdbc/Hibernate"          auth="Container"
type="javax.sql.DataSource"
  factory="org.apache.commons.dbcp.BasicDataSourceFactory"
maxTotal="20" maxIdle="5"
  maxWaitMillis="10000" username="hibuser" password="password"
  driverClassName="org.postgresql.Driver"
url="jdbc:postgresql://localhost:5432/hibernate"
  validationQuery="select version();" />

<Resource          name          =          "jdbc/Quartz"          auth="Container"
  type="javax.sql.DataSource"
  factory="org.apache.commons.dbcp.BasicDataSourceFactory"
maxTotal="20" maxIdle="5"
  maxWaitMillis="10000"          username="pentaho_user"
  password="password"
  driverClassName="org.postgresql.Driver"
url="jdbc:postgresql://localhost:5432/quartz"
  validationQuery="select 1"/>
</Context>

```

- Editamos el archivo \webapps\pentaho\WEB-INF\web.xml comentando las secciones enmarcadas entre los tags [BEGIN HSQLDB DATABASES] y [BEGIN HSQLDB STARTER].
- En el directorio pentaho-solutions editamos /hibernate/hibernate-settings.xml Cambiando el parámetro de configuración <config-file> a

```
<config-file>system/hibernate/postgresql.hibernate.cfg.xml</config-file>
```

- Bajo el mismo directorio editamos el siguiente archivo para indicar el uso de postgres.

```

jdbc.driver=org.postgresql.Driver
jdbc.url=jdbc:postgresql://localhost:5432/hibernate
jdbc.username=hibuser
jdbc.password=password
hibernate.dialect=org.hibernate.dialect.PostgreSQLDialect

```

- Para finalizar editaremos alguna líneas, para el correcto funcionamiento con el uso de la base de datos postgresQL. Dentro del siguiente archivo XML `\pentaho-solutions\system\jackrabbit\repository.xml` modificaremos.

- **Configuración del FileSystem:** descomentar de la línea 61 a la 68. Comentar de la línea 69 a la 83
- **Configuración del DataStore:** descomentar de la línea 123 a la 134. Comentar de la línea 135 a la 155
- **Configuración del FileSystem:** descomentar de la línea 236 a la 243. Comentar de la línea 244 a la 259
- **Configuración del PersistenceManager:** descomentar de la línea 293 a la 300. Comentar de la línea 301 a la 317
- **Configuración del FileSystem:** descomentar de la línea 386 a la 393. Comentar de la línea 394 a la 408
- **Configuración del PersistenceManager:** descomentar de la línea 443 a la 450. Comentar de la línea 451 a la 466

- Con estos pasos podemos considerar que nuestro servidor Pentaho está listo para ser usado.

CAPÍTULO VI

RESULTADOS

6.1 GENERALIDADES

Para la validación de resultados de este trabajo se hizo uso del juicio de expertos, solicitando el apoyo de los principales involucrados dentro del proyecto y que tienen el total conocimiento del proceso de producción del molino, siendo éstos:

- Gerente de Producción
- Jefe de Producción
- Supervisor de producción

Con su apoyo se pudo realizar una validación de los datos presentados en la solución implementada, ya que se mantuvo el proceso manual con el fin de corroborar la fidelidad y certeza de los datos que se muestran en el sistema automatizado. A partir del criterio de los expertos es que se llegó a las conclusiones presentadas en este capítulo, además del reconocimiento de patrones de comportamiento de los indicadores y su significado.

Cabe señalar que para el desarrollo de este trabajo de tesis, se sostuvo el uso de un instrumento de recolección de información al final del proyecto: la entrevista semi estructurada (ver anexo 2), la misma que se desarrolló con el procedimiento siguiente:

- Coordinación de reunión.
- Entrevista final.
- Elaboración de acta de culminación de proyecto.

Luego de este procedimiento se obtuvo los siguientes resultados a las preguntas planteadas:

Pregunta 1.- ¿La solución es de utilidad para el control de la producción? ¿Se percibe alguna mejora sobre el proceso?

Los expertos señalaron que la solución obtenida sí es de utilidad, tanto para el monitoreo y control de la producción en tiempo real, como para observar comportamientos y patrones dentro la producción asociados a diferentes criterios, que fueron establecidos en reuniones previas. En cuanto al proceso, se denota una mejora en la producción ya que se optimizaron sus tiempos de respuesta a eventos propios de la línea de producción que comprometían la eficiencia del molino, además que gradualmente van eliminando tareas manuales del supervisor. Por otro lado señalaron que la facilidad de obtener un reporte de producción en sus correos corporativos de manera automática les daba una herramienta de acción y planificación oportuna.

Pregunta 2.- ¿La solución implementada es confiable? ¿Representa valores reales sobre la producción?

A esta pregunta, se precisó que para obtener un diagnóstico de los resultados mostrados por la solución, se realizaron los procesos de toma de datos y cálculo de indicadores de manera manual por parte del supervisor con ayuda del jefe de producción. Los datos correspondieron a los mostrados por la solución, de esta manera validando que la solución representa datos confiables y veraces de la línea de producción sobre la cual opera.

Pregunta 3.- Acerca de los valores y gráficas que se representan en la interfaz, ¿es posible tomar decisiones con esos datos?

La solución implementada sí permite la toma de decisiones oportunas a los principales responsables, ya que las tendencias descritas por las gráficas representadas fueron asociadas a eventos o comportamientos dentro de la producción. En la tabla representada a continuación se describen tales comportamientos y su respectiva descripción.

Tabla 14.- Comportamiento de Indicadores

N°	Comportamiento		Situación
	KPI-1	KPI-2	Descripción
1	—	—	Funcionamiento normal del molino
2	—	↗	Reincorporación de subproducto
3	—	↘	Atoro de sub producto
4	↗	—	Se mejora regulación del molino
5	↗	↗	Reincorporación de harina y/o subproductos en proceso.
6	↗	↘	Atoro de sub producto y mejora la regulación del molino.
7	↘	—	Regulación del molino inadecuada.
8	↘	↗	Regulación del molino inadecuada e incorporación de sub productos.
9	↘	↘	Atoro de productos en proceso.

Fuente: Elaboración conjunta con personal de la empresa

Pregunta 4.- La plataforma de análisis de datos, ¿es de ayuda para un planeamiento estratégico o táctico a futuro?

Se precisó que sí es posible elaborar planes de acción futuros o corregir ciertos comportamientos al poder visualizar esa información, aunque al momento de culminación del proyecto no existían suficientes datos históricos. Con el paso del tiempo y al recopilar un mayor volumen del tiempo será posible replantear objetivos o métodos de trabajo.

Con esta salvedad, es que se presenta en el actual capítulo, una valorización, ya no física, sino monetaria. Para esto se considera los resultados obtenidos que son considerados mejoras, así como también los cálculos referenciales que la eficiencia ha permitido a la empresa en una forma de monetización.

6.2 PROCESAMIENTO DE DATA

Para poder considerar cada uno de estos puntos, se resume en la siguiente tabla las mejoras propuestas:

Tabla 15.- Comparativo de Sistema en dos tiempos

Actividad	Antes de la implementación del sistema	Después de la implementación del sistema
Ingreso de información	El ingreso era manual y estaba a cargo del supervisor, él mismo emplea un papel simple (sin formato) para recopilar información de cada balanza y esta pasaba al Parte de producción, con el cual se desarrollaban los procesos futuros.	El ingreso es automatizado, desde los PLC hacia una base de datos, siendo el supervisor el que debe estar pendiente de los datos que proporciona el sistema.
Tiempo de ingreso de información	Al final de cada turno de producción, se recopilaba la información manualmente, que considera un tiempo alrededor de 30 minutos por turno, lo que significa un total de 90 minutos por día, lo que hace al mes 2,700 minutos y al año 32,400 minutos.	El tiempo de recolección de datos es de 15 segundos, tiempo en el que el sistema se actualiza automáticamente. No existe participación de una persona para dicha recolección.
Procesamiento de información	El supervisor y el asistente de planta recopilaban los datos del Parte de Producción y se trasladaban a una hoja de cálculo, para poder generar sus reportes en el formato predeterminado de la misma hoja de cálculo. Revisaban la recolección histórica del día anterior para presentar el reporte del día.	El usuario observa la plantilla de reporte por cada turno en un tiempo menor, es decir, a disposición del tiempo del usuario (a demanda), con una serie de herramientas que le permiten apreciar los indicadores históricos, describiendo la eficiencia de la planta, sus volúmenes de producción y los volúmenes de desecho o pérdida de producción, todos estos por cada turno.
Tiempo de procesos de información	El tiempo, estaba bajo el condicionante de lo que deseaban hacer con la información, por lo que se da un menor escenario en 20 minutos y un máximo de 50 minutos. Esta actividad se ejecutaba una vez al día, entonces esta actividad mensual requería de 600 minutos a 1,500 minutos al mes o 7,200 minutos a 18,000 minutos al año.	El tiempo es similar al de la actualización de la data, que se da cada 15 segundos.
Resultados de información	El asistente de producción elaboraba un solo reporte de producción diario, con el uso de la hoja de cálculo, con un formato propio de la hoja de cálculo, se mostraba solo el dato del día y no presentaba indicadores o gráficas que permitieran un análisis más eficiente de la producción.	El usuario (gerente, jefe de planta, asistente de producción, supervisor y el jefe de mantenimiento) observa la plantilla de reporte por cada turno en un tiempo menor, es decir, a disposición del tiempo del usuario (a demanda), donde se aprecia las tendencias de los indicadores en el día de producción y en el mes. El reporte también se envía

		automáticamente por correo electrónico a los usuarios.
Tiempo de resultados de información	El tiempo que se daba en esta actividad estaba alrededor de los 30 minutos diarios, lo que al mes hace 900 minutos y al año hace 10,800 minutos	El tiempo es de 1 minuto y la actualización de la data se da cada 15 segundos de manera automática.

Fuente: Elaboración propia

Con lo plasmado en la tabla anterior, se presentan los resultados en los siguientes puntos

6.2.1 Administrativos

La parte administrativa, considera siempre la gestión que se pueda dar de los tres factores relevantes como son: recurso humano, recurso económico y recurso material, de lo cual se busca la eficiencia de estos, basados en las mejores tomas de decisión, las mismas que solo se pueden dar con la información adecuada y en forma oportuna. La mejora del proceso de producción de una empresa productora de harinas aplicando Inteligencia de Negocios con herramientas Open Source, ha ocasionado lo siguiente:

- Los usuarios que tienen acceso al sistema (gerente, jefe de producción, asistente de producción, supervisor y el jefe de mantenimiento), mantienen información diaria la misma que se actualiza cada 15 segundos, para poder determinar si se está dando el uso correcto de los recursos materiales y humanos en el proceso de producción.
- Los tiempos de horas hombre destinados se han reducido incluso hasta en más de un 98%, pasando de 90 minutos al día a 1 un minuto
- La información puede ser compartida con los demás usuarios, en forma inmediata, lo que ha mejorado los tiempos en un 98%.

6.2.2 Producción

Dentro de la producción se considera las mejoras en los siguientes puntos:

- Los procesos de producción son mejor controlados, por el tiempo de actualización de la información obtenida.

- Las relaciones entre los indicadores de “KPI1” y “KPI2”, permiten la ejecución de alguna actividad, o reconocer un comportamiento específico. Esto se comprende mejor en la tabla anteriormente mostrada. (Tabla N° 14).
- Los niveles de producción mejoraron en el KPI-1, pasando de un promedio de 78% a 84%, lo que significa un incremento del 7.7%.

6.2.3 Técnicos

Se dejó de lado la adquisición de licencias adicionales al SCADA para lograr este control, así como también el pago del especialista de desarrollo sobre la plataforma SCADA que sería el responsable de implementar estos ajustes. Este presupuesto podría ser hasta 4 veces mayor al que se ha implementado.

6.2.4 Económicos

Con los resultados plasmados en los puntos anteriores es que se consideran los cálculos referenciales para los resultados económicos en la empresa:

Tabla 16.- Referencial del cálculo

	Sin implementación	Con implementación	Diferencia	Valorización por producción (cantidad)	por Valorización producción (S/)
Horas hombre	4600 minutos al mes	Se considerará un 1 minuto por día, porque la recolección y procesamiento de datos fueron automatizados. Son 30 minutos al mes.	4570 minutos de ahorro al mes	Venta aproximada por mes 90,000 sacos de 50 kilos a 70 soles cada uno.	6,300,000.00 por mes
Resultado de ahorro					666,458.00

Incremento de producción			Aumento del 7.7%	Venta aproximada por mes 90,000 sacos de 50 kilos a 70 soles cada uno.	6,300,000.00 por mes
Resultado de aumento de producción					485,100.00 por mes

Fuente: Elaboración propia

6.3 PROYECCIONES

El desarrollo de las mejoras, al haber empleado Open Source, se considera que pueden darse actualizaciones y correcciones, si se consideran necesarias, como el caso del MySQL que actualizó al 2018, implementando nuevas funcionalidades que se pueden adaptar y hacerlo más eficiente el sistema. Se considera que los cambios implementados lograrán toda la eficiencia plasmada, dentro de un tiempo de 5 años posteriores, ya que con el SCADA han trabajado por 2 años sin el sistema. Esta es la conversación que se sostuvo con el jefe de molinos.

CONCLUSIONES

1. Tras haber implementado una solución completa, se logró mejorar el proceso de producción al existir un correcto y oportuno manejo de la información y de los datos de la producción. Esto se ve reflejado en el incremento del porcentaje del primer indicador, logrando una mejora de un aproximado de 7.7% dado que se pueden tomar decisiones de manera oportuna, sobre todo en caso de atoros.
2. Con la solución desarrollada, se realizó una comunicación directa entre el sistema SCADA y la base de datos PostgreSQL permitiendo extraer la información desde los tags correspondientes. Dado el caso, el supervisor ya no requiere realizar la toma de datos manualmente, ya que la recopilación de datos relacionados a la producción se realiza de manera automática a través de software.
3. Para la implementación del repositorio de datos se escogió el motor de base de datos PostgreSQL, el mismo que a la fecha de la implementación era el más adecuado de acuerdo a sus características y beneficios que brinda. Es necesario tomar en cuenta que dicho repositorio cumple las expectativas de desempeño y calidad para el tratamiento de datos, además que la base de datos fue desarrollada con estándares de calidad.
4. A través del uso de la plataforma de código abierto de Pentaho, se logró la implementación de una solución en la cual pueden realizarse consultas de los datos históricos de producción, aplicando diferentes tipos de filtros y cortes de manera dinámica. El uso de la solución de BI ha permitido alcanzar una mayor perspectiva del negocio y del proceso de producción al observar fácilmente las tendencias y comportamientos que anteriormente no era fácil de realizar.
5. Los reportes, diseñados sobre la plataforma Jaspersoft, lograron una absolución del tiempo empleado para generar éstos. Una vez programados sobre el servidor de reportes de Jaspersoft, los documentos son generados automáticamente con

una frecuencia diaria o mensual y enviados de manera automatizada a los principales responsables, para que ellos puedan establecer el plan de acción y asegurar la sostenibilidad de la producción de la empresa.

6. La visualización y monitoreo de los indicadores se dan por el desarrollo de una interfaz estilo dashboard en Java y Javascript, funcionando dentro del servidor Tomcat; facilitando la percepción de los indicadores en tiempo real. Este cuenta con el valor del indicador actual, y de horas previas, graficando tendencias de los indicadores para la toma de decisiones sobre la producción.



RECOMENDACIONES

1. Se recomienda considerar la implementación de soluciones adicionales que permitan remover la dependencia del sistema SCADA para la recolección de datos, al ser este el punto más crítico de funcionamiento del resto de aplicaciones. Para esta recomendación se pueden aplicar soluciones como un servidor OPC y un Data Logger.
2. Se recomienda la integración de más procesos que pertenecen al área de producción de Molinos, con el fin de disponer de mayor información y poder generar conocimiento de superior calidad, como por ejemplo la causa de parada de motor que interrumpe el funcionamiento de la línea de producción. Adicionalmente, mediante la integración de más áreas se puede obtener conocimiento de mayor impacto en el negocio.
3. Se sugiere evaluar la herramienta de servidor de base de datos MySQL, en su última versión, contra la versión usada en esta investigación, para poder determinar una posible migración o persistencia con la solución en funcionamiento.

BIBLIOGRAFÍA

- Asociación del Parque Industrial de Arequipa. (2018). <http://www.adepia.com.pe/>.
Obtenido de <http://www.adepia.com.pe/>: <http://www.adepia.com.pe/historia-2/>
- Bonér, J., Farley, D., Kuhn, R., & Thompson, M. (05 de 07 de 2018). *The Reactive Manifesto*. Obtenido de <https://www.reactivemanifesto.org/>
- Cabero Almenara, J., & Llorente Cejudo, M. (2013). La aplicación del Juicio de Experto como técnica de evaluación de las tecnologías de la información y comunicación (TIC). *Revista de Tecnología de Información y Comunicación en Educación*, 11-22.
- Carrión Albán, A. R., & Torres Sosa, M. C. (2016). *Análisis comparativo de herramientas open source para soluciones de inteligencia de negocios (BI) y su aplicación práctica para la toma decisiones en la empresa lácteos de Marco's*. Quito: Pontificia Universidad Católica del Ecuador.
- Córdova Ulloa, J. P. (2016). *Implementación de Inteligencia de Negocios en el Sistema Red Socio Empleo utilizando Pentaho Reporting*. Quito: Facultad de Ingeniería y Ciencias Agropecuarias. UDLA.
- Díaz Piraquive, F. (2008). Gestión de procesos de BPM, TICs y crecimiento empresarial. *Universidad & Empresa*, 151-176.
- Escobar Pérez, J., & Cuervo Martínez, Á. (2008). Validez de contenido y juicio de expertos: Una aproximación a su utilización. *Avances en Medición*, 27-36.
- Fernández González, J., & Mayol Sarroca, E. (04 de Febrero de 2012). *Dataprix*. Recuperado el 01 de Julio de 2018, de <http://www.dataprix.com/factores-criticos-exito-proyecto-business-intelligence>
- Flórez Fernández, H. A. (2012). Inteligencia de negocios como apoyo a la toma de decisiones en la gerencia. *Vínculos*, 11-23.
- Flórez Salgado, D. F., & Tejada Cardona, L. M. (2015). *Guía para la implementación de una solución de inteligencia de negocios para pequeñas y medianas empresas*. 2015: Universidad tecnológica de Pereira.
- Gangadharan, G. R., & Swami, S. (2004). Business Intelligence Systems: Design and Implementation Strategies. *26th International Conference on Information Technology Interfaces*, (págs. 139-144). Cavtat.

- Guevara Vega, C. P. (2015). *Desarrollo de una plataforma de business intelligence para facilitar el análisis de datos de las competencias generales de formación aplicadas en el desempeño laboral de los egresados de la Universidad Técnica del Norte*. Ibarra: Universidad de las Fuerzas Armadas ESPE.
- IBM. (2012). *A step-by-step approach to succesful Business Intelligence*.
- Ícaro Fernández, M. (2013). *Implantación de la metodología BPM en la EPS: Aplicación para la gestión de comisiones*. Universidad de Alicante, Alicante.
- Ícaro Fernández, M. (2013). *Implantación de la metodología BPM en la EPS: Aplicación para la Gestión de Comisiones*. Trabajo de Fin de Máster, Universidad de Alicante, Alicante.
- Khan, R. A., & Quadri, S. (2012). Business Intelligence: An integrated Approach. *Business Intelligence Journal*, 64-70.
- Lorenzo, O. (2003). *www.tendenciasdigitales.com*. Obtenido de Evolución y Tendencias en los Sistemas Empresariales. : www.tendenciasdigitales.com/td/articulo_crm.htm
- Luna Mancinas, G. (2014). *Aplicación de la metodología Seis Sigma para mejorar el proceso de acondicionamiento del grano de trigo*. Tesis, Universidad de Sonora, Sonora.
- Montalvo Jaramillo, P., Hernandez Cevallos, M. I., Ledesma Marcalla, D. A., & Santillán Gallegos, M. (2010). *Desarrollo de un Sistema Scada para la Medición de Voltajes con Sistemas Embebidos para el Laboratorio de Mecatrónica de la Facultad de Mecánica*. Chimborazo: Escuela superior politécnica de Chimborazo.
- Moss, L., & Atre, S. (2003). *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*. Addison Wesley.
- Olivares, J. P. (2011). *SOLUCION DE INTELIGENCIA DE NEGOCIOS PARA UNA PYME*. Valparaíso: Pontificia Universidad Católica de Valparaíso.
- Ramirez, P. (2004). Rol y contribución de los sistemas ERP's. En http://fondosdigitales.us.es/public_thesis/pdf. Valparaiso.
- Ranjan, J. (2009). Business Intelligence: Concepts, Components, Techniques and Benefits. *Journal of Theoretical and Applied Information Technology*, 60-70.
- Salazar Ramírez, J. (2016). *Implementación de una solución BPM para agilizar los procesos del área de abastecimiento en la Municipalidad de Chiclayo*. Tesis, Chiclayo.

- Soto Olivares, J. P. (2011). *Solución de Inteligencia de Negocios para una PYME*. Valparaíso: PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO.
- Turban, E., Sharda, R., Aronson, J. E., & king, D. (2008). *Business Intelligence: A Managerial Approach*. Pearson/Prentice Hall.
- Vanegas Lago, E., & Guerra Cantero, L. (2013). Sistema de inteligencia de negocios para el apoyo al proceso de toma de decisiones. *Revista Ingeniería UC* 20 (3), 25 - 34.



ANEXOS

ANEXO 1.- ENTREVISTA SEMI ESTRUCTURADA PREVIA A IMPLEMENTACIÓN

GUÍA DE ENTREVISTA SEMIESTRUCUTURADA

Indicaciones

Para el desarrollo de la entrevista se debe de tener los siguientes requerimientos:

Materiales

- Grabadora de voz
- Cuaderno de apuntes
- Bebida (agua)

Ambiente

- Oficina del jefe de producción

Involucrados

- Gerente de producción
- Jefe de producción
- Jefe de mantenimiento

Procedimiento

Se debe de considerar la comunicación para las reuniones de coordinación, las mismas que serán tres:

- Reunión inicial: se establece los requerimientos
- Reunión de propuesta: se establece la aceptación del desarrollo a ejecutar
- Reunión de avances: se considera la evaluación y alcances.

La ficha de entrevista contempla las siguientes preguntas:

Pregunta 1: ¿Qué es lo que requiere?

Pregunta 1.2: ¿Existe un presupuesto y tiempo? Derivada de la pregunta 1

Pregunta 1.3: ¿Intervendrá el área de TI?

Pregunta 1.4: ¿Se contratará un personal encargado?

Pregunta 2: ¿Cuáles son los factores o indicadores a controlar? ¿Por qué?

Pregunta 3: ¿Dónde desean se visualice la información local o nacional y quien tendrá acceso?

Pregunta 4: ¿Qué recursos se podrán a disposición para el desarrollo del software y cuánto tiempo demorará?

A continuación, las respuestas a las preguntas planteadas.

Pregunta 1: ¿Qué es lo que requiere?

A esta pregunta se indicó que la empresa requiere reducir tiempos en los procesos manuales, para realizar un mejor control evitando en la medida de lo posible la adquisición de nuevas licencias, es decir, que se puede hacer con lo hay y se tiene.

A esta respuesta, se indicó que se analizaría cómo optimizar los recursos que se tienen.

Pregunta 1.2: ¿Existe un presupuesto y tiempo? Derivada de la pregunta 1

La respuesta fue, que no se había dado en el presupuesto de la empresa, sin embargo, se daría todo el apoyo en el tiempo que se requiera. Para eso fue necesario revisar la bibliografía existente, con la finalidad de encontrar la mejor solución posible.

Pregunta 1.2: ¿Intervendrá el área de TI?

Sobre este punto se indicó que el área de TI se encuentra en las instalaciones de Lima, que se puede realizar coordinaciones con ellos en el tema de infraestructura requerida. Sobre todo tienen a su disposición el servicio de mesa de ayuda, quienes normalmente brindan asistencia remota.

Pregunta 1.4: ¿Se contratará personal encargado? ¿Quién será el responsable para futura administración?

Se respondió que no está dentro del presupuesto el poder contratar a una persona. El sistema funcionará dentro de las instalaciones de la empresa y para efectos de administración o mantenimiento se solicitó que se considerara este en una propuesta económica independiente.

Pregunta 2: ¿Cuáles son los factores o indicadores a controlar? ¿Por qué?

El resultado de esa pregunta fue que deseaban medir:

- Indicador de aprovechamiento de materia prima.- se necesita saber cuánto es lo que se emplea de materia primaria en nuestro proceso de producción, porque necesitamos tener el conocimiento de cuanto es lo que se está aprovechando, para poderlo valorizar en físico y económico.
- Indicador de excesos.- se necesita saber a qué distancia de cero se tiene el registro, este determina la pérdida del producto; porque al ser más alejado de cero, de manera positiva, se pierde más producto, que en la industria se considera como dinero, y cuanto más cercano a cero esté, todos los procesos son correctos y positivos económicamente hablando.
- Indicador de eficiencia del molino.- se necesita calcular, si el proceso es eficiente o no, porque nos permitirá tomar decisiones correctivas, para estos se sostiene el cálculo de la relación “toneladas hora producidas / toneladas hora de capacidad de planta” y reflejada en un valor porcentual.

Pregunta 3: ¿Dónde desean se visualice la información y quien tendrá acceso?

Sobre esta pregunta se precisó que era necesario poder acceder a la solución desde cualquier punto dentro de las instalaciones de la empresa. Los principales usuarios disponen de equipos portátiles y suelen movilizarse bastante, solo en el caso de los supervisores los que disponen de un equipo de escritorio.

Pregunta 4: ¿Qué recursos se podrán a disposición para el desarrollo del software y cuánto tiempo demoraría solicitar nueva infraestructura?

Actualmente se dispone de 2 equipos de escritorio, uno de ellos que contiene el sistema SCADA que pueden ser libremente usados. En caso se requiera de nueva infraestructura se puede gestionar a través de mesa de ayuda la solicitud y puede ser atendida en el plazo de 1 mes calendario.

Después de finalizar la implementación del proyecto:

Pregunta 5: ¿Cuál fue el resultado de su evaluación?

ANEXO 2.- ENTREVISTA SEMI ESTRUCTURADA AL FINAL DE LA IMPLEMENTACIÓN

GUÍA DE ENTREVISTA SEMIESTRUCUTURADA

Indicaciones

Para el desarrollo de la entrevista se debe de tener los siguientes requerimientos:

Materiales

- Grabadora de voz
- Cuaderno de apuntes
- Bebida (agua)

Ambiente

- Oficina del jefe de producción

Involucrados

- Gerente de producción
- Jefe de producción
- Supervisor

Procedimiento

Se consideró hacer preguntas abiertas con el fin de obtener respuesta de los tres expertos involucrados en la reunión. Se formuló la siguiente guía de preguntas para la entrevista:

Pregunta 1.- ¿La solución es de utilidad para el control de la producción?

Pregunta 2.- ¿La solución implementada es confiable? ¿Representa valores reales sobre la producción?

Pregunta 3.- Acerca de los valores y gráficas que se representan en la interfaz, ¿es posible tomar decisiones con esos datos?

Pregunta 4.- La plataforma de datos históricos, ¿le es de ayuda para un planeamiento estratégico o táctico a futuro?

ANEXO 3.- PLAN DE TESIS

UNIVERSIDAD CATOLICA DE SANTA MARIA

**FACULTAD DE CIENCIAS E INGENIERIAS FISICAS Y
FORMALES**

ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS



PLAN DE TESIS

**MEJORA DEL PROCESO DE PRODUCCIÓN DE UNA EMPRESA
PRODUCTORA DE HARINAS APLICANDO INTELIGENCIA DE
NEGOCIOS CON HERRAMIENTAS OPEN SOURCE**

ALUMNO:

Montánchez Montesinos Arturo Renato

**AREQUIPA - PERÚ
2017**

1. PLANTEAMIENTO DEL PROBLEMA

1.1. Caracterización del Problema

Al día de hoy, convertir los datos en información es de vital importancia para cualquier organización que desee desarrollarse y ser competitiva. Las organizaciones líderes están en la búsqueda constante de caminos novedosos e inteligentes para mejorar su desempeño en varios aspectos dentro de la misma; y a su vez esperan que cada individuo sea capaz de aportar para estos resultados.

La inteligencia de negocios o BI (Business Intelligence) es un factor clave que permite alcanzar esos resultados ya que permite una toma de decisiones informada en todos los niveles de la organización permitiendo a cada uno de estos actuar de la manera más adecuada y efectiva en una situación dada. Aun así, para cada empresa representa un desafío encontrar la mejor estrategia que se adecue a sus recursos y presupuesto.

En Alicorp S.A.A – Arequipa, en su área de producción “Molino Victoria” disponen de un sistema SCADA (Supervisory Control And Data Acquisition) que les permite visualizar el estado actual y funcionamiento de sus equipos, así como el flujo actual de sus productos medidos en balanzas específicas para cada producto. Sin embargo, la visualización de datos exactos sobre producción de cada bien, o el acceso a un histórico visualmente amigable no es posible ya que el sistema SCADA no tiene dicho fin.

Es así que identificada la falta de información precisa y confiable en Alicorp – Arequipa, se requiere de una solución que pueda brindarles las herramientas necesarias para acceder a la información sobre la producción de sus Molinos, así como poder consultar un histórico, visualizar tendencias y disponer de diferentes tipos de criterios en cuanto a las consultas realizadas.

1.2. Línea y Sub-línea de Investigación

Línea.- Sistemas de Información y Bases de Datos

Sublínea.- Base de datos y Datawarehouse

2. OBJETIVOS DEL PROYECTO

2.1. General

Mejorar el proceso de producción de una empresa productora de harinas aplicando inteligencia de negocios con herramientas Open Source.

2.2. Específicos

- 2.2.1. Automatizar y optimizar el trabajo de recolección de datos acerca del proceso de producción.
- 2.2.2. Implementar un repositorio de datos óptimo para el almacenamiento de los datos.
- 2.2.3. Implementar una solución basada en herramientas de BI.
- 2.2.4. Implementar reportes de acuerdo a las necesidades del usuario.
- 2.2.5. Implementar una solución de monitoreo.

3. FUNDAMENTOS TEÓRICOS

3.1. Estado del arte

Según Flórez (2015) “En la actualidad, la información histórica de un negocio es el fundamento para comprender su comportamiento. La inteligencia de negocios les permite a los gerentes y personal directivo de una organización identificar claramente su prospectiva, debido a que esta le provee resultados que indican el comportamiento de dicha organización.

La inteligencia de negocios consiste en extraer y analizar la información de una organización con el fin de obtener su conocimiento para aplicarlo en la organización, de manera que pueda orientar adecuadamente la toma de

decisiones y permita un mejoramiento incremental de los procesos para la obtención de mejores resultados”.

Es una práctica común de hoy el hacer uso de herramientas de inteligencia de negocios que permitan a las empresas evolucionar y pasar de ser reactivas a ser proactivas, esta evolución está definida en 03 etapas (Gartner, 2012) que son la empresa centrada en TI, donde se enfoca en la mejora de recolección de datos y selección de herramientas de análisis; La administración de información, integrando las herramientas de análisis y sus resultados con las herramientas llaves del negocio; y para finalizar la Visión predictiva, con la que la empresa puede predecir el futuro comportamiento de algo en específico y adelantarse al mismo para sacar el mejor provecho o prevenir algún mal porvenir.

3.2. Bases Teóricas del proyecto

Inteligencia de Negocios (o Business Intelligence, BI).- Se considera su concepto referido a la inteligencia como aquella información apreciada por su valor y relevancia. Es información experta, conocimiento y tecnologías eficientes en el manejo de negocios organizacionales e individuales. Por lo tanto, la inteligencia de negocios es una amplia categoría de aplicaciones y tecnologías para recolectar, proveer acceso a, y analizar data con el propósito de ayudar a que los usuarios empresariales puedan tomar mejores decisiones sobre negocios. (J. Ranjan, 2009).

Data warehouse (DW).- Es el componente significativo de la inteligencia de negocios. Es orientado a temas, integrado. El data warehouse soporta la propagación física de data manejando los numerosos registros empresariales para tareas de integración, limpieza, agregación y consulta. También puede contener la data operativa, la cual puede ser definida como un conjunto integrado de data actualizable usada para una amplia toma de decisiones tácticas de un área en particular. Contiene data viva, no fotografías, y retienen un mínimo de historia. Las fuentes de datos pueden ser bases de datos operativas, data histórica, data externa como por ejemplo de compañías de investigación de

mercado o de Internet, o información del ambiente de un data warehouse que ya exista. (J. Ranjan, 2009)

Data Mart.- Según lo descrito por Inmon (1999), es una colección de campos de información organizados para el soporte a la toma de decisiones basado en la necesidad de un departamento en específico. Cada departamento tiene su propia interpretación de cómo debería verse su datamart, y cada datamart de cada departamento es peculiar y específico a sus propias necesidades.

OLAP (On-line analytical processing): De acuerdo a Ranjan (2009) Se refiere a la manera en la que los usuarios del negocio pueden cortar y arreglar su camino a través de los datos utilizando herramientas sofisticadas que permiten la navegación de dimensiones tales como el tiempo o las jerarquías. Procesamiento analítico en línea u OLAP proporciona vistas multidimensionales y resumidas de datos comerciales y se utiliza para generar informes, análisis, modelado y planificación para optimizar el negocio. Las técnicas y herramientas OLAP se pueden usar para trabajar con almacenes de datos o data marts diseñados para sistemas sofisticados de inteligencia empresarial. Estos sistemas procesan consultas necesarias para descubrir tendencias y analizar factores críticos. El software de generación de informes genera vistas agregadas de datos para mantener a la administración informada sobre el estado de su negocio

4. PRESENTACIÓN DEL PROYECTO

4.1. Justificación

El acceso a la información es de vital importancia para cada compañía. El hecho de no disponer de la información de producción de manera accesible y oportuna genera una baja en el desempeño de la planta, así como la incapacidad de medir el rendimiento real del Molino. Adicionalmente, para la toma de decisiones se precisaba consultar un parte diario de producción que se llena a mano y que no representa información fidedigna ya que la toma de datos tiene un nivel de incertidumbre.

Para finalizar, no existe un reporte automatizado que permita visualizar una tendencia o estadística de los indicadores establecidos ni de los productos propios del Molino.

4.2 Resumen del Proyecto

4.2.1 Descripción del Proyecto a medio y largo plazo.

Una vez iniciado el proyecto, y habiendo definido los indicadores a medir, así como los productos se establecerá la herramienta de base de datos con PostgreSQL, la cual además de ser libre, es una de las más potentes en el mercado de bases de datos. Se definirá el modelo entidad relación en el que se almacenarán los diversos datos que luego explotaremos. Una vez se haya instalado la base de datos y definido el modelo y esquema, procederemos a realizar la conexión del SCADA hacia la base de datos, estableciendo el criterio de escritura de los datos.

Se definirá la funcionalidad necesaria al interior de la base de datos, como lo son funciones, procedimientos y disparadores.

Posteriormente se continuará con la construcción del datamart respectivo, se definirá el modelo ETL, haciendo uso de la herramienta Spoon de Pentaho, que cargará los datos limpios y transformados hacia el datamart y se programará su ejecución automática con el propio programador de tareas de Windows.

Para finalizar utilizaremos la herramienta Schema Workbench de Pentaho para crear el esquema del cubo OLAP correspondiente y cargarlo a la herramienta Pentaho Business Analytics Server.

Culminaremos el proyecto con las respectivas pruebas y capacitación del usuario final. Para la herramienta de monitoreo, se desarrollará haciendo uso de las tecnologías JAVA y Javascript.

4.2.2 Usuarios del Proyecto.

Los principales usuarios de este proyecto son aquellos que pertenecen al nivel táctico y estratégico de la misma empresa, quienes pueden hacer consultas de

acuerdo a sus propios parámetros y extraer la información necesaria y pertinente para una toma de decisión acertada.

4.2.3 Beneficios.

Este proyecto permitirá a la empresa:

- Implementar una solución de inteligencia de negocios basada en la tecnología actual, accesible, disponible y adaptable. Además, las herramientas utilizadas serán Open Source, es decir, totalmente gratuitas lo que permitirá un ahorro considerable en costos de licencias.
- Tener un conocimiento actual y preciso de manera oportuna sobre la situación actual de la producción en sus molinos; además podrá analizar el comportamiento o tendencia de cada uno de sus productos a través del tiempo, relacionando información clave para disponer del conocimiento adecuado.
- A través del detalle de los datos, se podrá mantener un control más efectivo sobre los indicadores de producción, logrando un control en tiempo real.

4.2.4 Localización.

El proyecto se desarrollará en una empresa productora de harinas ubicada en Arequipa, Perú. Esta empresa ubica uno de sus molinos y líneas de producción de Harinas en Arequipa, a pesar que la mayor parte de su área administrativa se ubica en Lima.

4.2.5 Análisis del futuro del Proyecto.

Una vez se disponga de la estructura base claramente definida, se podrá explotar información de diferentes maneras, poder armar diversos reportes y dashboards que permitan hacer un seguimiento y monitoreo constante del rendimiento del área de producción. Además, con el éxito de este proyecto se puede plantear

ampliar el alcance hacia otros ámbitos con el fin de complementar la información disponible y poder adquirir mejor conocimiento.

4.2.6 Riesgos que debemos afrontar.

- **Económicos:** En caso que el riesgo tecnológico (abajo descrito) sea muy grave y no permita que la solución se desenvuelva con normalidad, existe la probabilidad de requerir una mejor infraestructura con características adecuadas que puedan dar soporte al software utilizado. La adquisición de un servidor o equipo con dichas características es un costo que la empresa tardaría en evaluar y aprobar, por lo que el proyecto se desenvolvería con dicha incertidumbre.
- **Competencia:** La competencia como tal no representa un riesgo directo para el desenvolvimiento de este proyecto ya que los usuarios y todo lo involucrado pertenecen al ambiente interno de la empresa, sin interactuar con el medio externo.
- **Tecnológicos:** La capacidad de la infraestructura tecnológica actual de la empresa, quienes ponen a disposición dos terminales PC para la implementación de la solución completa. Dichas computadoras tienen una limitación en cuanto a memoria RAM de cada terminal que puede afectar directamente el desempeño de las herramientas Open Source a instalar.
- **No tecnológicos:** El riesgo conocido es la resistencia al cambio del propio personal de la empresa, quienes pueden no estar de acuerdo a aprender el uso de una nueva herramienta para extraer la información necesaria.

5. PLAN DE IMPLANTACIÓN DEL PROYECTO

5.1 Definición del Proyecto

5.1.1 Aspectos Técnicos

Para el desarrollo del proyecto se desarrollará una solución de inteligencia de negocios, que consiste en una base de datos especializada para almacenar información específica. El desarrollo del mismo se llevará a cabo siguiendo el ciclo de inteligencia de negocios, y las mejores prácticas para la implementación de un datawarehouse, y los datamart respectivos.

5.1.2 Aspectos Económicos

En cuanto al aspecto económico, se han evitado costos de las herramientas al escoger tecnologías Open Source o gratuitas, las cuales permiten hacer uso de versiones comunitarias gratuitas, que aunque presentan limitaciones, pueden abarcar y satisfacer las necesidades del cliente de efectivamente.

5.1.3 Aspectos Comerciales

Este proyecto hará uso de herramientas Open Source con el fin de evitar el pago de licencias de software. Dentro de las herramientas consideradas se encuentra JasperSoft Report Server, Pentaho Tools, Tomcat Server y la base de datos PostgreSQL. Al culminar el proyecto permitirá a la empresa optimizar su proceso de producción y obtener beneficios económicos con dicha optimización.

5.1.4 Recursos del Proyecto

Dentro del proyecto se necesitarán recursos tecnológicos, como lo son las computadoras donde funcionarán las herramientas, recursos humanos, como lo son los usuarios finales a cargo de validar las pruebas y el desarrollador de la solución.

6. Metodología a emplear.

a) Análisis de la situación actual: Se realizará un breve estudio de viabilidad del proyecto estudiando la situación actual, igualmente se definirán los recursos disponibles; posterior a este estudio previo se definirán los requisitos del proyecto en sí, al igual que se recabará información de los clientes o usuarios finales para tener en claro el objetivo y las metas a alcanzar. Se definirá el enfoque que se ha de tomar para el desenvolvimiento de este proyecto.

b) Diseño: Diseño de la arquitectura de la solución; comprende desde el diseño de la base de datos que ha de soportar toda la solución, el data mart con sus respectivas tablas dimensionales y de hechos, el diseño del proceso ETL que llevará a cabo la carga de datos, el esquema del cubo OLAP y para culminar los reportes que el usuario final considere pertinentes.

c) Implementación: Se implementará la solución completa, iniciando por instalar base de datos PostgreSQL y definir toda su funcionalidad, posteriormente se instalarán las herramientas de Pentaho desde Spoon para el desarrollo del ETL, hasta el servidor de Pentaho y el plugin Mondrian.

d) Pruebas: Cabe mencionar que las pruebas se desarrollarán de manera modular y en serie, es decir, por cada etapa de la implementación se realizarán pruebas unitarias con el fin de comprobar el correcto funcionamiento de cada implementación dada. Al final se realizarán pruebas automáticas y manuales de la solución integral.

7. PLAN DE TRABAJO

El proyecto constará de varias etapas, dentro de las cuales se han de tomar ciertos pasos para la culminación de cada una de ellas; así tenemos:

1. Fase previa al inicio del proyecto (10 días calendario)

- a. Estudio de viabilidad e identificación de recursos
- b. Entrevistas con usuarios.
- c. Establecimiento del alcance y objetivo del proyecto.

2. Fase I: Base de Datos (20 días calendario)

- a. Diseño del diagrama Entidad-Relación de la base de datos transaccional.
- b. Creación de Tablas y dependencias.
- c. Programación de funciones y disparadores.
- d. Diseño de DataMart

3. Fase II: Desarrollo de herramienta de monitoreo (10 días calendario)

4. Fase III: ETL y OLAP (07 días calendario)

- a. Identificación de entidades y campos a usar.
- b. Diseño del diagrama ETL y estandarización.
- c. Diseño del esquema del cubo OLAP, identificando dimensiones medidas y miembros calculados.

5. Fase IV: Instalación de Herramientas (03 días calendario)

- a. Instalación de base de Datos PostgreSQL en terminal del usuario.
- b. Instalación de Tomcat Server.
- c. Instalación de Herramientas de Pentaho y sus respectivos requisitos (Java).
- d. Instalación de servidor de Reportes Jaspersoft Community.
- e. Configuración de herramientas instaladas.

6. Fase V: Pruebas Integrales (07 días calendario)

- a. Pruebas modulares
- b. Pruebas integrales.

8. REFERENCIAS BIBLIOGRÁFICAS

1. J. Ranjan , 2009, Business intelligence: concepts, components, Techniques and Benefits
2. S. Rouhani et. al, 2012, Review Study: Business Intelligence Concepts and Approaches
3. N. Bernal, C. Suarez, 2015, Agile methodology for modeling and design of data warehouses
4. S. Negash, 2004, Business Intelligence, Communications of the Association for Information Systems
5. C. Olszak, E. Ziemba, 2007, Approach to Building and Implementing Business Intelligence Systems.
6. H. Watson, 2007, The Current State of Business Intelligence.
7. Flórez, 2015, Guía para la implementación de una solución de inteligencia de negocios para pequeñas y medianas empresas.
8. Gartner, 2012, A step-by-step approach to successful Business Intelligence.