

# **UNIVERSIDAD NACIONAL DEL SANTA**

## **FACULTAD DE INGENIERIA**

**Escuela Profesional de Ingeniería de Sistemas e Informática**



### **SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS PARA MEJORAR LA TOMA DE DECISIONES EN LA DIRECCIÓN EJECUTIVA DEL HOSPITAL LA CALETA**

**Tesis para optar el Título de Ingeniero de Sistemas e Informática**

#### **TESISTAS:**

Bach. ALVA QUILICHE MIGUEL ANGELLO

Bach. CALLAN CARBAJAL MILAGROS ROXANA

#### **ASESOR:**

MS. YIM ISAIAS APESTEGUI FLORENTINO

**NUEVO CHIMBOTE -PERU**

**2017**



**UNIVERSIDAD NACIONAL DEL SANTA**

**FACULTAD DE INGENIERIA**

**Escuela Profesional de Ingeniería de Sistemas e Informática**

**TITULO**

**SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS PARA  
MEJORAR LA TOMA DE DECISIONES EN LA DIRECCIÓN  
EJECUTIVA DEL HOSPITAL LA CALETA**

**Tesis para optar el Título de Ingeniero de Sistemas e Informática**

**REVISADO Y APROBADO POR:**

---

**Asesor**

**NUEVO CHIMBOTE –PERU**

**2017**

**UNIVERSIDAD NACIONAL DEL SANTA**

**FACULTAD DE INGENIERIA**

**Escuela Profesional de Ingeniería de Sistemas e Informática**

**TITULO**

**SOLUCIÓN DE INTELIGENCIA DE NEGOCIOS PARA  
MEJORAR LA TOMA DE DECISIONES EN LA DIRECCIÓN  
EJECUTIVA DEL HOSPITAL LA CALETA**

**Tesis para optar el Título de Ingeniero de Sistemas e Informática**

**REVISADO Y APROBADO POR EL JURADO EVALUADOR:**

---

**Presidente**

---

**Secretario**

---

**Miembro**

---

**Accesitario**

**NUEVO CHIMBOTE –PERU**

**2017**

## **DEDICATORIA A:**

Dios por mostrarme día a día que con humildad, paciencia y sabiduría que todo es posible.

Mi padre, que siempre lo he sentido presente en mi vida. Y sé que está orgulloso de la persona en la cual me he convertido.

Mi madre, por brindarme su confianza y apoyo, que sin duda en el trayecto de mi vida me ha demostrado su amor, corrigiendo mis faltas y celebrando mis triunfos.

A todas las personas, que siempre estuvieron listas para brindarme toda su ayuda, ahora me toca regresar un poquito de todo lo inmenso que me han otorgado.

Mi hermana, que con sus consejos me ha ayudado a afrontar los retos que se me han presentado a lo largo de mi vida.

De: Miguel Angello Alva Quiliche

## **DEDICATORIA A:**

Dios, que siempre me ilumina, me protege y cuida; dándome fuerzas para seguir adelante, conservando la fe y la esperanza.

¡Mi madre! Sra. Flora Rosa Carbajal Romero por darme su apoyo y amor para seguir adelante, inculcándome el deseo de superación bajo cualquier circunstancia.

A mis amigos y Compañeros, que nunca dejaron de confiar en mí y siempre estaban conmigo cuando los necesite.

A todas aquellas personas quienes con su amor, apoyo y comprensión incondicional estuvieron siempre a lo largo de mi vida universitaria, a ellos que siempre tuvieron una palabra de aliento en los momentos difíciles y que han incentivado mi vida.

Mis profesores, por su tiempo, por su apoyo así como por la sabiduría que me transmitieron en el desarrollo de mi formación profesional.

De: Milagros Roxana Callan Carbajal

## **AGRADECIMIENTO**

Deseamos expresar nuestro más sincero agradecimiento a:

- Dios por darnos la fuerza de voluntad para seguir adelante.
  
- Nuestros Padres por habernos apoyado en todo lo que respecta a nuestros estudios hasta la culminación.
  
- Nuestros Profesores, que mediante sus enseñanzas han ido moldeando, durante nuestra vida universitaria, todos los requisitos que son necesarios para la Formación Académica de todo buen profesional.
  
- Nuestro Asesor por ayudarnos a culminar el informe.
  
- Nuestros amigos y compañeros de Escuela de Ingeniería de Sistemas e Informática.

De: Los Autores

# INDICE

## CONTENIDO

DEDICATORIA A:.....	v
AGRADECIMIENTO .....	vii
INDICE.....	viii
LISTA DE TABLAS .....	xiii
LISTA DE FIGURAS .....	xiv
RESUMEN .....	xvii
ABSTRACT .....	xviii
PRESENTACIÓN .....	xix
INTRODUCCIÓN.....	1
CAPITULO I.....	3
PROBLEMA DE INVESTIGACIÓN .....	3
1.1 Planteamiento y fundamentación del problema de investigación .....	3
1.1.1.Planteamiento del Problema.....	3
1.1.2.Fundamentación del problema. ....	5
1.2 Antecedentes de la investigación.....	7
A nivel internacional: .....	7
A nivel nacional: .....	9
1.3 Formulación del problema de investigación .....	11
1.4 Delimitación del estudio.....	11
1.5 Justificación e importancia de la investigación .....	11
Justificación teórica .....	11
Justificación metodológica .....	12
Justificación práctica .....	12
Importancia.....	13
1.6 Objetivos de la investigación: general y específicos .....	13
1.6.1Objetivo general. ....	13
1.6.2Objetivos específicos.....	13
CAPITULO II.....	14
MARCO TEÓRICO .....	14
2.1 Fundamentos teóricos de la investigación.....	14
2.1.1.Toma de Decisiones. ....	14

2.1.2.Inteligencia de Negocios. ....	15
2.1.2.1.Definición. ....	15
2.1.2.2.Proceso de Inteligencia de Negocios. ....	16
2.1.2.3.Ventajas. ....	16
2.1.3.ETL(Extracción, Transformación y Carga).....	17
2.1.3.1.Definición. ....	17
2.1.4.Data Warehouse (DWH). ....	17
2.1.4.1.Definición. ....	17
2.1.4.2.Ventajas. ....	18
2.1.5.Proceso Analítico en Línea (OLAP).....	18
2.1.6.Pentaho. ....	19
2.1.6.1.Definición. ....	19
2.1.6.2.Herramientas y Áreas de Aplicación. ....	20
2.1.6.2.1.Pentaho Data Integration. ....	20
2.1.6.2.2.Pentaho Schema Workbench.....	21
2.1.6.2.3.Pentaho Analysis. ....	21
2.2 Marco conceptual .....	22
CAPITULO III .....	24
MARCO METODOLÓGICO .....	24
3.1 Hipótesis central de la investigación .....	24
3.2 Variables e indicadores de la investigación .....	24
3.2.1.Variables.....	24
3.2.2.Definición operacional. ....	24
3.3 Métodos de la investigación .....	26
3.4 Diseño o esquema de la investigación.....	26
3.5 Población y muestra .....	27
3.2.3.Población. ....	27
3.2.4.Muestra.....	29
3.6 Actividades del proceso investigativo .....	30
3.7 Técnicas e instrumentos de la investigación .....	30
Entrevista.....	30
Encuesta.....	30
Observación.....	30

3.8	Procedimiento para la recolección de datos .....	31
3.9	Técnicas de procesamiento y análisis de datos.....	31
	CAPITULO IV .....	32
	RESULTADOS Y DISCUSIÓN .....	32
4.1	Resultados de la inteligencia de negocios .....	32
4.1.1	Dirigir y planear. ....	32
4.1.2	Recolección de información.....	33
4.1.3	Procesamiento de datos. ....	39
4.1.3.1	<i>Análisis de Requerimientos.....</i>	39
4.1.3.1.1	<i>Identificar Preguntas. ....</i>	39
4.1.3.1.2	<i>Identificar Indicadores y Perspectivas.....</i>	40
4.1.3.1.3	<i>Modelo Conceptual. ....</i>	40
4.1.3.2	<i>Análisis de las Fuentes. ....</i>	41
4.1.3.2.1	<i>Conformar Indicadores.....</i>	41
4.1.3.2.2	<i>Establecer Correspondencias. ....</i>	42
4.1.3.2.3	<i>Nivel de Granularidad .....</i>	44
4.1.3.2.4	<i>Modelo conceptual ampliado. ....</i>	46
4.1.3.3	<i>Modelo lógico del data warehouse.....</i>	48
4.1.3.3.1	<i>Tipo de modelo lógico del data warehouse.....</i>	48
4.1.3.3.2	<i>Establecer tabla de dimensiones.....</i>	48
4.1.3.3.3	<i>Establecer tabla de hechos.....</i>	52
4.1.3.3.4	<i>Uniones. ....</i>	53
4.1.3.4	<i>Integración de datos. ....</i>	54
4.1.3.4.1	<i>Carga inicial. ....</i>	54
4.1.3.4.2	<i>Actualización.....</i>	63
4.1.3.5	<i>Creación de cubos multidimensionales. ....</i>	64
4.1.3.5.1	<i>Creación de indicadores. ....</i>	65
4.1.3.5.1.1	<i>Cubo productividad. ....</i>	65
4.1.3.5.1.2	<i>Cubo cantidad.....</i>	65
4.1.3.5.1.3	<i>Cubo cumplimiento.....</i>	66
4.1.3.5.1.4	<i>Cubo promedio. ....</i>	67
4.1.3.5.2	<i>Creación de atributos.....</i>	68
4.1.4	Análisis y producción.....	76

4.1.5 Difusión .....	84
4.2 Discusión .....	85
4.2.1 Contrastación de la hipótesis .....	85
4.2.1.1 Indicadores cualitativos .....	86
4.2.1.1.1 Cálculo para hallar el Nivel De Satisfacción .....	88
4.2.1.1.2 Cálculo para medir la Calidad de la información .....	96
4.2.1.2 Indicadores cuantitativos .....	103
4.2.1.2.1 Prueba de la Hipótesis para el indicador cuantitativo .....	103
4.2.1.2.2 Prueba de la Hipótesis para el indicador cuantitativo .....	109
4.2.1.2.3 Prueba de la Hipótesis para el indicador cuantitativo .....	115
4.3 Resultados .....	121
4.3.1 Indicador cualitativo: nivel de satisfacción de los usuarios .....	121
4.3.2 Indicador cualitativo: calidad de la información .....	122
4.3.3 Indicador cuantitativo: tiempo de generación de reportes que soportan la t... ..	124
4.3.4 Indicador cuantitativo: costo de generación de reportes .....	125
4.3.5 Indicador cuantitativo: tiempo para analizar información y formular alternat. ..	126
CAPITULO V .....	128
CONCLUSION Y RECOMENDACIONES .....	128
5.1 Conclusiones .....	128
5.2 Recomendaciones .....	130
REFERENCIAS BIBLIOGRÁFICAS .....	132
REFERENCIAS DE PÁGINAS EN LA WEB .....	133
ANEXO A .....	134
CONEXIÓN SQL SERVER 2008 .....	134
6.1 Configuración de la base de datos en SQL Server 2008 .....	134
6.2 Creación de la base de datos DW en SQL Server 2008 .....	138
ANEXO B .....	139
CONEXIÓN A PENTAHO .....	139
7.1 Configuración de la conexión de Pentaho con SQL Server 2008 .....	139
ANEXO C .....	141
CUESTIONARIOS .....	141
8.1 Cuestionarios para hallar el Nivel de Satisfacción de los Usuarios .....	141
8.2 Cuestionarios para hallar la Calidad de Información .....	143

ANEXO D .....	145
CONFIABILIDAD POR ALFA DE CRONBACH .....	145
9.1 Encuesta para indicador: nivel de satisfacción de los usuarios .....	145
9.2 Encuesta para indicador: calidad de la información.....	146

## LISTA DE TABLAS

<b>Tabla N° 1:</b> Matriz de Operacional de variables. ....	24
<b>Tabla N° 2:</b> Indicadores.....	26
<b>Tabla N° 3:</b> Recursos Humanos del Hospital la Caleta.....	27
<b>Tabla N° 4:</b> Conclusión del Modelo Conceptual “Respectivas”.....	43
<b>Tabla N° 5:</b> Conclusión del Modelo Conceptual “Indicadores.....	43
<b>Tabla N° 6:</b> Perspectiva “Hora” .....	44
<b>Tabla N° 7:</b> Perspectiva "Médico". .....	45
<b>Tabla N° 8:</b> Perspectiva "Servicio". .....	45
<b>Tabla N° 9:</b> Perspectiva "Tiempo". .....	45
<b>Tabla N° 10:</b> Perspectiva "Paciente". .....	46
<b>Tabla N° 11:</b> Perspectiva "Especialidad". .....	46
<b>Tabla N° 12:</b> Determinación de pruebas estadísticas de los indicadores. ....	85
<b>Tabla N° 13:</b> Rango de valores.....	86
<b>Tabla N° 14:</b> Lista usuarios. ....	86
<b>Tabla N° 15:</b> Ponderación del Usuario del Sistema - Nivel de Satisfacción.....	88
<b>Tabla N° 16:</b> Tabulación usuario del sistema Pre Test - Nivel de Satisfacción .....	89
<b>Tabla N° 17:</b> Tabulación usuario del sistema Post Test - Nivel de Satisfacción .....	90
<b>Tabla N° 18:</b> Contrastación Pre & Post Test - Nivel de Satisfacción. ....	91
<b>Tabla N° 19:</b> Ponderación del Usuario del Sistema – Calidad de la Información. ....	96
<b>Tabla N° 20:</b> Tabulación usuario del sistema Pre Test - Calidad de la Información .....	97
<b>Tabla N° 21:</b> Tabulación usuario del sistema Post Test – Calidad de la Información. ...	98
<b>Tabla N° 22:</b> Contrastación Pre & Post Test – Calidad de la Información. ....	98
<b>Tabla N° 23:</b> Toma de Tiempo de Generación de Reportes.....	105
<b>Tabla N° 24:</b> Toma de Costos de Generación de Reportes. ....	111
<b>Tabla N° 25:</b> Toma de Tiempos para analizar información y formular alternativas.....	118
<b>Tabla N° 26:</b> Presentación de Resultados Nivel de Satisfacción del Usuario.....	121
<b>Tabla N° 27:</b> Presentación de Resultados de Calidad de la Información.....	123
<b>Tabla N° 28:</b> Presentación de Resultados de Tiempo de generación de reportes que soportan la toma de decisiones. ....	124
<b>Tabla N° 29:</b> Presentación de Resultados de Costo de generación de reportes. ....	126
<b>Tabla N° 30:</b> Presentación de Resultados del Tiempo de analizar la información y formular alternativas.....	127
<b>Tabla N° 31:</b> Estadísticos de la encuesta para el Indicador Nivel de Satisfacción. ....	145
<b>Tabla N° 32:</b> Estadísticos de la encuesta para el Indicador Calidad de la Información	146

## LISTA DE FIGURAS

<b>Figura N° 1</b> : El proceso de la toma de decisiones en la empresa. ....	15
<b>Figura N° 2:</b> Fases del proceso BI. ....	16
<b>Figura N° 3</b> : Arquitectura Data Warehouse. ....	18
<b>Figura N° 4:</b> Plataforma Pentaho. ....	20
<b>Figura N° 5:</b> BD-Tabla de Atenciones. ....	34
<b>Figura N° 6:</b> BD-Tabla de Empleados. ....	35
<b>Figura N° 7:</b> BD-Tabla de Especialidades. ....	35
<b>Figura N° 8:</b> BD-Tabla de Medicos. ....	36
<b>Figura N° 9:</b> BD-Tabla de ProgramaciónMedica. ....	36
<b>Figura N° 10:</b> BD- Tabla de Pacientes. ....	37
<b>Figura N° 11:</b> BD- Tabla de Servicios. ....	38
<b>Figura N° 12:</b> BD- Tabla de MedicosEspecialidad. ....	38
<b>Figura N° 13:</b> Modelo conceptual del proceso de Gestión Asistencial. ....	41
<b>Figura N° 14:</b> Modelo Conceptual Ampliado del proceso de Gestión Asistencial. ....	47
<b>Figura N° 15:</b> Tabla de Dimensión "PROGRAMACIÓN" ....	48
<b>Figura N° 16:</b> Tabla de Dimensión "MÉDICO". ....	49
<b>Figura N° 17:</b> Tabla de Dimensión "SERVICIO". ....	49
<b>Figura N° 18:</b> Tabla de Dimensión "FECHA". ....	50
<b>Figura N° 19:</b> Tabla de Dimensión "PACIENTE" ....	51
<b>Figura N° 20:</b> Tabla de Dimensión "ESPECIALIDAD". ....	51
<b>Figura N° 21:</b> Tabla de Hechos del proceso de Gestión Asistencial. ....	52
<b>Figura N° 22:</b> Uniones correspondientes. ....	53
<b>Figura N° 23:</b> PENTAHO- Carga de Dimensión PROGRAMACION. ....	55
<b>Figura N° 24:</b> PENTAHO- Carga de Dimensión MEDICO. ....	56
<b>Figura N° 25:</b> PENTAHO - Carga de Dimensión SERVICIO. ....	57
<b>Figura N° 26:</b> PENTAHO - Carga de Dimensión FECHA. ....	57
<b>Figura N° 27:</b> PENTAHO - Carga de Dimensión PACIENTE. ....	58
<b>Figura N° 28:</b> PENTAHO - Carga de Dimensión ESPECIALIDAD. ....	59
<b>Figura N° 29:</b> PENTAHO - Carga de Hechos PRODUCTIVIDAD. ....	60
<b>Figura N° 30:</b> PENTAHO - Carga de Hechos CANTIDAD. ....	60
<b>Figura N° 31:</b> PENTAHO - Carga de Hechos CUMPLIMIENTO. ....	61
<b>Figura N° 32:</b> PENTAHO - Carga de Hechos PROMEDIO. ....	62

<b>Figura N° 33:</b> PENTAHO - Trabajo ETL de Actualización.....	63
<b>Figura N° 34:</b> Cubo Multidimensional. ....	64
<b>Figura N° 35:</b> Cubo PRODUCTIVIDAD, indicadores.....	65
<b>Figura N° 36:</b> Cubo CANTIDAD, indicadores.....	66
<b>Figura N° 37:</b> Cubo CUMPLIMIENTO, indicadores.....	67
<b>Figura N° 38:</b> Cubo PROMEDIO, indicadores.....	67
<b>Figura N° 39:</b> Cubo CUMPLIMIENTO, atributos. ....	69
<b>Figura N° 40:</b> Cubo PRODUCTIVIDAD, atributos. ....	70
<b>Figura N° 41:</b> Cubo CANTIDAD, atributos. ....	70
<b>Figura N° 42:</b> Cubo CANTIDAD, atributos. ....	71
<b>Figura N° 43:</b> Modelo Conceptual del Hospital La Caleta. ....	73
<b>Figura N° 44:</b> PENTAHO – Trabajo de Carga de Dimensiones.....	74
<b>Figura N° 45:</b> PENTAHO – Trabajo de Carga de Hechos. ....	74
<b>Figura N° 46:</b> Modelo Entidad Relación Físico, obtenido como resultado del Modelo Lógico del DW del Hospital La Caleta. ....	75
<b>Figura N° 47:</b> La tabla de productividad hora medico.....	76
<b>Figura N° 48:</b> La gráfica de productividad hora médico. ....	77
<b>Figura N° 49:</b> La tabla de total de pacientes atendidos por servicio.....	77
<b>Figura N° 50:</b> La gráfica de total de pacientes atendidos por servicio.....	78
<b>Figura N° 51:</b> La tabla de pacientes atendidos detallado por servicio y mes.....	78
<b>Figura N° 52:</b> La gráfica de pacientes atendidos detallado por servicio y mes. ....	79
<b>Figura N° 53:</b> La tabla de concentración de consultas.....	80
<b>Figura N° 54:</b> La gráfica de concentración de consultas. ....	81
<b>Figura N° 55:</b> La tabla de concentración de consultas de los 10 primeros pacientes. ....	81
<b>Figura N° 56:</b> La gráfica de concentración de consultas de los 10 primeros pacientes. ....	82
<b>Figura N° 57:</b> La tabla de grado de cumplimiento de la programación por servicio. ....	82
<b>Figura N° 58:</b> La gráfica grado de cumplimiento de la programación por servicio.....	83
<b>Figura N° 59:</b> La tabla de grado de cumplimiento de la programación por especialidad . ....	83
<b>Figura N° 60:</b> La gráfica grado de cumplimiento de la programación por especialidad . ....	84
<b>Figura N° 61:</b> Zona de Aceptación y Rechazo – Nivel de Aceptación.....	95
<b>Figura N° 62:</b> Zona de Aceptación y Rechazo – Calidad de la Información.....	102
<b>Figura N° 63:</b> Zona de Aceptación y Rechazo – Tiempo para generar los reportes. ....	108
<b>Figura N° 64:</b> Zona de Aceptación y Rechazo – Costo de generación de reportes.....	114
<b>Figura N° 65:</b> Zona de Aceptación y Rechazo – Tiempo para analizar información y formular alternativas. ....	120

<b>Figura N° 66:</b> Gráfico de Resultados Indicador Nivel de Satisfacción del Usuario. ....	122
<b>Figura N° 67:</b> Gráfico de Resultados del Indicador Calidad de la Información. ....	123
<b>Figura N° 68:</b> Gráfico de Resultados del Indicador Tiempo de generación de reportes que soportan la toma de decisiones. ....	125
<b>Figura N° 69:</b> Gráfico de Resultados del Indicador Costo de generación de reportes.....	126
<b>Figura N° 70:</b> Gráfico de Resultados del Indicador Tiempo para analizar información y formular alternativas. ....	127
<b>Figura N° 71:</b> Pantalla de Inicio y Conexión al Servidor de SQL SERVER 2008. ....	134
<b>Figura N° 72:</b> Explorador de Objetos SQL SERVER 2008.....	134
<b>Figura N° 73:</b> Opción de Nueva Base de datos.....	135
<b>Figura N° 74:</b> Opciones para restaurar la base de datos. ....	135
<b>Figura N° 75:</b> Restaurar base de datos -SIGH. ....	136
<b>Figura N° 76:</b> Especificación de copia de seguridad. ....	136
<b>Figura N° 77:</b> Datos Generales de la restauración. ....	137
<b>Figura N° 78:</b> Base de datos SIGH restaurada.....	137
<b>Figura N° 79:</b> Creación del Data WareHouse BI-SIGH. ....	138
<b>Figura N° 80:</b> Carpeta Utilitarios Pentaho .....	139
<b>Figura N° 81:</b> Pentaho Data Integration.....	139
<b>Figura N° 82:</b> Configuración Conexión Base de Datos .....	140
<b>Figura N° 83:</b> Cuestionario de Nivel de Satisfacción - Sistema Actual.....	141
<b>Figura N° 84:</b> Cuestionario de Nivel de Satisfacción - Solución BI.....	142
<b>Figura N° 85:</b> Cuestionario de Calidad de Información - Sistema Actual.....	143
<b>Figura N° 86:</b> Cuestionario de Calidad de Información - Solución BI.....	144

## **RESUMEN**

Se presenta una Solución de Inteligencia de Negocios, que permitió agilizar la toma de decisiones en la Dirección Ejecutiva del Hospital la Caleta, a través de herramientas y tecnologías de inteligencia de negocios que convirtió los datos en información útil. Usamos la metodología Hefesto, así como de la plataforma Pentaho, la cual es una herramienta Open Source que permitirá gestionar los procesos ETL (extracción, transformación y carga de datos), informes interactivos y análisis multidimensionales de información (OLAP). Como resultado se obtuvo información relevante en tiempo real y de forma rápida para la toma de decisiones en la dirección ejecutiva, quienes pudieron planificar, prever y proyectarse a cumplir las metas con una mejor calidad de gestión administrativa.

Autores: Bach. Alva Quiliche Miguel Angello

Bach. Callan Carbajal Milagros Roxana

Asesor : Ms. Apestegui Florentino Yim Isaias

## **ABSTRACT**

A Business Intelligence Solution is presented, which made it possible to streamline decision making in the Executive Direction of Hospital la Caleta, through business intelligence tools and technologies that turned the data into useful information. We use the Hephaestus methodology, as well as the Pentaho platform, which is an Open Source tool that will allow the management of ETL processes, interactive reports and multidimensional information analysis (OLAP). As a result, relevant information was obtained in real time and quickly for decision making in the executive direction, who were able to plan, forecast and project to meet the goals with a better quality of administrative management.

## **PRESENTACIÓN**

Señores miembros del Jurado Evaluador:

En cumplimiento a lo dispuesto por el Reglamento General de Grados y Títulos de la Universidad Nacional del Santa, ponemos a vuestra consideración el presente informe de tesis intitulado: “Solución de Inteligencia de Negocios para Mejorar la Toma de Decisiones en la Dirección Ejecutiva del Hospital La Caleta” requisito para optar el Título Profesional de Ingeniero de Sistemas e Informática.

La presente investigación tiene como lugar de aplicación al Hospital La Caleta, y tiene como propósito mejorar la toma de decisiones en la Dirección Ejecutiva, logrando tener un acceso fácil y rápido a información útil y valiosa de la institución, para dar soluciones operativas, tácticas y estratégicas de mejora, basados en el análisis de esta información.

Por lo todo lo expuesto, a ustedes señores miembros del jurado evaluador ponemos a su disposición nuestro informe, para su revisión, esperando cumpla con los requisitos mínimos para su aprobación.

Atentamente,

Los autores.

# INTRODUCCIÓN

En la actualidad, las Tecnologías de la Información adquieren cada vez más importancia dentro de las instituciones. Debido a que son herramientas de apoyo clave en los principales procesos e indispensables para el crecimiento y la competitividad. El desarrollo de los sistemas de información en las instituciones genera que los directivos pueden acceder a mucha más información, de más calidad y con mayor rapidez. El potencial que ello ofrece para mejorar la toma de decisiones y para guiar a las organizaciones hacia la consecución de sus objetivos es enorme. Sin embargo, muchos directivos se enfrentan a la paradoja de que “cada vez tienen más información y menos tiempo para analizarla” (Oltra Badenes, R. ,2012).

La inteligencia de negocios es una solución que permite a las instituciones acceder a información que les permite tomar decisiones, esta información generalmente proviene de diversas fuentes por lo que la solución de inteligencia de negocios se encarga de integrarla y transformarla para almacenarla en una estructura especial conocida como Data Warehouse. El Data Warehouse cumple con ciertas características para su construcción contando principalmente con 2 tipos de elementos, las tablas de hechos y las dimensiones. Este tipo de modelado nos facilita el usar herramientas multidimensionales que nos permiten hacer el cruce de los hechos con diversas dimensiones que les den un sentido real para el negocio. El resultado de una solución de inteligencia de negocios generalmente lo vemos de manera visible por medio de herramientas web.

El presente informe de Tesis tiene por objetivo una Solución de Inteligencia de Negocios para mejorar la toma de decisiones en la Dirección Ejecutiva del Hospital La Caleta; usando la plataforma Business Intelligence Open Source Pentaho y siguiendo la metodología Hefesto.

**En el Capítulo I**, Se presenta un estudio de la problemática existente. Luego se formula el problema, se justifica la investigación y posteriormente se definen los objetivos del informe.

**En el Capítulo II**, se describe la teoría usada para la comprensión e interpretación de los diversos conceptos aplicados durante el desarrollo del informe.

**En el Capítulo III**, en este capítulo se describe la hipótesis, variables e indicadores, los métodos de investigación que fueron utilizados para comprobar la veracidad de las hipótesis planteada en el proyecto.

**En el Capítulo IV**, se presentan los resultados de la investigación y se realiza la discusión de los resultados con el marco teórico, antecedentes y los datos empíricos en la investigación.

**En el Capítulo V**, se detalla las conclusiones del trabajo así como las recomendaciones del caso.

Además se incluye la bibliografía utilizada para el correcto desarrollo del proyecto, así como los anexos y los índices de figuras y tablas.

**Finalmente en el Anexo**, se explica el procedimiento para identificar perspectivas de análisis e indicadores de rendimiento, y los procesos ETL (Extracción, Transformación y Carga) que se utilizaron para poblar el almacén de datos temporal y los modelos dimensionales.

Finalmente, se realizará el modelamiento de cubos OLAP (Pentaho Schema Worbench), reportes mediante el módulo Pentaho Report Designer, para mostrarlos al usuario final a través del módulo Pentaho BI Server.

# CAPITULO I

## PROBLEMA DE INVESTIGACIÓN

### 1.1 Planteamiento y fundamentación del problema de investigación

#### 1.1.1. Planteamiento del Problema.

Los hospitales en el Perú reciben una gran demanda de servicios, muchas veces no satisfecha en su totalidad, por parte de los sectores menos favorecidos de la población peruana. Los sectores pobres del país acuden a los centros hospitalarios de atención pública como única alternativa, ya que no cuentan con recursos suficientes para acceder a un seguro privado. La falta de hospitales en el país y la centralización de los mismos en la capital peruana, sumado al crecimiento urbano, hacen necesario la optimización de los procesos administrativos de los hospitales. Esta optimización de los procesos administrativos debe considerarse como parte del proceso de perfeccionamiento de los servicios que ofrecen los centros de atención asistencial, orientado a satisfacer la demanda y mejorar la calidad de los servicios hospitalarios. El proceso de perfeccionamiento es iniciado por cada hospital, ya que son conscientes que para lograr la equidad y eficiencia en la atención al público deben mejorar los procesos administrativos.

Decisiones médicas oportunas y acertadas requieren de un soporte administrativo de calidad. Los sistemas de información creados pensando en las necesidades y características particulares del usuario contribuyen a que este proceso de toma de decisiones, muchas veces ligado a salvar vidas humanas, utilice la mejor y mayor información disponible. Es, por esto, importante resaltar la necesidad de contar con sistemas de información creados especialmente para los sistemas administrativos de salud pública, considerando todos los detalles que servirán de apoyo

durante el proceso de toma de decisiones a distintos niveles de la administración hospitalaria. En contraste, parece imposible crear un único sistema de información que resuelva la problemática entera de un hospital, es decir, que integre cada sub-proceso, por lo que debe considerarse que existen diversas fuentes de información distintas, de cada uno de los diferentes sistemas, con diferentes estándares o formas de representar los datos.

Actualmente las actividades de internamiento hospitalario absorben una gran proporción de recursos, y el primer nivel de atención en salud que recibe el paciente, es débil y requiere fortalecerse. Es importante una adecuada gestión de recursos en la entidad de salud. Hay varias direcciones hacia las que se pueden dirigir los recursos, como tratamientos, equipos, personal, infraestructuras, investigación y capacitación. Es por esto importante encontrar un balance y una proporción adecuada entre los recursos dirigidos a cada uno de estos puntos. Esto beneficiaría tanto a pacientes como al personal de la entidad de salud, y crearía una mejor equidad, lo que terminaría en una mejora en la calidad general del servicio brindado por el hospital.

Es por todo lo antes mencionado, que la dirección ejecutiva necesita tomar decisiones oportunas y acertadas, ya que la toma de decisiones en el Hospital La Caleta es vital para una mejora significativa en la atención a las personas que acuden a este centro de salud, que muchas veces son personas provenientes de provincia o de lugares alejados al distrito donde se encuentra el hospital, que esperan una mejor atención a comparación de su centro de salud asignado, es decir, de su distrito o provincia de residencia. Las decisiones tomadas por el directorio se ven reflejadas en: la atención que recibe el paciente; el tiempo que demora en ser atendido un paciente; adquisición de tecnología de última generación; políticas de crecimiento y expansión de servicios.

### **1.1.2. Fundamentación del problema.**

Uno de los principales problemas del Hospital La Caleta es que el directorio está conformado, en su mayoría, por médicos no especializados en labores administrativas. Es por esta razón que se debe de aprovechar las herramientas informáticas para que el directorio pueda tomar decisiones correctas y pertinentes basándose en información acertada y veraz de la situación dentro de la institución.

Actualmente, generar un reporte demora días e incluso semanas, ya que se tiene que cumplir obligatoriamente una serie de procesos administrativos para elaborarlo. En cambio con herramientas tecnológicas, estos reportes son generados, en forma personalizada, en minutos o incluso segundos.

Los Sistemas de Soporte de Decisiones (DSS) proveen información adecuada para la toma de decisiones, ya que adquieren, manipulan, transforman, validan y presentan datos almacenados en una forma resumida y bastante visual.

El Hospital La Caleta tiene los datos almacenados, pero no son capaces de aprovecharlos, ya que no generan ninguna información con ellos; por esto, es necesario crear un repositorio central de datos, que integre todos estos datos de los diferentes departamentos de servicios del hospital, les aplique un control de calidad y los almacene con un estándar. Con los DSS se puede utilizar los datos para transformarlos en información, y así conseguir que el directorio pueda planificar, preveer y proyectarse a cumplir las metas con una mejor calidad de gestión administrativa, por medio de reportes, como por ejemplo, Balance Scorecard o reportes presupuestales estadísticas y datos epidemiológicos de dicho nosocomio.

Cabe resaltar que los Sistemas de Soporte de Decisiones usan tecnologías informáticas, documentos de datos, conocimiento y modelos analíticos para identificar y solucionar problemas, ya que se brinda un sistema de fácil manejo en el que se puede hacer comparaciones cuantitativas con años, meses, semanas y días anteriores, para que el usuario pueda ser capaz de optimizar el proceso de toma de decisiones.

Es por ello, que en esta institución se solicitan al área de informática y estadística reportes para medir sus indicadores de servicios prestados por parte del cuerpo médico, lo cual genera problemas ya que el tiempo promedio en la elaboración de distintos reportes es de 2 a 3 días dependiendo del tipo de información requerida.

Por lo mencionado anteriormente podemos determinar los siguientes problemas:

- Retardo en la generación de reportes e informes administrativos.
- Disconformidad del personal administrativo por la demora en la elaboración de los reportes.
- Pérdida de tiempo del personal operativo para realizar los reportes estadísticos.
- Búsqueda de información lenta y desorganizada.
- Falta de reportes personalizados.

## 1.2 Antecedentes de la investigación

### A nivel internacional:

**Puenayán y Aynaguano (2012)** realizó la investigación: *Estudio comparativo de ETLs Software propietario vs Software libre para la implementación de una Solución Business Intelligence*. en la Escuela Superior Politécnica de Chimborazo, Ecuador. La investigación llegó a las siguientes conclusiones:

- Las herramientas SSIS(Software Propietario SQL Server 2008 Integration de servicios) y PDI(Software Libre Pentaho Data Integration) resultaron las mejoras en cuanto a soporte de fuente de datos, es decir tiene una conectividad sin ningún problema con todas las fuentes de datos antes indicadas: por esta razón, alcanzaron el puntaje máximo que es de 4. La herramienta TOS(Talent Open Studio) tuvo dificultades con los archivos dbf, por esta razón obtuvo un puntaje de 3.52.
- En el proyecto se ha utilizado la metodología HEFESTO facilitando la construcción del data warehouse, la misma que cuenta con fases flexibles que permiten desarrollar paso a paso un proyecto.

En este estudio comparativo de ETLs Software propietario vs Software libre Puenayán y Aynaguano (2012) demostraron que la mejor y más eficiente herramienta cuando se trata de una solución de business intelligence es SQL Server 2008 por esta razón haremos uso de esta herramienta; ya que esta es de mucha utilidad para nuestra investigación, porque ha sido demostrado que tiene un buen rendimiento en la carga de datos, velocidad de transmisión así como en la integración de los datos a utilizar por la Dirección Ejecutiva del Hospital la Caleta.

**Haro y Pérez (2014)**, realizó la investigación: *Datawarehouse para el Centro de Documentación Regional “Jun Bautista Vásquez”*, en la Universidad de Cuenca, México. La investigación llegó a las siguientes conclusiones:

- Hefesto es una metodología que proporciona los pasos necesarios para la implementación de un Data Warehouse de forma eficiente, existe mucha documentación al respecto. Además, se considera importante el hecho de que se parte de los requerimientos de los usuarios permitiendo así que los resultados persigan los objetivos del centro documental.
- La suite de Pentaho facilita herramientas necesarias para la implementación de un Data Warehouse y ayuda en cada etapa de la creación.
- El uso de la herramienta BI Server para realizar los reportes es muy intuitiva y permite que fácilmente los usuarios generen sus reportes de acuerdo a las necesidades que aparezcan.

Haro y Pérez (2014) demostraron que Pentaho es una herramienta de gran utilidad que se utiliza para el desarrollo del Data Warehouse, debido a la facilidad de uso, arquitectura flexible y de calidad, madurez de la tecnología, integración rápida con la infraestructura existente y bajo costo. Así como que la Metodología Hefesto permite la construcción de Data Warehouse de forma sencilla, ordenada e intuitiva.

## **A nivel nacional:**

**Gonzales (2014).** *Impacto de la Datawarehouse e Inteligencia de negocios en el desempeño de las empresas: Investigación empírica en Perú, como país en vías de desarrollo.* Universidad ESAN, Perú. La investigación llegó a las siguientes conclusiones:

Los principales constructos y sus componentes, que impactan en el desempeño de la DW y BI, de acuerdo a relevancia, son:

- Calidad de la Información (Información Actualizada, Fuente de Datos Adecuada, e Calidad de la Información en sí misma), la cual era mencionada por 9 de cada 10 entrevistados.
- Uso del Sistema (Uso), la cual era mencionada por 4 de cada 10 entrevistados.
- Calidad del Sistema (Herramientas, ETL de los datos, y Velocidad de Respuesta del Sistema), la cual era mencionada por 4 de cada 10 entrevistados.
- Calidad del Servicio (Entrenamiento del Usuario y Entrenamiento del personal de Tecnología de Información), la cual era mencionada por 3 de cada 10 entrevistados.
- Satisfacción del Usuario (Satisfacción del Usuario), la cual era mencionada por 3 de cada 10 entrevistados.

Complementariamente se determinaron otros factores relevantes a DW y BI, como son los factores de implementación, factores estratégicos, relación beneficio-costos y recursos financieros de la empresa.

También se establecieron las principales aplicaciones que le dan las empresas a la Inteligencia de Negocios: desarrollo de modelos predictivos, seguimiento de campañas comerciales, segmentación de clientes, análisis del retorno de la

inversión y resultados financieros, análisis de ventas, asignación de la parte variable de la remuneración de los empleados, mejor conocimiento de los clientes, análisis de rentabilidad de los clientes, análisis de riesgo, y otras. De esta investigación se tomarán en cuenta conceptos y definiciones que hacen referencia a la inteligencia de negocios para complementar y enriquecer nuestro marco teórico que servirá de base y guía para el desarrollo del presente proyecto de tesis.

**Gallo y Paz (2013).***Construcción de un Datawarehouse para mejorar la toma de decisiones en la gestión administrativa de la UBAP Metropolitano de Chimbote.*

La investigación llegó a las siguientes conclusiones:

- En el nivel de satisfacción del usuario en una escala valorada de 1 a 5 con el sistema actual es de 2.08 (41.60%) y cuando se implementó el Datawarehouse, las nuevas encuestas arrojaron un nivel de satisfacción del usuario del sistema 4.65 (93%). Lo cual comprende en un incremento del 2.57 (51.40%) en los usuarios con el nuevo sistema de trabajo.
- El tiempo promedio de elaboración de reportes semanales con el sistema actual era de 6002 minutos (100%) y con la implementación del datawarehouse el nuevo tiempo de elaboración de reportes fue de 904 minutos (15.06%), por lo que se mostró una reducción de 5098 minutos (84.94%) tiempo restante que se aprovecharía en otras actividades.
- En la cantidad de dinero que se pierde por la elaboración de reportes con el sistema actual se halló un total de 750.25 (100%) soles. Con la implementación del datawarehouse se obtuvo como resultado la cantidad de 113(15.6%) soles, logrando una disminución de 637.25 (84.94%), invirtiendo más en la calidad de atención al paciente.

Gallo y Paz (2013) demostraron que la construcción de un Data Warehouse mejoró la toma de decisiones en la gestión administrativa de la UBAP Metropolitano de Chimbote cumpliendo de esta manera los objetivos planteados de la investigación mejorando el nivel de satisfacción de los usuarios, reduciendo tiempos en la elaboración y búsquedas de reportes, y optimizando costos.

### **1.3 Formulación del problema de investigación**

¿En qué medida la solución de inteligencia de negocios mejorará la toma de decisiones en la Dirección Ejecutiva del Hospital la Caleta?

### **1.4 Delimitación del estudio**

El estudio está orientado a la Solución de Inteligencia de Negocios para mejorar la toma de decisiones en la Dirección Ejecutiva del Hospital la Caleta.

### **1.5 Justificación e importancia de la investigación**

#### **Justificación teórica**

La investigación propuesta busca, mediante la aplicación de la teoría y los conceptos básicos de Inteligencia de negocios, encontrar explicación al proceso interactivo que explora y analiza la información estructurada sobre un área (normalmente almacenada en un datawarehouse), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones. Esto gracias a la capacidad de explotar su información, con la intención de poder manipularlos de una manera más sencilla y entender el porqué de nuestro desempeño o, mejor aún, plantear escenarios a futuro, lo cual nos ayudará a tomar mejores decisiones.

Como resultado de la solución de inteligencia de negocios las decisiones en la dirección ejecutiva van a estar mejor fundamentadas y habrá menos posibilidades de equivocación, lo cual es favorable para Hospital y por lo tanto se justifica plenamente la realización de esta investigación.

### **Justificación metodológica**

Para lograr el cumplimiento de los objetivos de estudio, se acude al empleo de la Metodología Hefesto, la cual es una metodología cuya propuesta se fundamenta en una amplia investigación, comparación de metodologías existentes y la experiencia en la elaboración de almacenes de datos. La ventaja principal de esta metodología es que especifica puntualmente los pasos a seguir en cada fase a diferencia de otras metodologías que mencionan los procesos, más no explican cómo realizarlos. Además está enfocada al análisis de requerimientos, identificación de indicadores y perspectivas, que permitirán realizar el análisis de datos respectivos y obtener resultados que apoyen a la toma de definiciones del Hospital.

### **Justificación práctica**

La solución de inteligencia de negocios se justifica en la práctica ya que le será de gran utilidad a la dirección ejecutiva mediante los siguientes beneficios:

- Disposición de la información correcta en el momento adecuado para la toma de decisiones.
- Preverá la capacidad para evaluar distintos escenarios al mismo tiempo, con lo cual se pueden analizar diferentes situaciones que pudiesen afectar al Hospital, y de esta manera poder adelantar las posibles decisiones estratégicas a tomar y hacer que una tendencia negativa se convierta en acciones positivas para el Hospital.
- Permitirá agrupar información de distintas áreas en un solo cuadro, lo cual es muy favorable para el cambio de políticas o reorientaciones de los planes establecidos.

## **Importancia**

En la actualidad el manejo de la información es vital para obtener una ventaja competitiva, y para ello, el director y su equipo técnico, encargados de la toma de decisiones, requieren un acceso fácil y rápido a información útil y valiosa de la institución para dar soluciones operativas, tácticas y estratégicas de mejora, basados en el análisis de esta información.

La implementación de una solución de Business Intelligence (BI) dentro del Hospital La Caleta, para la dirección ejecutiva, representa una estrategia corporativa de crecimiento, optimización, productividad y competitividad, que no solo les permitirá tener tiempos de respuesta menores en la obtención de datos procesados, sino el mejor manejo de los mismos, la toma de decisiones basados en datos precisos y de calidad para su análisis.

### **1.6 Objetivos de la investigación: general y específicos**

#### **1.6.1 Objetivo general.**

Mejorar la toma de decisiones en la Dirección Ejecutiva del Hospital La Caleta a través de una solución de inteligencia de negocios.

#### **1.6.2 Objetivos específicos.**

- Realizar el levantamiento de los requerimientos
- Analizar la fuente de datos del sistema operacional
- Diseñar la solución de inteligencia de negocios
- Implementar la solución de inteligencia de negocios
- Realizar Pruebas mediante solución de inteligencia de negocios

## **CAPITULO II**

### **MARCO TEÓRICO**

#### **2.1 Fundamentos teóricos de la investigación**

##### **2.1.1. Toma de Decisiones.**

La toma de decisiones empresariales constituye un proceso de selección entre las diferentes alternativas posibles. A partir de la elección de una alternativa se podrá pasar al proceso de la planificación de actividades, desarrollo de las estrategias y formulación de programas y presupuestos. Para dirigir eficazmente la pequeña y mediana empresa se necesita utilizar los recursos disponibles de una manera eficiente, al ser los recursos limitados los gerentes deberán tratar de aprovechar todas las oportunidades (Maqueda, 2010, p.21).

La toma de decisiones se debe realizar en los diferentes niveles de la empresa, esto trae consigo que las decisiones individuales, correspondientes a cada nivel de responsabilidad, afectan de una manera directa al desarrollo de todo el sistema organizativo y condicionan tanto la consecución de los objetivos particulares como la de los objetivos generales o globales. En la siguiente figura se observa las fases de la planificación de toma de decisiones en función del análisis del valor. (Maqueda, 2010, p.21-22).



**Figura N° 1** : El proceso de la toma de decisiones en la empresa.

**Fuente:** (Maqueda, 2010, p.22).

## **2.1.2. Inteligencia de Negocios.**

### **2.1.2.1. Definición.**

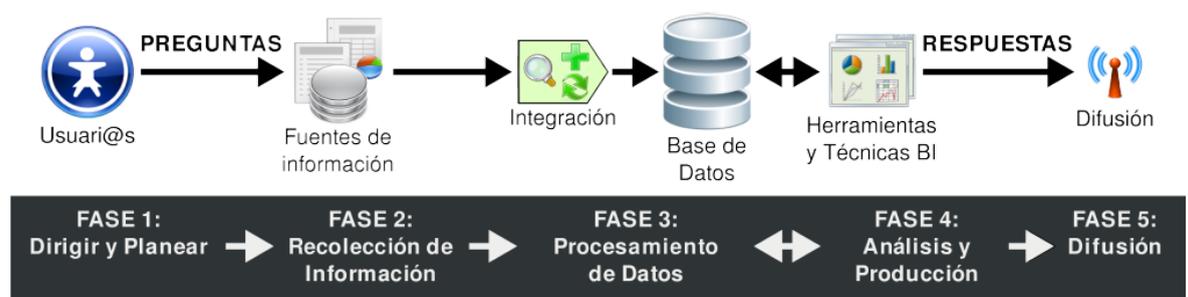
La inteligencia de negocios es un conjunto de tecnologías y procesos que permiten a personas de todos los niveles en una organización, tener acceso a datos y su análisis. (Howson, 2009, p.14).

Según el autor (Hatch, D., 2009) en el libro “Operational Business Intelligence What you Need to Know About Improving Your Customer-Facing Performance” menciona que la Inteligencia de Negocios es la combinación de prácticas, capacidades y tecnologías usadas por las compañías para recopilar e integrar la información, aplicar reglas del negocio y asegurar la visibilidad de la información en función de una mejor comprensión del mismo y, en última instancia, para mejorar el desempeño.

Según los autores (Sabherwal y Becerra, 2011, NJ: John Wiley & Sons, Inc., EUA) definen a la inteligencia de negocios como: la forma de brindarle a los decisores información y conocimientos valiosos, mediante el apalancamiento de una gama de fuentes de información, así como de la información estructurada y no estructurada.

### 2.1.2.2. *Proceso de Inteligencia de Negocios.*

Se describirá a continuación el proceso de BI. El mismo está dividido en cinco fases, las cuales serán explicadas teniendo como referencia el siguiente gráfico, que sintetiza todo el proceso:



**Figura N° 2:** Fases del proceso BI.

**Fuente:** Bernabeu, R. (2009). Proceso de BI. Recuperado de <http://www.dataprix.com/data-warehousing-y-metodologia-hefesto/1-business-intelligence/13-proceso-bi>

### 2.1.2.3. *Ventajas.*

Como señala (Howson, 2009, McGraw-Hill, España) "Las ventajas de utilizar inteligencia de negocios son":

- Reducción de costes y mayor eficiencia
- Más capacidad para tomar decisiones:
- Mejor capacidad de respuesta:
- Mejor visibilidad, mayor comprensión del negocio

### 2.1.3. ETL(Extracción, Transformación y Carga).

#### 2.1.3.1. *Definición.*

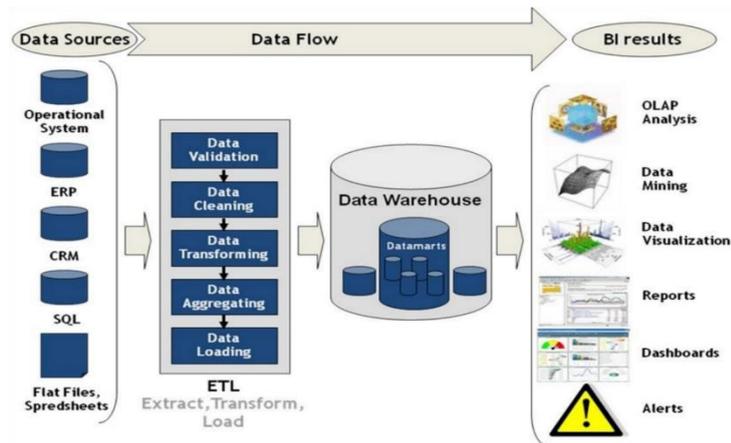
Según los autores (Kimball, R., y Ross, M., 2010) en el libro “The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence.” que **ETL** representan las siglas de “Extract, Transform and Load” que en idioma español significa: **Extraer, Transformar y Cargar**. ETL es el proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra **base de datos, data mart, o data warehouse**; para analizar, o en otro sistema operacional para apoyar un proceso de negocio

### 2.1.4. Data Warehouse (DWH).

#### 2.1.4.1. *Definición.*

Según los autores (Kimball, R., y Ross, M., 2010) en el libro “The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence.” sostiene que un *Data Warehouse* es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta.

Según el autor (Conesa y Curto, 2010) en el libro “Introducción al Business Intelligence” una data warehouse es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización independientemente de cómo se vayan a utilizar posteriormente por los consumidores o usuarios, con las propiedades siguientes: estable, coherente, fiable y con información histórica.



**Figura N° 3 :** Arquitectura Data Warehouse.

**Fuente:** (Kommineni Sivaganesh & Suresh Chandra, 2012).

#### 2.1.4.2. *Ventajas.*

Como señala (Kimball, R., y Ross, M., 2010) Las ventajas de usar un Data Warehouse en la empresa son las siguientes:

- Mejorar la Entrega de Información.
- Facilitar el Proceso de Toma de Decisiones.
- Impacto Positivo sobre los Procesos Empresariales

#### 2.1.5. **Proceso Analítico en Línea (OLAP).**

“OLAP es una tecnología que permite sacar provecho a como está estructurado un Data Mart, presentando textos y números bajo el concepto dimensional” (Salcedo Parra, Milena Galeano, & Rodriguez B, 2010, p.22).

Como señala Salcedo, Galeano, & Rodriguez (2010) “Presenta las siguientes características”:

- Esta optimizado para realizar consultas rápidas de los usuarios
- Poseen un motor robusto para realizar análisis numéricos

- Es un modelo de datos conceptual y altamente intuitivo.
- Proporciona una vista de datos multidimensional

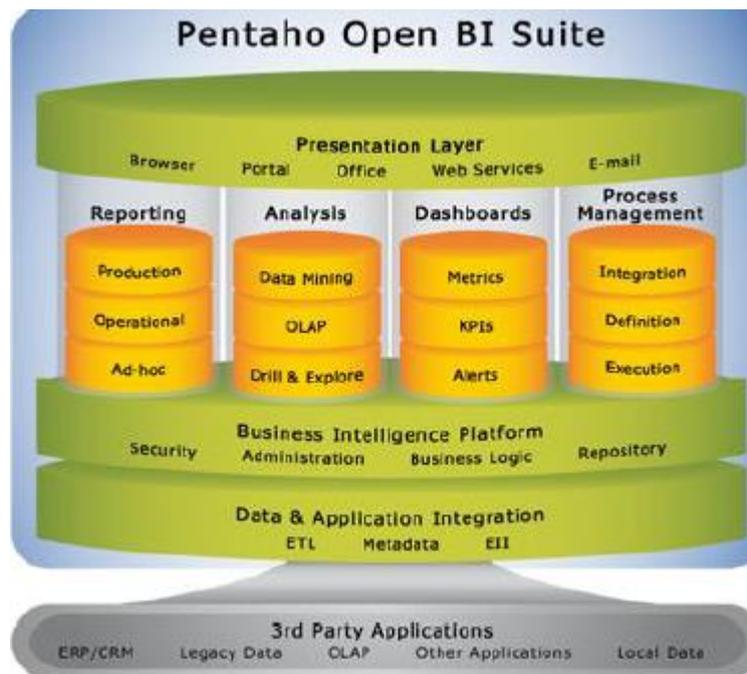
## 2.1.6. Pentaho.

### 2.1.6.1. *Definición.*

Según el autor (Conesa y Curto, 2010) en el libro “Introducción al Business Intelligence” el PENTAHO es una de las suites más completas y maduras del mercado OSBI que existe desde el año 2006. Existen dos versiones: Community y Enterprise y están compuestas por diferentes motores incluidos en el servidor de Pentaho:

- Pentaho Reporting: soporta informes estáticos, paramétricos y ad hoc.
- Pentaho Análisis: soporta OLAP (mediante Mondrian) y minería de datos (mediante Weka).
- Cuadros de mando: mediante CDF (Community Dashboard Framework).
- ETL: mediante la herramienta Kettle.
- Metadata: que proporciona una capa de acceso de información basada en lenguaje de negocio.
- Workflow: el servidor de Pentaho se basa en acciones que la mayoría de objetos de negocio permite lanzar.

Pentaho, es un proyecto iniciado por una comunidad Open Source, provee una alternativa de soluciones de BI en distintas áreas como en la arquitectura, soporte, funcionalidad e implantación (Gravitar, 2016).



**Figura N° 4:** Plataforma Pentaho.

**Fuente:** (Gravitar, 2016).

### 2.1.6.2. *Herramientas y Áreas de Aplicación.*

Bajo la integración de otros proyectos Open Source que brindan funcionalidad en áreas de BI, la comunidad Pentaho trabaja en formalizar estas herramientas y formar el Suite BI (Gravitar, 2016).

#### 2.1.6.2.1. *Pentaho Data Integration.*

Pentaho Data Integration abre, limpia e integra esta valiosa información y la pone en manos del usuario. Provee una consistencia, una sola versión de todos los recursos de información, que es uno de los más grandes desafíos para las organizaciones TI hoy en día. Pentaho Data Integration permite una poderosa ETL (Extracción, Transformación y Carga) (Gravitar, 2016).

#### **2.1.6.2.2.** *Pentaho Schema Workbench.*

Pentaho Schema Workbench, es una herramienta de desarrollo que permite crear, modificar y publicar un esquema Mondrian. Es un programa java multiplataforma. Permite crear todos los objetos que soporta Modrian: esquema, cubo, dimensiones, métricas, etc... (Conesa y Curto, 2010, p.106)

#### **2.1.6.2.3.** *Pentaho Analysis.*

Suministra a los usuarios un sistema avanzado de análisis de información. Con uso de las tablas dinámicas, el usuario puede navegar por los datos, ajustando los filtros de visualización, añadiendo o quitando los campos de agregación. Los datos pueden ser representados en una forma de SVG o Flash, los dashboards widgets, o también integrados con los sistemas de minería de datos y los portales web (Conesa y Curto, 2010, p.109)

## 2.2 Marco conceptual

- **Toma de Decisiones:** en una definición simple podemos decir que decidir es simplemente escoger, entre las alternativas, la mejor. La decisión puede ser en situación de certidumbre o de incertidumbre. Por lo general, cuando existe una situación para la toma de decisiones, el gerente se encuentra entre diferentes opciones de acción, y usualmente cada una de estas opciones está asociada a cierta incertidumbre. (*Namakforoosh, 2005*)
  
- **Inteligencia de Negocios:** “se entiende como BI al conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas en la creación administración de la información que permite tomar mejores decisiones a los usuarios de una organización.” (*Conesa y Curto, 2010, p.. 18* )
  
- **Data Warehouse:** “es un repositorio de datos que proporciona una visión global, común e integrada de los datos de la organización - independientemente de cómo se vayan a utilizar posteriormente por los consumidores o usuarios, con las propiedades siguientes: estable coherente, fiable y con información histórica.” (*Conesa y Curto , 2010, p. 30*)
  
- **Data Marts:** Un Datamart es una base de datos departamental, especializada en el almacenamiento de los datos de un área de negocio específica. Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento Sinnexus (2007, 2016).
  
- **ETL (Extracción, Transformación y Carga):** Según los autores (Larissa T. Moss & Shaku Atre, 2003) en el libro “Business Intelligence Roadmap: The Complete Project Lifecycle for Decision Support Applications” Los datos de origen para las aplicaciones de BI provendrán de una variedad de plataformas, que son administradas por una variedad de sistemas operativos y aplicaciones. El propósito del proceso ETL es fusionar los datos de estas plataformas heterogéneas en un formato estándar para las bases de datos de BI en el entorno de soporte de decisiones de BI

- **OLAP(Procesamiento Analítico en Línea ):** es una nueva forma de presentar la información. Con este, los datos se ven en cubos con ejes, dimensiones/medidas, rebanadas y niveles. Los ejes se refieren a la estructura física de la presentación, como renglones y columnas. Las dimensiones son características de los datos que se colocan en los ejes. Las medidas son los valores de los datos a ser desplegados. Las rebanadas son los atributos del cubo (dimensiones o medidas) que se mantendrán constantes en la presentación. El nivel es un atributo de una dimensión que describe su posición en una jerarquía.

“El termino cubo se utiliza tanto para referirse a la estructura semántica implícita que se usa para interpretar datos, como para la materialización particular de datos en esa estructura semántica.” (Kroenke, 2003, p. 547).

Consiste en hechos numéricos, llamadas medidas, a partir de tablas de hechos. OLAP proporciona análisis interactivo, multi-dimensional con diferentes dimensiones y diferentes niveles de detalle ( *Technics Publications LLC, 2010* )

- **Pentaho:** Conesa y Curto (2010, p.26) define Pentaho como “una de las suites más complejas y maduras del mercado OSBI que existe desde el 2006. Existen dos versiones: Community y Enterprise. Está compuesta por diferentes motores incluidos en el Servidor de Pentaho:”

## CAPITULO III

### MARCO METODOLÓGICO

#### 3.1 Hipótesis central de la investigación

La Solución de Inteligencia de Negocios mejora la toma de decisiones en la Dirección Ejecutiva del Hospital La Caleta.

$H_0$ , La Solución de Inteligencia de Negocios no mejora la toma de decisiones en la Dirección Ejecutiva del Hospital La Caleta.

$H_1$ , La Solución de Inteligencia de Negocios mejora la toma de decisiones en la Dirección Ejecutiva del Hospital La Caleta.

#### 3.2 Variables e indicadores de la investigación

##### 3.2.1. Variables.

**Variable independiente:** Inteligencia de negocios

**Variable dependiente:** Mejora la Toma de decisiones

##### 3.2.2. Definición operacional.

**Tabla N° 1:** Matriz de Operacional de variables.

Variable	Definición conceptual	Definición operacional	Indicadores	Unidad de Medida
Inteligencia de negocios	“Conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada en información estructurada, para su explotación directa o para su análisis y conversión en conocimiento, dando así soporte a la	Actúa como un factor estratégico para la dirección ejecutiva, diseñado para realizar consultas de manera fácil y personalizada, apoyando el proceso de toma de decisiones.	Calidad de Información	Escala de Likert (1-5)
			Tiempo de generación de reportes que soportan la toma de decisiones.	Minutos

	toma de decisiones sobre el negocio” (Howson, 2009, p.14).		Costo por la generación de reportes	Soles
Mejora la Toma de decisiones	“La toma de decisiones empresariales constituye un proceso de selección entre las diferentes alternativas posibles” (Maqueda, 2010, p.21).	Es un proceso racional que se da al presentarse distintas alternativas, analizando seleccionando e implementando la que más se adecue al objetivo estratégico, obteniendo así una ventaja competitiva.	Nivel de satisfacción de los usuarios.	Escala de Likert (1-5)
			Tiempo para analizar reportes y formular alternativas	Minutos

Indicadores	Descripción	Objetivo	Técnica/ Instrumento	Modo de Calculo
Nivel de satisfacción de los usuarios. <i>NSU</i>	Determina el nivel de satisfacción de los usuarios internos.	Incrementar el nivel de satisfacción en el proceso de toma de decisiones de la dirección ejecutiva del Hospital la Caleta.	Encuesta	$NSU = \sum_{j=1}^5 (F_{ij} \times P_j)$ <p>F = Frecuencia P= Peso n = usuarios del sistema</p>
Calidad de Información <i>CI</i>	Determina la calidad de la información que se entrega a los usuarios.	Incrementar calidad de la información que se entrega a la dirección ejecutiva del Hospital.	Encuesta	$CI = \sum_{j=1}^5 (F_{ij} \times P_j)$ <p>F = Frecuencia P= Peso n = usuarios del sistema</p>
Tiempo de generación de reportes que soportan la toma de decisiones. <i>TGR</i>	Determina el tiempo promedio que se demora el usuario en generar reportes.	Disminuir el tiempo de generación de reportes que soportan la toma de decisiones	Observación/Ficha de recolección de datos.	$\overline{TGR} = \frac{\sum_{i=1}^n TGR_i}{n}$ <p>n = número de tiempos.</p>

Costo por la generación de reportes <i>CGR</i>	Determina el costo promedio que cuesta generar los reportes.	Disminuir el costo por la generación de reportes	Observación/Ficha de recolección de datos.	$\overline{CGR} = \frac{\sum_{i=1}^n CGR_i}{n}$ n = número de costos.
Tiempo para analizar reportes y formular alternativas <i>TAIFA</i>	Determina el tiempo promedio que se demora el usuario en analizar los reportes y formular alternativas,	Disminuir el tiempo para analizar reportes y formular alternativas	Observación/Ficha de recolección de datos.	$\overline{TAIFA} = \frac{\sum_{i=1}^n TAIFA_i}{n}$ n = número de tiempos.

**Tabla N° 2:** Indicadores.

### 3.3 Métodos de la investigación

El método de investigación es Inductivo – Deductivo porque luego de definir la realidad problemática y, por ser una investigación aplicada, planteamos una hipótesis que permitirá hacer pruebas a los indicadores o características de la variable dependiente para observar o percibir si las consecuencias de la hipótesis son viables o verificados con los resultados de la propuesta de una solución de inteligencia de negocios que permita mejorar la toma de decisiones en la dirección ejecutiva del hospital la caleta.

### 3.4 Diseño o esquema de la investigación

El diseño de la investigación que aplicaremos será: El método de sucesión o en línea con un solo grupo, también denominado “Pre test-Pos test”.

El método a utilizar está contemplado por el tipo de investigación Cuasi Experimental, que nos sirve para comparar la situación anterior a la experimentación con los resultados obtenidos después de la misma.

Así tenemos:

- Una medición previa de la variable independiente a ser utilizada antes de la aplicación de la variable independiente (Pre test).

- La aplicación de la variable independiente a los sujetos de la muestra
- Una nueva medición de la variable dependiente después de la aplicación de la variable independiente(Pos test)

Grupo Único de Control:

$$G1 \quad O_1 \rightarrow X \rightarrow O_2$$

Donde:

$O_1$  = Situación previa a la Solución

$X_1$  = Solución de inteligencia de negocios

$O_2$  = Situación posterior a la Solución

### 3.5 Población y muestra

#### 3.2.3. Población.

En el presente trabajo de investigación, la población está compuesta por las áreas del hospital La Caleta, las cuales son: Administración, Planeamiento estratégico, Medicina, Cirugía, Pediatría, Ginecoobstetra, Odontoestomatología, Emergencia, Centro Quirúrgico, Patología, Diagnóstico por imágenes, Apoyo al tratamiento, Estadística e informática, Asesoría jurídica, Epidemiología y Gestión de la calidad.

Las cuales constan de los siguientes profesionales.

**Tabla N° 3:** Recursos Humanos del Hospital la Caleta.

<b>RECURSOS HUMANOS</b>	<b>CANTIDAD</b>
Directivos	06
Profesionales médicos	46
Profesionales no médicos	88
Otros médicos	05

Técnicos administrativos	52
Técnicos asistenciales	110
<b>Total</b>	307

Fuente: ROF- Unidad de Personal – Hospital La Caleta

La población para la evaluación de la solución de BI está representada por los registros de toma de tiempo, costos y los miembros de la Dirección ejecutiva del Hospital la Caleta.

### **Cálculo de Población**

**Indicador 1:** Nivel de satisfacción de los usuarios

$$N_1 = 6 \text{ Directivos}$$

$$N_1 = 6$$

**Indicador 2:** Calidad de Información

$$N_2 = 6 \text{ Directivos}$$

$$N_2 = 6$$

**Indicador 3:** Tiempo de generación de reportes que soportan la toma de decisiones

$$N_3 = \frac{4 \text{ tiempo}}{1 \text{ pruebas}} * 16 \text{ pruebas}$$

$$N_3 = 64 \text{ tiempo}$$

**Indicador 4:** Costo por la generación de reportes

$$N_4 = \frac{4 \text{ costo}}{1 \text{ pruebas}} * 16 \text{ pruebas}$$

$$N_4 = 64 \text{ costo}$$

**Indicador 5:** Tiempo para analizar reportes y formular alternativas

$$N_5 = \frac{1 \text{ tiempo}}{1 \text{ directivo}} * \frac{6 \text{ directivos}}{1 \text{ prueba}} * 6 \text{ pruebas}$$

$$N_5 = 36 \text{ tiempos}$$

### 3.2.4. Muestra.

Para la determinación de la muestra se empleara la técnica del muestreo aleatorio simple.

$$n = \frac{(N \cdot z^2 \cdot p \cdot q)}{[(N - 1) \cdot e + z^2 \cdot p \cdot q]}$$

Nota: El muestreo se aplica cuando se conoce la población (n) y si  $n > 80$  en caso contrario no se realiza muestreo, es decir si  $n \leq 80$ .

#### Cálculo de Muestra

**Indicador 1:** Nivel de satisfacción de los usuarios

La población  $N_1$  es de 6 directivos, siendo menor a 80 ( $N_1 < 80$ ), se asume la población como muestra.

$$n_1 = 6$$

**Indicador 2:** Calidad de Información

La población  $N_2$  es de 6 directivos, siendo menor a 80 ( $N_2 < 80$ ), se asume la población como muestra.

$$n_2 = 6$$

**Indicador 3:** Tiempo de generación de reportes que soportan la toma de decisiones

La población  $N_3$  es de 64 tiempos, siendo menor a 80 ( $N_3 < 80$ ), se asume la población como muestra.

$$n_3 = 64$$

**Indicador 4:** Costo por la generación de reportes

La población  $N_4$  es de 64 costos, siendo menor a 80 ( $N_4 < 80$ ), se asume la población como muestra.

$$n_4 = 64$$

**Indicador 5:** Tiempo para analizar reportes y formular alternativas

La población  $N_5$  es de 36 tiempos, siendo menor a 80 ( $N_5 < 80$ ), se asume la población como muestra.

$$n_5 = 36$$

### 3.6 Actividades del proceso investigativo

- Definición de Variables
- Hipótesis Estadísticas
- Nivel de Significancia
- Estadígrafo de contraste
- Resultados de la Estadística

### 3.7 Técnicas e instrumentos de la investigación

#### **Entrevista.**

Se aplicará la entrevista, para recopilar información en forma verbal, a través de preguntas que propondremos como entrevistadores. Quienes responderán son las personas involucradas en la toma de decisiones los cuales serán los usuarios de la solución de inteligencia de negocios.

#### **Encuesta.**

Con esta técnica de recolección de datos da lugar a establecer contacto con las unidades de observación por medio de los cuestionarios previamente establecidos (los cuestionarios consisten en un conjunto de preguntas respecto a una o más variables).

#### **Observación.**

Es el registro visual de lo que ocurre en las reuniones de toma de decisiones, clasificando consignando los acontecimientos pertinentes de acuerdo con algún esquema previsto y según el problema que se estudia.

### **3.8 Procedimiento para la recolección de datos**

- Se gestionará la carta de presentación a través de la Dirección de la Escuela de Sistemas a la Dirección del Hospital La Caleta para la ejecución del proyecto.
- Con el documento de autorización y de presentación respectiva al área de Informática y Estadística se procederá a realizar las entrevistas.
- Para el proceso de recolección de datos se identificarán a los usuarios involucrados en los procesos administrativos del Hospital La Caleta que reúnan los criterios de selección.
- Durante la entrevista se pretenderá crear un clima agradable de comodidad y sobre todo de confianza en la que las partes puedan expresarse libremente.

### **3.9 Técnicas de procesamiento y análisis de datos**

Para el procesamiento, análisis contrastación e interpretación de resultados de la investigación se aplicaran las técnicas estadísticas:

- Ordenamiento y codificación de datos
- Tabulación
- Tablas estadísticas
- Gráficos
- Análisis e interpretación

## CAPITULO IV

### RESULTADOS Y DISCUSIÓN

#### 4.1 Resultados de la inteligencia de negocios

Se desarrolla las Fases de la Inteligencia de Negocios:

##### 4.1.1 Dirigir y planear.

En esta etapa revisaremos todos los aspectos relacionados con nuestra investigación, comenzaremos por describir lo siguiente:

##### **¿Qué se quiere hacer?**

El presente trabajo de investigación tiene por finalidad analizar la Base de Datos del Hospital La Caleta, a fin de poder generar una solución de Inteligencia de Negocios que integra por un lado el almacenamiento y por el otro el procesamiento de grandes cantidades de datos, con el principal objetivo de transformarlos en conocimiento y en decisiones en tiempo real, a través de un sencillo análisis y exploración.

Reduciendo el tiempo mínimo que se requiere para recoger toda la información relevante de un tema en particular, ya que la misma se encontrará integrada en una fuente única de fácil acceso y los usuarios podrán consultar y analizar los datos de manera sencilla e intuitiva.

##### **¿Qué información necesita el Hospital La Caleta?**

El Hospital La Caleta como toda institución de Salud requiere información que les permita establecer mejoras en los procesos hospitalarios mediante la medición y análisis del desempeño a través de la relación cuantitativa entre dos variables que participan en un mismo proceso. El Hospital La Caleta diariamente almacena información de los

pacientes que atienden y en base a esta información podemos encontrar ciertos patrones que se repiten.

Aprovechemos esta información para alimentar este trabajo, de tal forma que podamos dar respuestas a las siguientes preguntas.

- ¿Cuál es la productividad hora médico por cada hora programada de trabajo del médico?
- ¿Cuál es la cantidad de pacientes atendidos por cada servicio en un período de tiempo?
- ¿Cuál es el promedio de consultas que se han otorgado a cada paciente que acudió al hospital durante un período de tiempo?
- ¿Cuál es el grado de cumplimiento de la programación de las horas médico por servicio o especialidad en un periodo de tiempo?
- ¿Cuál es el promedio de permanencia de los pacientes que egresaron en un periodo de tiempo?

#### **4.1.2 Recolección de información.**

**En cuanto a Software ¿Con que cuenta el Hospital La Caleta para realizar este trabajo?**

En esta etapa de la investigación debemos analizar los datos con los que se cuenta y los campos que se encuentran poblados a fin de poder determinar que se puede obtener de ellos.

Lo encontrado en la institución es lo siguiente:

- Nombre de la Base de Datos: SIGH\_EXTERNA
- Plataforma de la BD: SQL Server 2008
- Tamaño del archivo Log: 369 MB

## Tablas de la Base de Datos

De la copia de la Base de Datos seleccionaremos solamente las tablas que necesitaremos para realizar nuestra Solución de Inteligencia de Negocios.

A continuación se muestran las tablas que utilizaremos así como la estructura de cada una:

Atenciones			
	Nombre de columna	Tipo de datos	Permitir valores NULL
🔑	IdAtencion	int	<input type="checkbox"/>
	IdPaciente	int	<input type="checkbox"/>
	Edad	int	<input type="checkbox"/>
	FechaIngreso	datetime	<input type="checkbox"/>
	HoraIngreso	char(5)	<input type="checkbox"/>
	IdDestinoAtencion	int	<input checked="" type="checkbox"/>
	IdTipoCondicionAlServicio	int	<input checked="" type="checkbox"/>
	IdTipoCondicionALEstab	int	<input checked="" type="checkbox"/>
	IdServicioIngreso	int	<input type="checkbox"/>
	IdMedicoIngreso	int	<input type="checkbox"/>
	IdEspecialidadMedico	int	<input checked="" type="checkbox"/>
	IdMedicoEgreso	int	<input checked="" type="checkbox"/>
	FechaEgreso	datetime	<input checked="" type="checkbox"/>
	HoraEgreso	char(5)	<input checked="" type="checkbox"/>
	IdOrigenAtencion	int	<input checked="" type="checkbox"/>
	FechaEgresoAdministrativo	datetime	<input checked="" type="checkbox"/>
	HoraEgresoAdministrativo	char(5)	<input checked="" type="checkbox"/>
	IdCondicionAlta	int	<input checked="" type="checkbox"/>
	IdTipoAlta	int	<input checked="" type="checkbox"/>
	IdServicioEgreso	int	<input checked="" type="checkbox"/>
	IdCamaIngreso	int	<input checked="" type="checkbox"/>
	IdCamaEgreso	int	<input checked="" type="checkbox"/>
	IdTipoGravedad	int	<input checked="" type="checkbox"/>
	IdTipoEdad	int	<input checked="" type="checkbox"/>
	IdCuentaAtencion	int	<input type="checkbox"/>
	IdCuentaAtencion	int	<input type="checkbox"/>
	IdTipoServicio	int	<input checked="" type="checkbox"/>
	IdFormaPago	int	<input checked="" type="checkbox"/>
	idFuenteFinanciamiento	int	<input checked="" type="checkbox"/>
	idEstadoAtencion	int	<input checked="" type="checkbox"/>
	EsPacienteExterno	bit	<input checked="" type="checkbox"/>
	idSunasaPacienteHistorico	int	<input checked="" type="checkbox"/>

**Figura N° 5:** BD-Tabla de Atenciones.

**Fuente:** Hospital La Caleta

Empleados			
	Nombre de columna	Tipo de datos	Permitir valores NULL
🔑	IdEmpleado	int	<input type="checkbox"/>
	ApellidoPaterno	varchar(50)	<input type="checkbox"/>
	ApellidoMaterno	varchar(50)	<input type="checkbox"/>
	Nombres	varchar(50)	<input type="checkbox"/>
	IdCondicionTrabajo	int	<input type="checkbox"/>
	IdTipoEmpleado	int	<input type="checkbox"/>
	DNI	char(20)	<input type="checkbox"/>
	CodigoPlanilla	varchar(8)	<input type="checkbox"/>
	FechaIngreso	datetime	<input checked="" type="checkbox"/>
	FechaAlta	datetime	<input checked="" type="checkbox"/>
	Usuario	varchar(20)	<input checked="" type="checkbox"/>
	Clave	varchar(200)	<input checked="" type="checkbox"/>
	loginEstado	int	<input checked="" type="checkbox"/>
	loginPC	varchar(50)	<input checked="" type="checkbox"/>
	FechaNacimiento	datetime	<input checked="" type="checkbox"/>
	idTipoDestacado	int	<input checked="" type="checkbox"/>
	IdEstablecimientoExterno	int	<input checked="" type="checkbox"/>
	HisCodigoDigitador	varchar(20)	<input checked="" type="checkbox"/>
	ReniecAutorizado	bit	<input checked="" type="checkbox"/>
	idTipoDocumento	int	<input checked="" type="checkbox"/>
	idSupervisor	int	<input checked="" type="checkbox"/>
	esActivo	bit	<input checked="" type="checkbox"/>

**Figura N° 6:** BD-Tabla de Empleados.

**Fuente:** Hospital La Caleta

Especialidades			
	Nombre de columna	Tipo de datos	Permitir valores NULL
🔑	IdEspecialidad	int	<input type="checkbox"/>
	Nombre	varchar(50)	<input type="checkbox"/>
	IdDepartamento	int	<input type="checkbox"/>
	TiempoPromedioAtenc...	int	<input checked="" type="checkbox"/>

**Figura N° 7:** BD-Tabla de Especialidades.

**Fuente:** Hospital La Caleta

Medicos			
	Nombre de columna	Tipo de datos	Permitir valores NULL
🔑	IdMedico	int	<input type="checkbox"/>
	Colegiatura	char(6)	<input type="checkbox"/>
	IdEmpleado	int	<input checked="" type="checkbox"/>
	LoteHIS	varchar(2)	<input checked="" type="checkbox"/>
	idColegioHIS	varchar(2)	<input checked="" type="checkbox"/>
	rne	varchar(50)	<input checked="" type="checkbox"/>
	egresado	bit	<input checked="" type="checkbox"/>

**Figura N° 8:** BD-Tabla de Medicos.

**Fuente:** Hospital La Caleta

ProgramacionMedica			
	Nombre de columna	Tipo de datos	Permitir valores NULL
🔑	IdProgramacion	int	<input type="checkbox"/>
	IdMedico	int	<input type="checkbox"/>
	IdDepartamento	int	<input type="checkbox"/>
	Fecha	datetime	<input type="checkbox"/>
	HoraInicio	char(5)	<input type="checkbox"/>
	HoraFin	char(5)	<input type="checkbox"/>
	IdTipoProgramacion	int	<input type="checkbox"/>
	Descripcion	varchar(100)	<input checked="" type="checkbox"/>
	IdTurno	int	<input checked="" type="checkbox"/>
	IdEspecialidad	int	<input checked="" type="checkbox"/>
	Color	int	<input checked="" type="checkbox"/>
	IdServicio	int	<input checked="" type="checkbox"/>
	IdTipoServicio	int	<input checked="" type="checkbox"/>
	FechaReg	datetime	<input checked="" type="checkbox"/>
	TiempoPromedioAtencion	int	<input checked="" type="checkbox"/>
	HoraFinProgramacion	char(5)	<input checked="" type="checkbox"/>

**Figura N° 9:** BD-Tabla de ProgramaciónMedica.

**Fuente:** Hospital La Caleta

Pacientes			
	Nombre de columna	Tipo de datos	Permitir valores NULL
🔑	IdPaciente	int	<input type="checkbox"/>
	ApellidoPaterno	varchar(40)	<input checked="" type="checkbox"/>
	ApellidoMaterno	varchar(40)	<input checked="" type="checkbox"/>
	PrimerNombre	varchar(40)	<input checked="" type="checkbox"/>
	SegundoNombre	varchar(40)	<input checked="" type="checkbox"/>
	TercerNombre	varchar(40)	<input checked="" type="checkbox"/>
	FechaNacimiento	datetime	<input checked="" type="checkbox"/>
	NroDocumento	varchar(12)	<input checked="" type="checkbox"/>
	Telefono	varchar(10)	<input checked="" type="checkbox"/>
	DireccionDomicilio	varchar(100)	<input checked="" type="checkbox"/>
	Autogenerado	varchar(30)	<input checked="" type="checkbox"/>
	IdTipoSexo	int	<input checked="" type="checkbox"/>
	IdProcedencia	int	<input checked="" type="checkbox"/>
	IdGradoInstruccion	int	<input checked="" type="checkbox"/>
	IdEstadoCivil	int	<input checked="" type="checkbox"/>
	IdDocIdentidad	int	<input checked="" type="checkbox"/>
	IdTipoOcupacion	int	<input checked="" type="checkbox"/>
	IdCentroPobladoNa...	int	<input checked="" type="checkbox"/>
	IdCentroPobladoDo...	int	<input checked="" type="checkbox"/>
	NombrePadre	varchar(20)	<input checked="" type="checkbox"/>
	NombreMadre	varchar(20)	<input checked="" type="checkbox"/>
	NroHistoriaClinica	int	<input checked="" type="checkbox"/>
	IdTipoNumeracion	int	<input checked="" type="checkbox"/>
	IdCentroPobladoPr...	int	<input checked="" type="checkbox"/>
	Observacion	varchar(150)	<input checked="" type="checkbox"/>
	IdPaisDomicilio	int	<input checked="" type="checkbox"/>
	IdPaisProcedencia	int	<input checked="" type="checkbox"/>
	IdPaisNacimiento	int	<input checked="" type="checkbox"/>
	IdDistritoProcedencia	int	<input checked="" type="checkbox"/>
	IdDistritoDomicilio	int	<input checked="" type="checkbox"/>
	IdDistritoNacimiento	int	<input checked="" type="checkbox"/>
	FichaFamiliar	varchar(20)	<input checked="" type="checkbox"/>
	IdEtnia	varchar(2)	<input checked="" type="checkbox"/>
	GrupoSanguineo	varchar(10)	<input checked="" type="checkbox"/>
	FactorRh	varchar(10)	<input checked="" type="checkbox"/>
	UsoWebReniec	bit	<input checked="" type="checkbox"/>
	IdIdioma	int	<input checked="" type="checkbox"/>
	Email	varchar(50)	<input checked="" type="checkbox"/>
	madreDocumento	varchar(12)	<input checked="" type="checkbox"/>
	madreApellidoPaterno	varchar(20)	<input checked="" type="checkbox"/>
	madreApellidoMaterno	varchar(20)	<input checked="" type="checkbox"/>
	madrePrimerNombre	varchar(20)	<input checked="" type="checkbox"/>
	madreSegundoNom...	varchar(20)	<input checked="" type="checkbox"/>
	NroOrdenHijo	int	<input checked="" type="checkbox"/>
	madreTipoDocumento	int	<input checked="" type="checkbox"/>
	Sector	varchar(2)	<input checked="" type="checkbox"/>
	Sectorista	int	<input checked="" type="checkbox"/>

**Figura N° 10:** BD- Tabla de Pacientes.

**Fuente:** Hospital La Caleta

Servicios			
	Nombre de columna	Tipo de datos	Permitir valores NULL
🔑	IdServicio	int	<input type="checkbox"/>
	Nombre	varchar(50)	<input type="checkbox"/>
	IdEspecialidad	int	<input type="checkbox"/>
	IdTipoServicio	int	<input type="checkbox"/>
	Codigo	varchar(6)	<input type="checkbox"/>
	SVG	varchar(1000)	<input checked="" type="checkbox"/>
	IdProducto	int	<input checked="" type="checkbox"/>
	soloTipoSexo	int	<input checked="" type="checkbox"/>
	maximaEdad	int	<input checked="" type="checkbox"/>
	codigoServicioSEM	varchar(2)	<input checked="" type="checkbox"/>
	ubicacionSEM	varchar(5)	<input checked="" type="checkbox"/>
	codigoServicioHIS	varchar(6)	<input checked="" type="checkbox"/>
	CostoCeroCE	varchar(1)	<input checked="" type="checkbox"/>
	MinimaEdad	int	<input checked="" type="checkbox"/>
	idEstado	int	<input checked="" type="checkbox"/>
	Triaje	bit	<input checked="" type="checkbox"/>
	EsObservacionEmergencia	bit	<input checked="" type="checkbox"/>
	UsaModuloNinoSano	bit	<input checked="" type="checkbox"/>
	UsaModuloMaterno	bit	<input checked="" type="checkbox"/>
	UsaGalenHos	bit	<input checked="" type="checkbox"/>
	TipoEdad	int	<input checked="" type="checkbox"/>
	UsaFUA	bit	<input checked="" type="checkbox"/>
	codigoServicioSuSalud	varchar(7)	<input checked="" type="checkbox"/>
	codigoServicioFUA	varchar(6)	<input checked="" type="checkbox"/>
	FuaTipoAnexo2015	int	<input checked="" type="checkbox"/>

**Figura N° 11:** BD- Tabla de Servicios.

**Fuente:** Hospital La Caleta

MedicosEspecialidad			
	Nombre de columna	Tipo de datos	Permitir valores NULL
🔑	IdMedicoEspecialidad	int	<input type="checkbox"/>
	IdEspecialidad	int	<input type="checkbox"/>
	IdMedico	int	<input type="checkbox"/>

**Figura N° 12:** BD- Tabla de MedicosEspecialidad.

**Fuente:** Hospital La Caleta

### 4.1.3 Procesamiento de datos.

A continuación, para poder realizar el procesamiento de datos haremos uso de la Metodología Hefesto que se detalla a continuación:

#### 4.1.3.1 *Análisis de Requerimientos.*

##### 4.1.3.1.1 *Identificar Preguntas.*

El proceso más importante en las actividades cotidianas del Hospital la Caleta es el de Gestión Asistencial, conformado por los sub procesos de Consulta Externa y Atenciones de Emergencia.

Luego de recopilar la información del Hospital, sus objetivos y necesidades se procedieron a analizar esta información y extraer los requerimientos del Data Warehouse. Como resultado de este análisis, se obtuvo que el Hospital necesita obtener respuesta a las siguientes preguntas:

##### **Gestión Asistencial:**

- **R1:** ¿Cuál es la productividad hora médico por cada hora programada de trabajo del médico?
- **R2:** ¿Cuál es la cantidad de pacientes atendidos por cada servicio en un período de tiempo?
- **R3:** ¿Cuál es el promedio de consultas que se han otorgado a cada paciente que acudió al hospital durante un período de tiempo?
- **R4:** ¿Cuál es el grado de cumplimiento de la programación de las horas médico por servicio o especialidad en un periodo de tiempo?
- **R5:** ¿Cuál es el promedio de permanencia de los pacientes que egresaron en un periodo de tiempo?

#### **4.1.3.1.2** *Identificar Indicadores y Perspectivas.*

En síntesis, los indicadores que se utilizaron y perspectivas de análisis son:

##### **Indicador:**

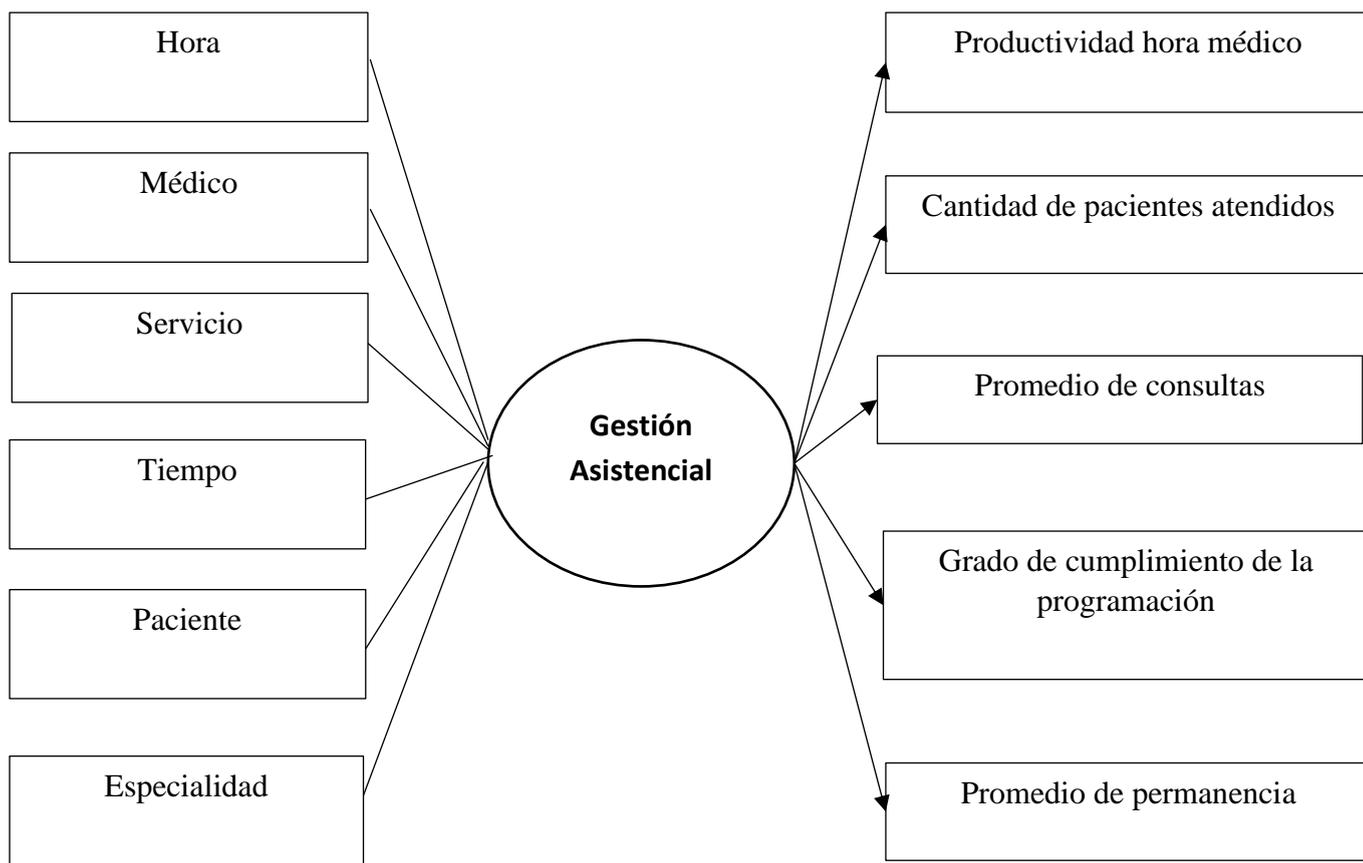
- ✓ Productividad hora médico
- ✓ Cantidad de pacientes atendidos
- ✓ Promedio de consultas
- ✓ Grado de cumplimiento de la programación
- ✓ Promedio de permanencia

##### **Perspectivas:**

- ✓ Hora
- ✓ Médico
- ✓ Servicio
- ✓ Tiempo
- ✓ Paciente
- ✓ Especialidad

#### **4.1.3.1.3** *Modelo Conceptual.*

Considerando los indicadores y perspectivas analizadas en el paso anterior se plantea el modelo conceptual, que permite observar con mayor facilidad el alcance del proyecto en base al proceso de:



**Figura N° 13:** Modelo conceptual del proceso de Gestión Asistencial.

#### 4.1.3.2 *Análisis de las Fuentes.*

En esta sección se continúa con la aplicación de la metodología Hefesto.

##### 4.1.3.2.1 *Conformar Indicadores.*

En este paso se explica las operaciones para calcular los indicadores de cada tabla de hecho que forman parte de los procesos del Hospital. Los indicadores se calcularán de la siguiente manera.

- **Productividad hora médico**
  - **Hechos:**  $(N^{\circ} \text{ de Atenciones Médicas realizadas en un período}) / (N^{\circ} \text{ de horas médico Programadas en el mismo período})$

- **Cantidad de pacientes atendidos**
  - **Hechos:** N° de atenciones (pacientes nuevos y reingresantes) por servicio en un periodo de tiempo.
  
- **Promedio de consultas**
  - **Hechos:** (N° de atenciones en un periodo)/ (N° de atendidos ( nuevos y reingresos) en el hospital en el mismo periodo)
  
- **Grado de cumplimiento de la programación**
  - **Hechos:**  $100 * (\text{N}^\circ \text{ de horas médico efectivas de atención en un periodo}) / (\text{N}^\circ \text{ de horas médico programadas en el mismo periodo})$
  
- **Promedio de permanencia**
  - **Hechos:** Total días de estancias de pacientes egresados en un periodo de tiempo / Total de egresos del mismo periodo.

#### 4.1.3.2.2 *Establecer Correspondencias.*

Una vez analizado este modelo de la Base de Datos del Hospital, y sus relaciones con el modelo conceptual, se obtuvieron las siguientes conclusiones:

### Gestión Asistencial:

**Tabla N° 4:** Conclusión del Modelo Conceptual  
“Respectivas”.

<b>PERSPECTIVA</b>	<b>TABLA/CAMPO</b>
<b>Hora</b>	ProgramacionMedica.Fecha ProgramacionMedica.HoraInicio ProgramacionMedica.HoraFin
<b>Médico</b>	Empleados.ApellidoPaterno Empleados.ApellidoMaterno Empleados.Nombres
<b>Servicio</b>	Servicios.Nombre
<b>Tiempo</b>	Atenciones.FechaIngreso
<b>Paciente</b>	Pacientes.ApellidoPaterno Pacientes.ApellidoMaterno Pacientes.PrimerNombre Pacientes.SegundoNombre Pacientes.TercerNombre
<b>Especialidad</b>	Especialidades.Nombre

**Tabla N° 5:** Conclusión del Modelo Conceptual  
“Indicadores.”

<b>Indicador</b>	<b>Tabla/Campo</b>
Productividad hora médico	Atenciones.IdAtencion ProgramacionMedica.HoraInicio ProgramacionMedica.HoraFin
Cantidad de pacientes atendidos	Atenciones. IdAtencion
Promedio de consultas	Atenciones.IdAtencion

Grado de cumplimiento de la programación	Atenciones.HoraIngreso ProgramacionMedica.HoraInicio ProgramacionMedica.HoraFin
Promedio de permanencia	Atenciones.IdAtencion Atenciones.FechaIngreso Atenciones.FechaEgreso

#### 4.1.3.2.3 Nivel de Granularidad

Conforme a las correspondencias determinadas anteriormente, se examinaron los campos existentes en la tabla a la que se hace referencia, para intuir los significados de cada campo que devuelve.

#### **Gestión Asistencial:**

- Con respecto a la perspectiva “**Hora**”, los datos disponibles son los siguientes:

**Tabla N° 6:** Perspectiva “Hora”

<b>Campo</b>	<b>Significado/Utilidad</b>
ProgramacionMedica.Fecha	Fecha Programada para realizar sus labores un médico.
ProgramacionMedica.HoraInicio	Hora Programada para el Inicio de sus labores de un médico.
ProgramacionMedica.HoraFin	Hora Programada para el Término de sus labores de un médico.

- Con respecto a la perspectiva “**Médico**”, los datos disponibles son los siguientes:

**Tabla N° 7:** Perspectiva "Médico".

<b>Campo</b>	<b>Significado/Utilidad</b>
Empleados.ApellidoPaterno	Apellido Paterno de un médico.
Empleados.ApellidoMaterno	Apellido Materno de un médico.
Empleados.Nombres	Nombres de un médico.

- Con respecto a la perspectiva “**Servicio**”, los datos disponibles son los siguientes:

**Tabla N° 8:** Perspectiva "Servicio".

<b>Campo</b>	<b>Significado/Utilidad</b>
Servicios.Nombre	Nombre de un servicio que brindado el hospital.

- Con respecto a la perspectiva “**Tiempo**”, los datos disponibles son los siguientes:

**Tabla N° 9:** Perspectiva "Tiempo".

<b>Campo</b>	<b>Significado/Utilidad</b>
Año	Año en la que ingreso el paciente para ser atendido.
Mes	Mes en la que ingreso el paciente para ser atendido.
Día	Día en la que ingreso el paciente para ser atendido.

- Con respecto a la perspectiva “**Paciente**”, los datos disponibles son los siguientes:

**Tabla N° 10:** Perspectiva "Paciente".

<b>Campo</b>	<b>Significado/Utilidad</b>
Pacientes.ApellidoPaterno	Apellido paterno del paciente.
Pacientes.ApellidoMaterno	Apellido materno del paciente.
Pacientes.PrimerNombre	Primer nombre del paciente.
Pacientes.SegundoNombre	Segundo nombre del paciente.
Pacientes.TercerNombre	Tercer nombre del paciente.

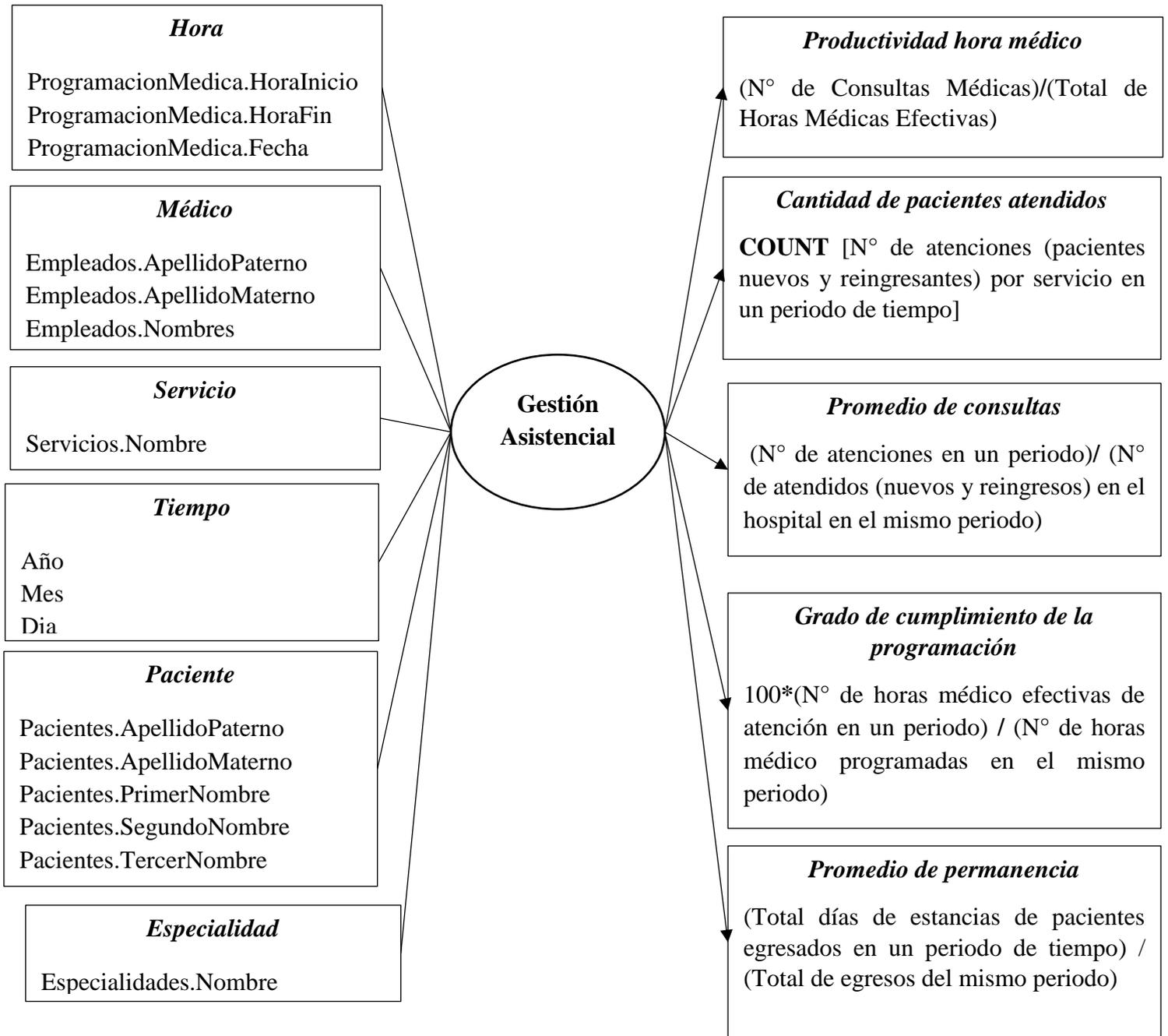
- Con respecto a la perspectiva “Especialidad” , los datos disponibles son los siguientes:

**Tabla N° 11:** Perspectiva "Especialidad".

<b>Campo</b>	<b>Significado/Utilidad</b>
Especialidades.Nombre	Nombre de una especialidad médica.

#### **4.1.3.2.4** *Modelo conceptual ampliado.*

En este paso, y con el fin de graficar los resultados obtenidos en los pasos anteriores, se ampliará el modelo conceptual, colocando bajo cada perspectiva los campos elegidos y bajo cada indicador su respectiva fórmula de cálculo. Gráficamente:



**Figura N° 14:** Modelo Conceptual Ampliado del proceso de Gestión Asistencial.

### 4.1.3.3 Modelo lógico del data warehouse

#### 4.1.3.3.1 Tipo de modelo lógico del data warehouse.

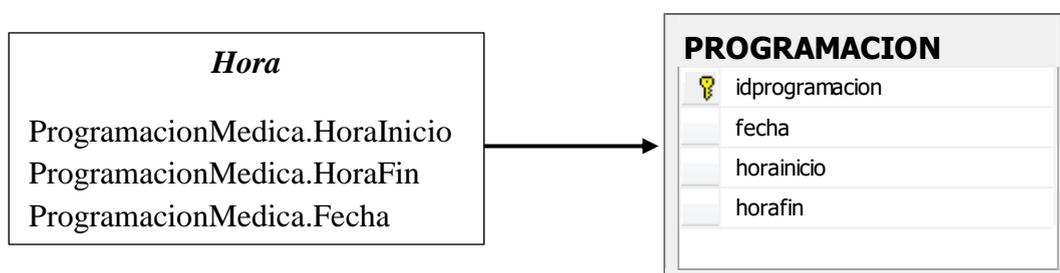
El esquema que se utilizará será en estrella, debido a sus características, ventajas y diferencias con los otros esquemas.

#### 4.1.3.3.2 Establecer tabla de dimensiones.

A continuación, se diseñaran las tablas de dimensiones:

- Perspectiva “**Hora**”:
  - La nueva Tabla de dimensión tendrá el nombre de “PROGRAMACION”.
  - Se le agregara una clave principal con el nombre idprogramacion.
  - El nombre de los campos no serán modificados.

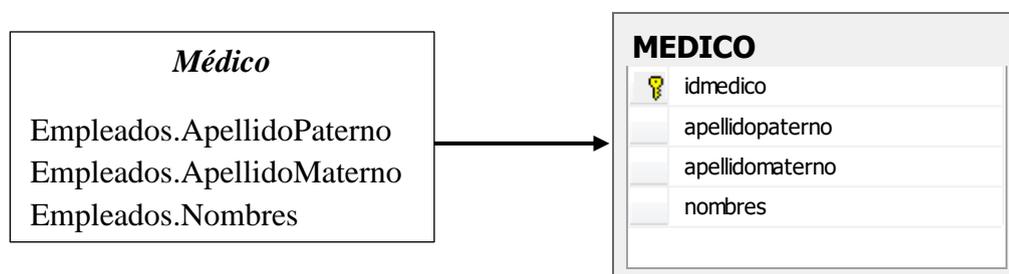
Se puede apreciar el resultado de estas operaciones en la siguiente gráfica:



**Figura N° 15:** Tabla de Dimensión "PROGRAMACIÓN".

- Perspectiva “**Médico**”:
  - La nueva Tabla de dimensión tendrá el nombre de “MEDICO”.
  - Se le agregara una clave principal con el nombre idmedico.
  - El nombre de los campos no serán modificados.

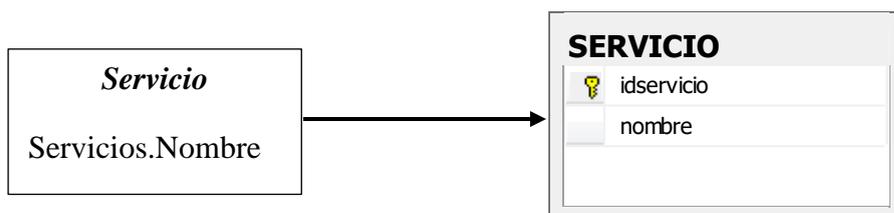
Se puede apreciar el resultado de estas operaciones en la siguiente gráfica:



**Figura N° 16:** Tabla de Dimensión "MÉDICO".

- Perspectiva “**Servicio**”:
  - La nueva Tabla de dimensión tendrá el nombre de “SERVICIO”.
  - Se le agregara una clave principal con el nombre idservicio.
  - El nombre de los campos no serán modificados.

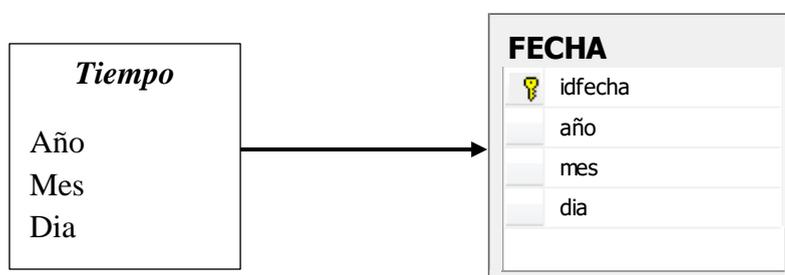
Se puede apreciar el resultado de estas operaciones en la siguiente gráfica:



**Figura N° 17:** Tabla de Dimensión "SERVICIO".

- Perspectiva “**Tiempo**”:
  - La nueva Tabla de dimensión tendrá el nombre de “FECHA”.
  - Se le agregara una clave principal con el nombre idfecha.
  - El nombre de los campos no serán modificados.

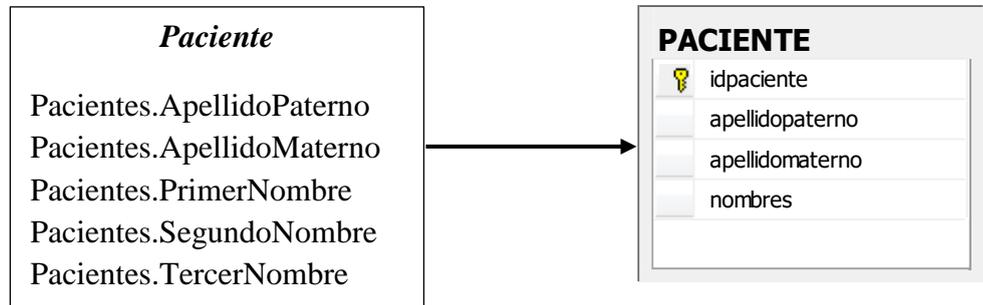
Se puede apreciar el resultado de estas operaciones en la siguiente gráfica:



**Figura N° 18:** Tabla de Dimensión "FECHA".

- Perspectiva “**Paciente**”:
  - La nueva Tabla de dimensión tendrá el nombre de “PACIENTE”.
  - Se le agregara una clave principal con el nombre idpaciente.
  - El nombre del campo ApellidoPaterno no será cambiado.
  - El nombre del campo ApellidoMaterno no será cambiado.
  - Los campos PrimerNombre, SegundoNombre y TercerNombre serán concatenado y se llamara Nombres.

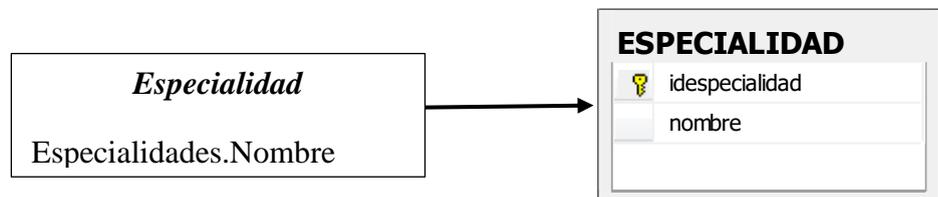
Se puede apreciar el resultado de estas operaciones en la siguiente gráfica:



**Figura N° 19:** Tabla de Dimensión "PACIENTE".

- Perspectiva “Especialidad”:
  - La nueva Tabla de dimensión tendrá el nombre de “ESPECIALIDAD”.
  - Se le agregara una clave principal con el nombre idespecialidad.
  - El nombre de los campos no serán modificados.

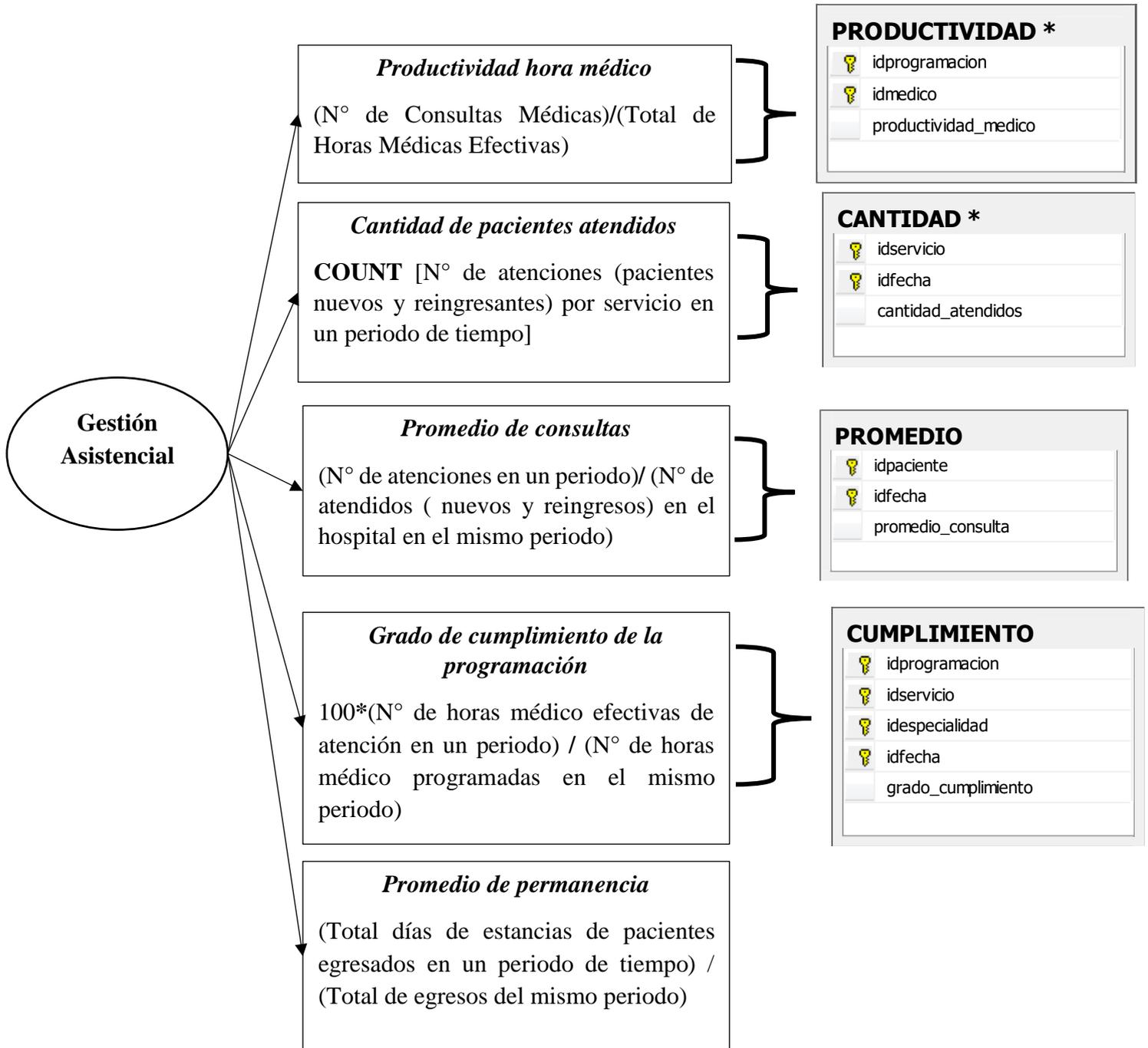
Se puede apreciar el resultado de estas operaciones en la siguiente gráfica:



**Figura N° 20:** Tabla de Dimensión "ESPECIALIDAD".

#### 4.1.3.3.3 Establecer tabla de hechos.

A continuación, se confeccionaran la tabla de hechos:



**Figura N° 21:** Tabla de Hechos del proceso de Gestión Asistencial.

#### 4.1.3.3.4 Uniones.

Se realizaran las uniones pertinentes, de acuerdo corresponda:

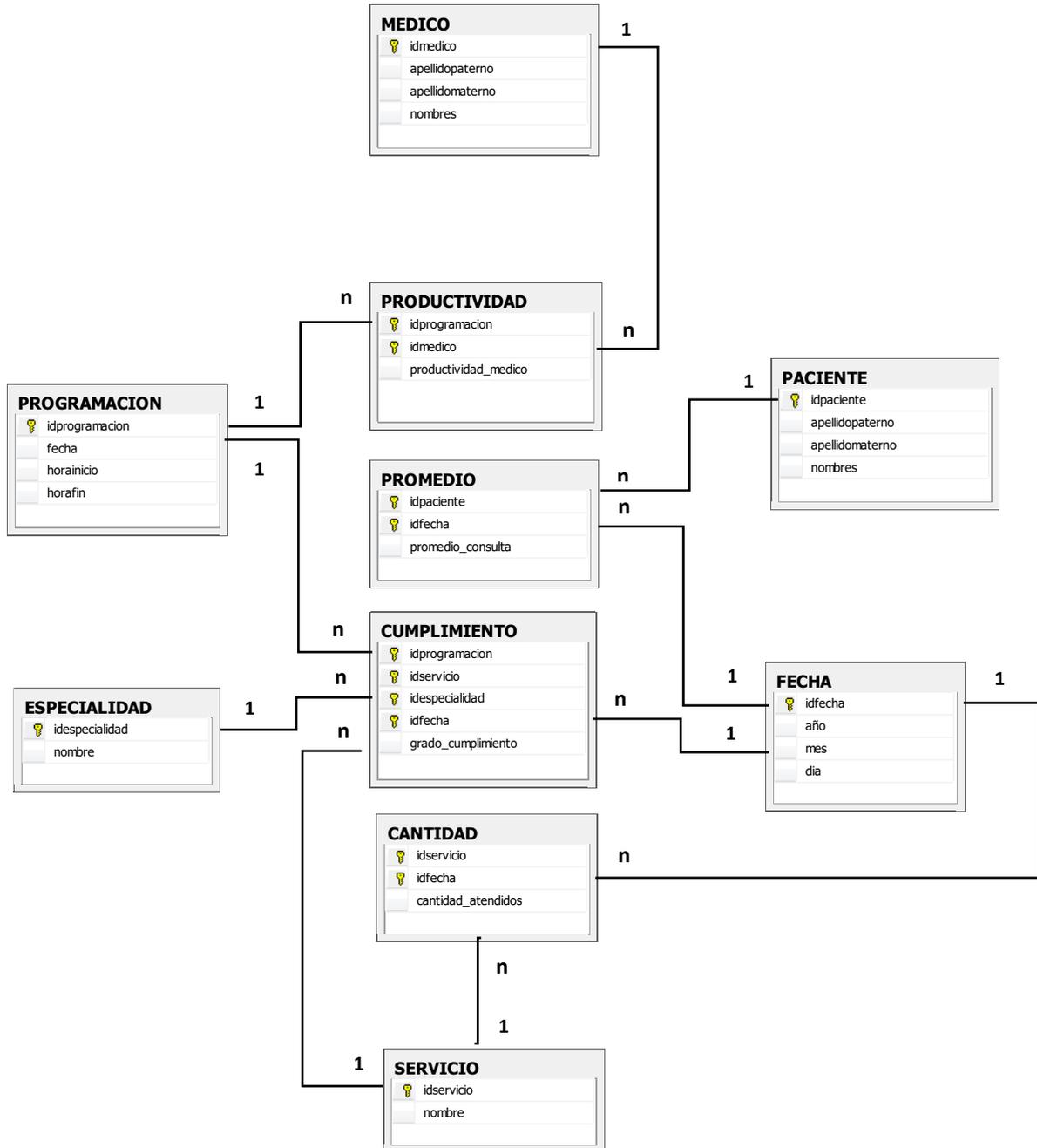


Figura N° 22: Uniones correspondientes.

#### 4.1.3.4 *Integración de datos.*

##### 4.1.3.4.1 *Carga inicial.*

Las tareas que lleva como objetivo el proceso de carga inicial son las siguientes:

- Inicio. - Se realiza una limpieza de los datos de las tablas de dimensiones y de hechos, eliminando todos los datos de pruebas.
- Establecer variables fecha.desde y fecha.hasta: establece dos variables globales que serán utilizadas posteriormente por algunos pasos.
  - ✓ Para la variable "fecha.desde" se obtiene el valor de la fecha de inicio del periodo de tiempo.
  - ✓ Para la variable "fecha\_hasta" se obtiene el valor de la fecha actual o de fin del periodo de tiempo.
- Cargar Programación. - Ejecuta la tarea que contiene los pasos que cargará la dimensión PROGRAMACION.
- Cargar Médicos. - Ejecuta la tarea que contiene los pasos que cargará la dimensión MEDICO.
- Cargar Servicios. - Ejecuta la tarea que contiene los pasos que cargará la dimensión SERVICIO.
- Cargar Fechas. - Ejecuta la tarea que contiene los pasos que cargará la dimensión FECHA.
- Cargar Pacientes. - Ejecuta la tarea que contiene los pasos que cargará la dimensión PACIENTE.

- Cargar Especialidades. - Ejecuta la tarea que contiene los pasos que cargará la dimensión ESPECIALIDAD.
- Cargar Productividad. - Ejecuta la tarea que contiene los pasos que cargará la tabla hechos PRODUCTIVIDAD.
- Cargar Cantidad. - Ejecuta la tarea que contiene los pasos que cargará la tabla hechos CANTIDAD.
- Cargar Cumplimiento. - Ejecuta la tarea que contiene los pasos que cargará la tabla hechos CUMPLIMIENTO.
- Cargar Promedio. - Ejecuta la tarea que contiene los pasos que cargará la tabla hechos PROMEDIO.

A continuación, se detallarán los pasos que realiza cada una de las tareas del proceso de carga inicial:

- **Cargar Programación.-**



**Figura N° 23:** PENTAHO- Carga de Dimensión PROGRAMACION.

- ✓ Obtener datos de OLTP.- Obtiene a través de una conexión con **MS SQL Server** los datos necesarios para cargar la dimensión PROGRAMACION.

- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Programación. - Almacena en la tabla de dimensión PROGRAMACION los datos obtenidos en los pasos anteriores.

- **Cargar Médicos.-**



**Figura N° 24:** PENTAHO- Carga de Dimensión MEDICO.

- ✓ Obtener datos de OLTP.- Obtiene a través de una conexión con **MS SQL Server** los datos necesarios para cargar la dimensión MEDICO.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Médicos. - Almacena en la tabla de dimensión MEDICO los datos obtenidos en los pasos anteriores.

- **Cargar Servicios.-**



**Figura N° 25:** PENTAHO - Carga de Dimensión SERVICIO.

- ✓ Obtener datos de OLTP.- Obtiene a través de una conexión con **MS SQL Server** los datos necesarios para cargar la dimensión **SERVICIO**.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Servicios. - Almacena en la tabla de dimensión **SERVICIO** los datos obtenidos en los pasos anteriores.

- **Cargar Fechas.-**

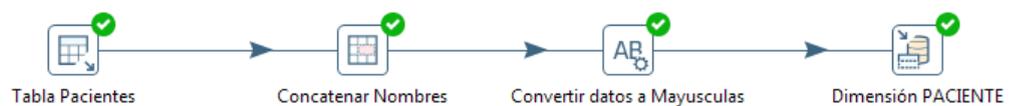


**Figura N° 26:** PENTAHO - Carga de Dimensión FECHA.

- ✓ Obtener datos de OLTP.- Se realiza el siguiente procedimiento para obtener los datos necesarios para cargar la dimensión **FECHA**.

- Recibe como parámetros la consulta SQL los valores de "fecha.desde" y "fecha.hasta".
- Recorre una a una las fechas que se encuentran dentro de este intervalo.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Fechas. - Almacena en la tabla de dimensión FECHA los datos obtenidos en los pasos anteriores.

• **Cargar Pacientes.-**

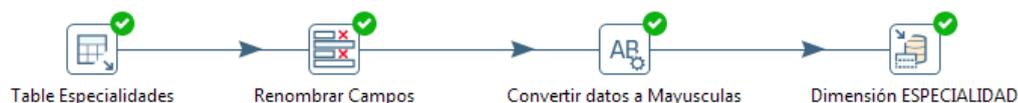


**Figura N° 27:** PENTAHO - Carga de Dimensión PACIENTE.

- ✓ Obtener datos de OLTP.- Obtiene a través de una conexión con **MS SQL Server** los datos necesarios para cargar la dimensión PACIENTE.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.

- ✓ Cargar Pacientes. - Almacena en la tabla de dimensión PACIENTE los datos obtenidos en los pasos anteriores.

- **Cargar Especialidades.-**



**Figura N° 28:** PENTAHO - Carga de Dimensión ESPECIALIDAD.

- ✓ Obtener datos de OLTP.- Obtiene a través de una conexión con **MS SQL Server** los datos necesarios para cargar la dimensión ESPECIALIDAD.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Especialidades. - Almacena en la tabla de dimensión ESPECIALIDAD los datos obtenidos en los pasos anteriores.

- **Cargar Productividad.-**



**Figura N° 29:** PENTAHO - Carga de Hechos PRODUCTIVIDAD.

- ✓ Obtener datos de OLTP. - Se tomaron como fuente la tabla “Atenciones” Y “ProgramacionMedica” que a través de una conexión con **MS SQL Server** obtiene los datos necesarios para cargar la tabla de Hechos PRODUCTIVIDAD.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Productividad. - Almacena en la tabla de hechos PRODUCTIVIDAD los datos obtenidos en los pasos anteriores.

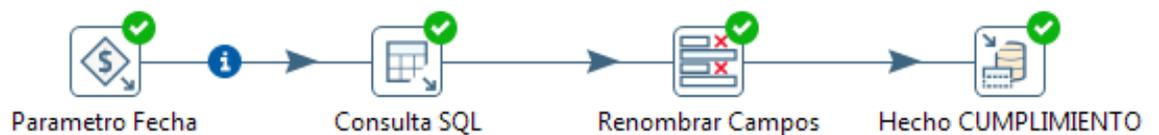
- **Cargar Cantidad.-**



**Figura N° 30:** PENTAHO - Carga de Hechos CANTIDAD.

- ✓ Obtener datos de OLTP. - Se tomaron como fuente la tabla “Atenciones” que a través de una conexión con **MS SQL Server** obtiene los datos necesarios para cargar la tabla de Hechos CANTIDAD.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Cantidad. - Almacena en la tabla de hechos CANTIDAD los datos obtenidos en los pasos anteriores.

- **Cargar Cumplimiento.-**



**Figura N° 31: PENTAHO - Carga de Hechos CUMPLIMIENTO.**

- ✓ Obtener datos de OLTP. - Se tomaron como fuente la tabla “Atenciones” y “ProgramacionMedica” que a través de una conexión con **MS SQL Server** obtiene los datos necesarios para cargar la tabla de Hechos CUMPLIMIENTO.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.

- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Cantidad. - Almacena en la tabla de hechos CUMPLIMIENTO los datos obtenidos en los pasos anteriores.

- **Cargar Promedio. -**



**Figura N° 32:** PENTAHO - Carga de Hechos PROMEDIO.

- ✓ Obtener datos de OLTP. - Se tomaron como fuente la tabla “Atenciones” que a través de una conexión con **MS SQL Server** obtiene los datos necesarios para cargar la tabla de Hechos PROMEDIO.
- ✓ Conversión de Datos. - Realiza la conversión de tipos de datos del origen para que coincidan con el tipo de dato del destino.
- ✓ Ordenar. - Se realiza una ordenación para evitar la duplicación de códigos.
- ✓ Cargar Promedio. - Almacena en la tabla de hechos PROMEDIO los datos obtenidos en los pasos anteriores.

#### 4.1.3.4.2 Actualización.



**Figura N° 33:** PENTAHO - Trabajo ETL de Actualización.

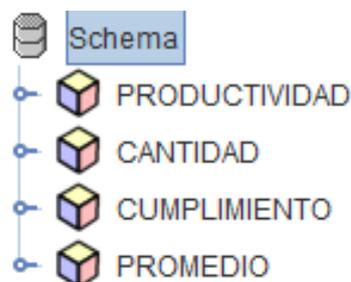
Las políticas de Actualización que se han convenido con los usuarios son las siguientes:

- **Inicio:** La información se refrescará cada vez que los usuarios lo requieran.
- Establecer variables fecha.desde y fecha.hasta:
  - La variable "fecha.desde" será ingresado por el usuario.
  - La variable "fecha.hasta" será ingresado por el usuario.
- **Limpiar Tablas.** - A la serie de pasos posteriores, se le antecederá un nuevo paso que borrará los datos de las tablas dimensiones y tablas hechos.
- **Cargar Dimensiones.**-Los datos de las tablas de dimensiones PROGRAMACION, MEDICO, SERVICIO, PACIENTE, FECHA y ESPECIALIDAD serán cargadas por completo cada vez que el usuario los actualicé. Pero la carga de Dimensión FECHA en este paso, se recibe los valores para la variable fecha.desde y fecha.hasta.

- **Cargar Hechos.**- Los datos de la tabla hechos PRODUCTIVIDAD, CANTIDAD, PROMEDIO y CUMPLIMIENTO se refrescarán cada vez que el usuario así lo necesite. En el paso obtener datos OLTP se tomará en cuenta solo los datos que estén entre el intervalo de “fecha.desde” y “fecha.hasta”.
- Los datos de la tabla de hechos que corresponden al último mes (30 días) a partir de la fecha actual, serán reemplazados cada vez.
- Estas acciones se realizarán durante un periodo de prueba, para analizar cuál es la manera más eficiente de generar las actualizaciones, basadas en el estudio de los cambios que se producen en los OLTP y que afectan al contenido del DW.

#### 4.1.3.5 Creación de cubos multidimensionales.

Para la implantación del cubo multidimensional se utilizó el módulo Pentaho Schema Workbench, el cual administra la conexión al almacén de datos, las tablas de hechos, de dimensiones y la ejecución del cubo multidimensional.



**Figura N° 34:** Cubo Multidimensional.

#### 4.1.3.5.1 Creación de indicadores.

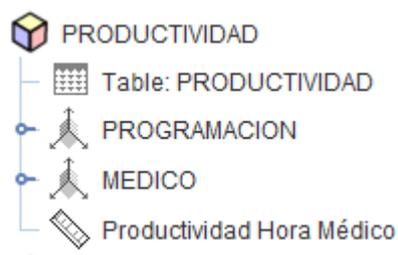
##### 4.1.3.5.1.1 Cubo productividad.

De la tabla hechos “PRODUCTIVIDAD”, se sumarizara el hecho “productividad\_medico” para crear el indicador denominado.

- Productividad hora médico

La fórmula utilizada para crear este indicador es la siguiente:

- “Productividad hora médico” =  
SUM(PRODUCTIVIDAD.  
productividad\_medico).



**Figura N° 35:** Cubo PRODUCTIVIDAD, indicadores.

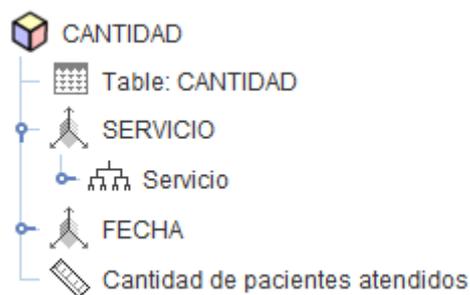
##### 4.1.3.5.1.2 Cubo cantidad.

De la tabla hechos “CANTIDAD”, se sumarizara el hecho “cantidad\_atendidos” para crear el indicador denominado.

- Cantidad de pacientes atendidos

La fórmula utilizada para crear este indicador es la siguiente:

- “Cantidad de pacientes atendidos”  
= SUM(CANTIDAD.cantidad\_atendidos).



**Figura N° 36:** Cubo CANTIDAD, indicadores.

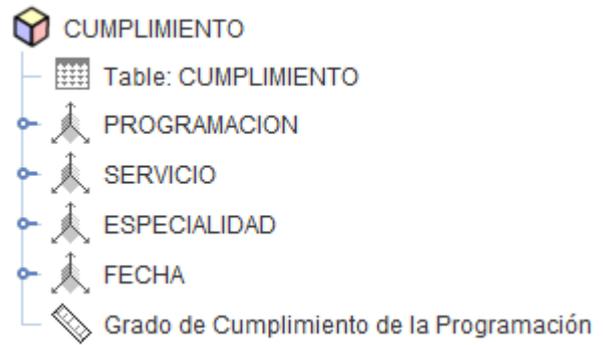
#### 4.1.3.5.1.3 *Cubo cumplimiento.*

De la tabla hechos “CUMPLIMIENTO”, se sumariza el hecho “grado\_cumplimiento” para crear el indicador denominado.

- Grado de cumplimiento de la programación

La fórmula utilizada para crear este indicador es la siguiente:

- “Grado de cumplimiento de la programación”  
= SUM(CUMPLIMIENTO.grado\_cumplimiento).



**Figura N° 37:** Cubo CUMPLIMIENTO, indicadores.

4.1.3.5.1.4 *Cubo promedio.*

De la tabla hechos “PROMEDIO”, se sumará el hecho “promedio\_consulta” para crear el indicador denominado.

- Promedio de consultas

La fórmula utilizada para crear este indicador es la siguiente:

- “Promedio de consultas” = SUM(PROMEDIO.promedio\_consulta).



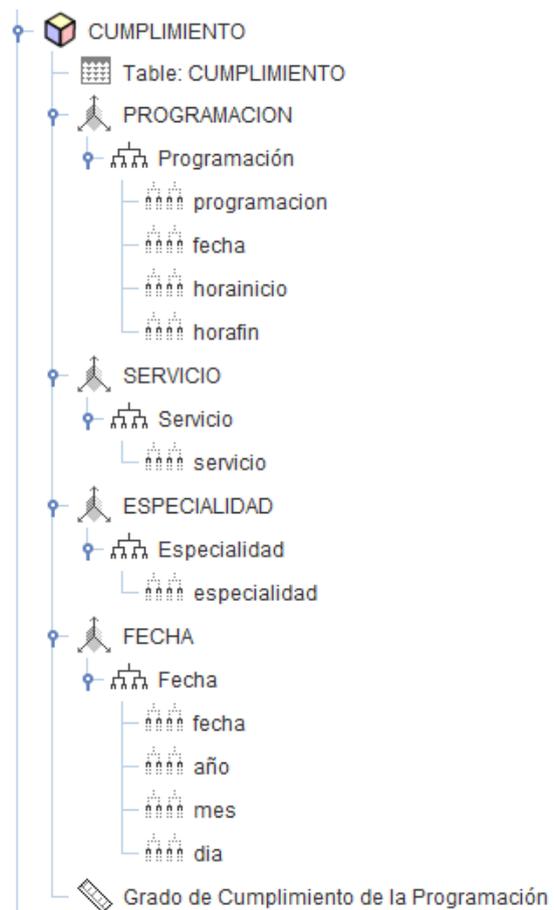
**Figura N° 38:** Cubo PROMEDIO, indicadores.

#### 4.1.3.5.2 Creación de atributos.

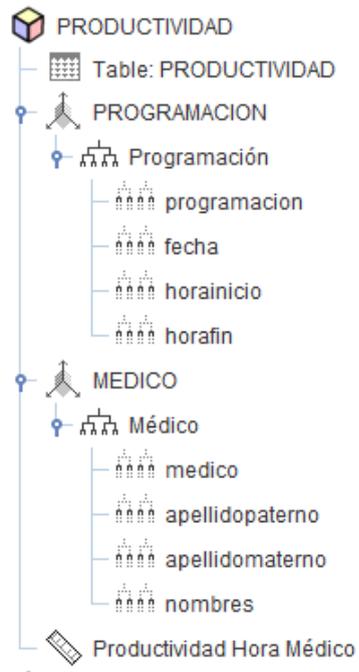
Se crearon y se agregaron al cubo los siguientes atributos:

- De la tabla dimensión “PROGRAMACIÓN”, se tomó los siguientes campos para la creación de sus respectivos atributos:
  - IDPROGRAMACION → PROGRAMACION
  - FECHA
  - HORAINICIO
  - HORAFIN
  
- De la tabla dimensión “MEDICO”, se tomó los siguientes campos para la creación de sus respectivos atributos:
  - IDMEDICO → MEDICO
  - APELLIDOPATerno
  - APELLIDOMATERNO
  - NOMBRES
  
- De la tabla dimensión “SERVICIO”, se tomó los siguientes campos para la creación de sus respectivos atributos:
  - NOMBRE → SERVICIO
  
- De la tabla dimensión “FECHA”, se tomó los siguientes campos para la creación de sus respectivos atributos:
  - IDFECHA → FECHA
  - AÑO
  - MES
  - DIA

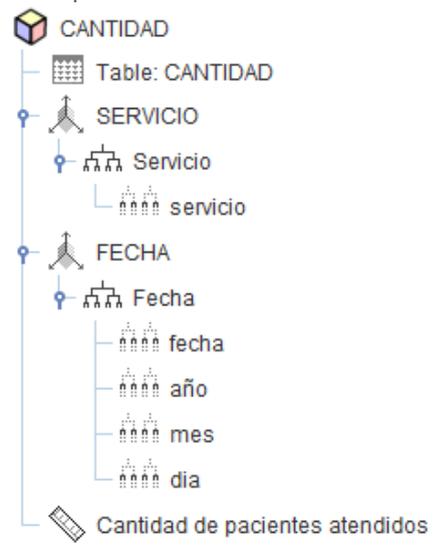
- De la tabla dimensión “PACIENTE”, se tomó los siguientes campos para la creación de sus respectivos atributos:
  - IDPACIENTE → PACIENTE
  - APELLIDOPATerno
  - APELLIDOMATERNO
  - NOMBRES
- De la tabla dimensión “ESPECIALIDAD”, se tomó los siguientes campos para la creación de sus respectivos atributos:
  - NOMBRE → ESPECIALIDAD



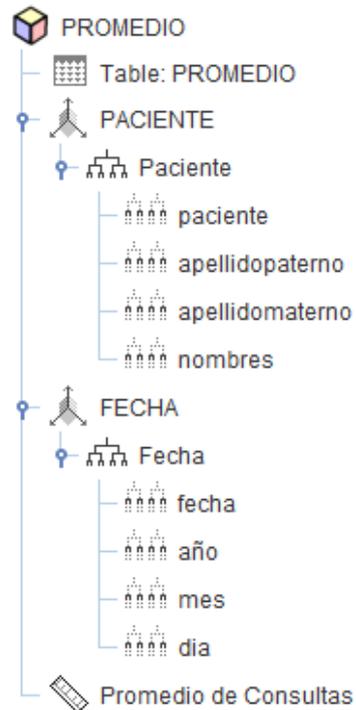
**Figura N° 39:** Cubo CUMPLIMIENTO, atributos.



**Figura N° 40:** Cubo PRODUCTIVIDAD, atributos.



**Figura N° 41:** Cubo CANTIDAD, atributos.



**Figura N° 42:** Cubo CANTIDAD, atributos.

A continuación, luego de haber desarrollado la Metodología Hefesto tenemos como resultado en esta fase de Procesamiento de Datos lo que se detalla a continuación:

**Requerimientos:**

- **R1:** Productividad hora médico por cada hora programada de trabajo del médico.
- **R2:** Cantidad de pacientes atendidos por cada servicio en un período de tiempo.
- **R3:** Promedio de consultas que se han otorgado a cada paciente que acudió al hospital durante un período de tiempo.
- **R4:** Grado de cumplimiento de la programación de las horas médico por servicio o especialidad en un periodo de tiempo.
- **R5:** Promedio de permanencia de los pacientes que egresaron en un periodo de tiempo.

### **Indicadores y perspectivas:**

En síntesis, los indicadores que se utilizaron y perspectivas de análisis son:

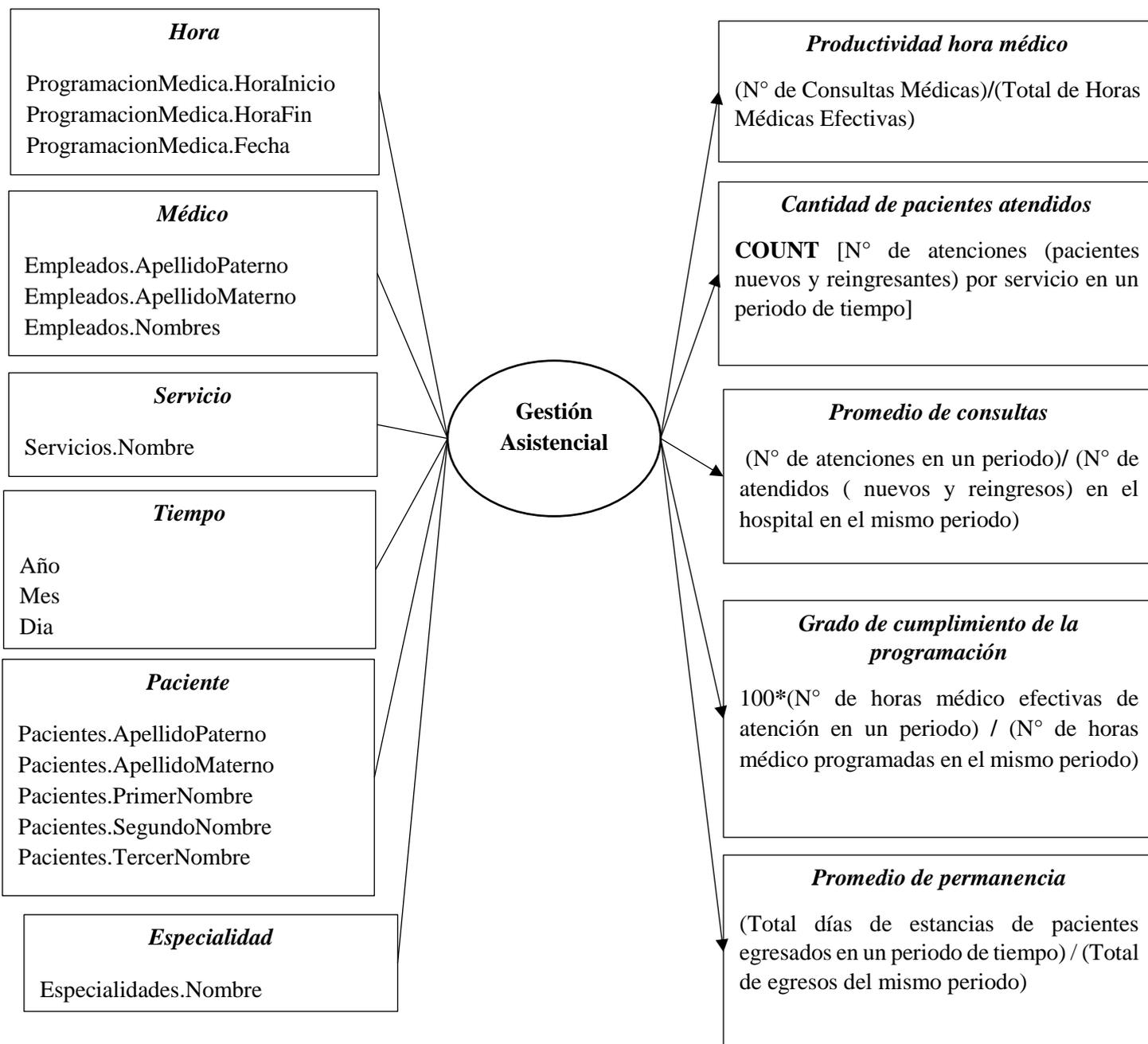
#### **Indicadores:**

- ✓ Productividad hora médico
- ✓ Cantidad de pacientes atendidos
- ✓ Promedio de consultas
- ✓ Grado de cumplimiento de la programación
- ✓ Promedio de permanencia

#### **Perspectivas:**

- ✓ Hora
- ✓ Médico
- ✓ Servicio
- ✓ Tiempo
- ✓ Paciente
- ✓ Especialidad

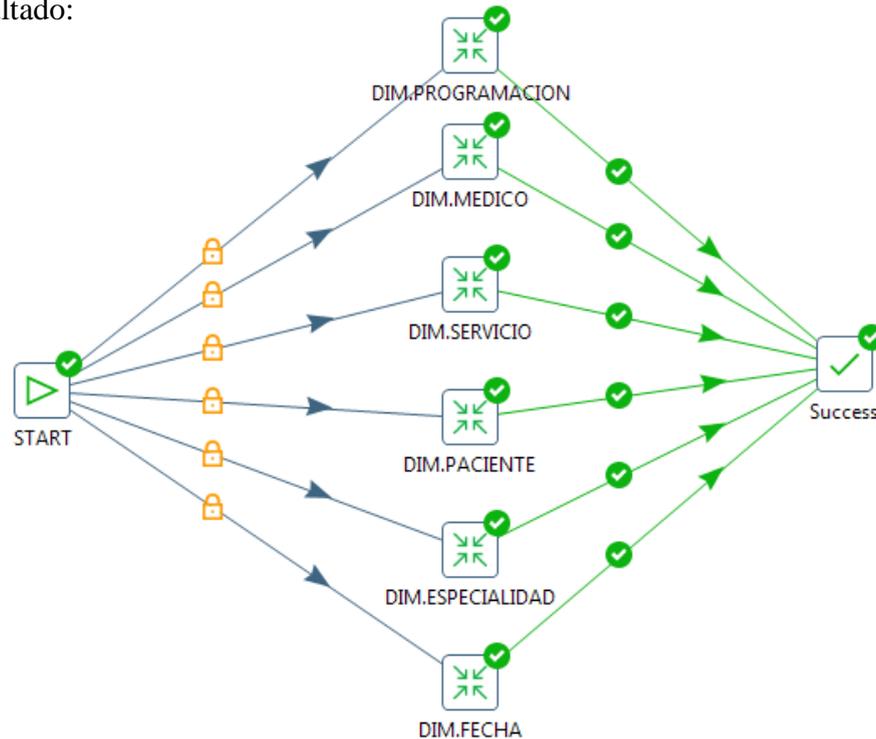
Una vez analizadas las fuentes OLTP para determinar cómo serán calculados los indicadores. Luego, se definirán qué campos se incluirán en cada perspectiva. Finalmente, se amplió el Modelo Conceptual con la información obtenida.



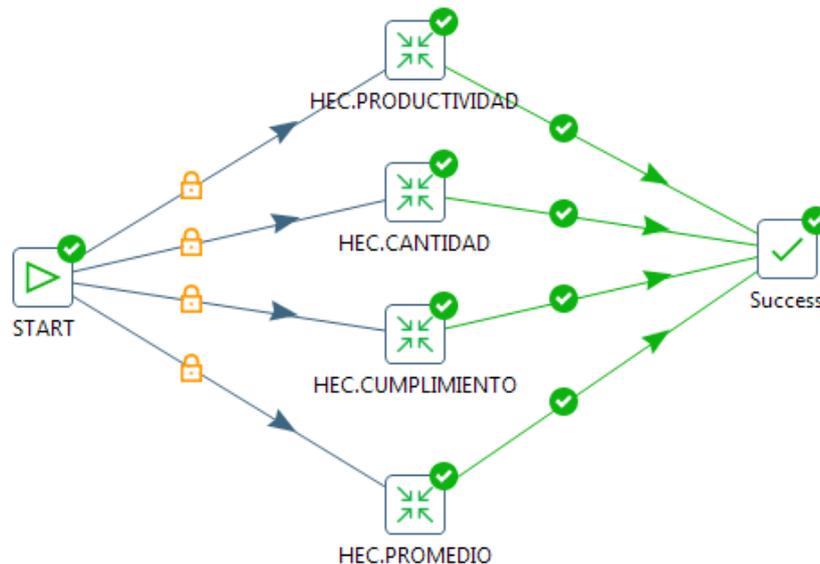
**Figura N° 43:** Modelo Conceptual del Hospital La Caleta.

### Modelado lógico del Data Warehouse

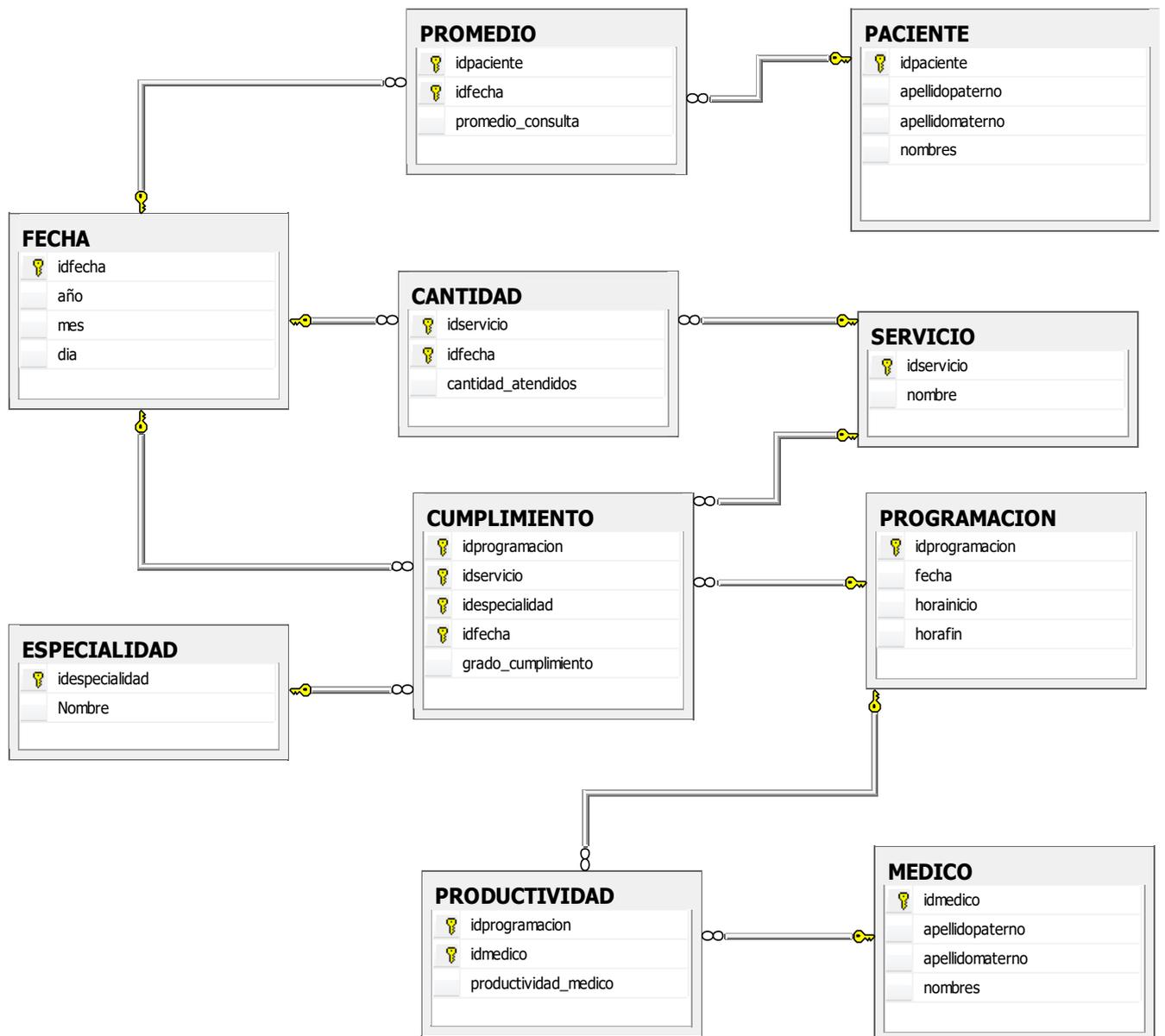
A continuación, teniendo como base el modelo conceptual anterior obtenido se confecciona el modelo lógico de la estructura del DW. Para ello, se definió el tipo de modelo que se utilizaría y se diseñó las tablas de dimensiones y de hechos en la base de datos SQL SERVER 2008. Se crearon haciendo uso de la herramienta PENTAHO. Finalmente, se realizaron las uniones pertinentes entre estas tablas, obteniendo como resultado:



**Figura N° 44:** PENTAHO – Trabajo de Carga de Dimensiones.



**Figura N° 45:** PENTAHO – Trabajo de Carga de Hechos.



**Figura N° 46:** Modelo Entidad Relación Físico, obtenido como resultado del Modelo Lógico del DW del Hospital La Caleta.

#### 4.1.4 Análisis y producción.

El componente de Pentaho utilizado para la creación de reportes con los indicadores del Hospital La Caleta, Analisis Report, el mismo que mediante conexión al Data Warehouse creado, nos permitió dar solución a dichos indicadores.

Los reportes generados con Pentaho Report Designer, nos han permitido dar solución a los indicadores:

##### Productividad hora médico

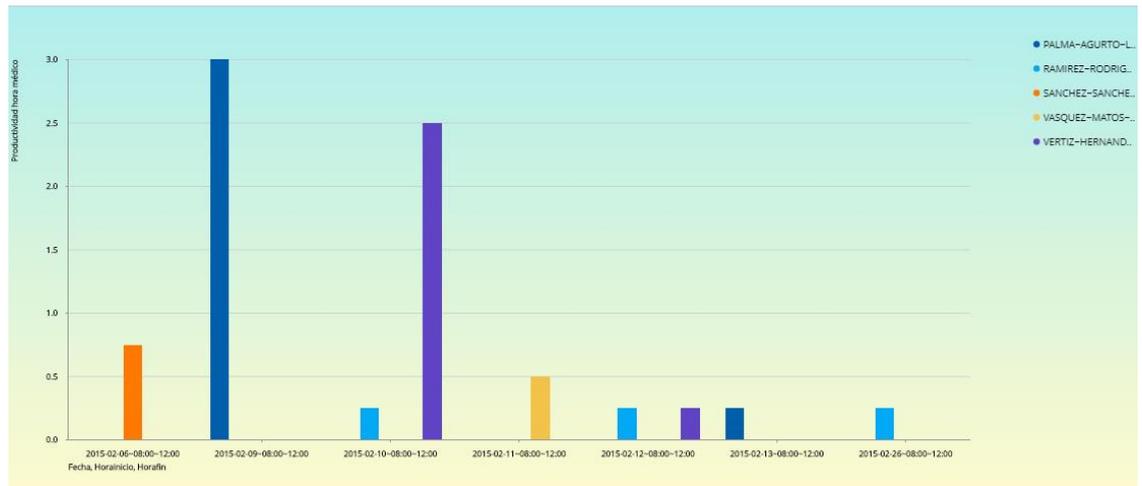
- Productividad hora médico por cada hora programada de trabajo del médico.

Perspectivas:

- **Programación:** Horas Programadas durante el mes de marzo
- **Medico :** Todos los médicos con programación en el mes de marzo

			Apellidopaterno				
			Apellidomaterno				
			Nombres				
			PALMA	RAMIREZ	SANCHEZ	VASQUEZ	VERTIZ
			AGURTO	RODRIGUEZ	SANCHEZ	MATOS	HERNANDEZ
			LUIS GONZAGA	VIRGILIO SANTOS	CARLOS ALBERTO	DORE MANZONI	MOISES ANDRES
Fecha	Horainicio	Horafin	Productividad hora médico				
2015-02-06	08:00	12:00	-	-	0.75	-	-
2015-02-09	08:00	12:00	3	-	-	-	-
2015-02-10	08:00	12:00	-	0.25	-	-	2.5
2015-02-11	08:00	12:00	-	-	-	0.5	-
2015-02-12	08:00	12:00	-	0.25	-	-	0.25
2015-02-13	08:00	12:00	0.25	-	-	-	-
2015-02-26	08:00	12:00	-	0.25	-	-	-

**Figura N° 47:** La tabla de productividad hora medico es el resultado del número de atenciones que realiza cada médico por cada hora efectiva de trabajo en base a su programación.



**Figura N° 48:** La gráfica de productividad hora médico, permite identificar rápidamente que medico es más productivo en base a su programación y en un periodo determinado.

### Cantidad de pacientes atendidos

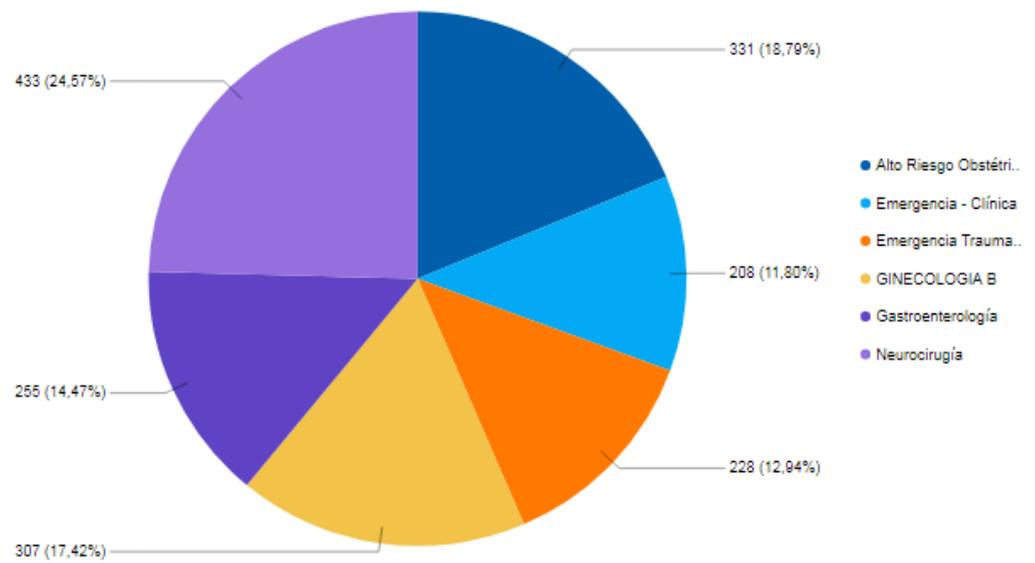
- Cantidad de pacientes atendidos por cada servicio en un período de tiempo.

Perspectivas:

- **Servicio:** Todos los Servicios
- **Periodo:** Enero- Marzo del 2015

Servicio	Cantidad de Pacientes Atendidos
Alto Riesgo Obstétrico II	331
Emergencia - Clínica	208
Emergencia Traumatolo...	228
GINECOLOGIA B	307
Gastroenterología	255
Neurocirugía	433

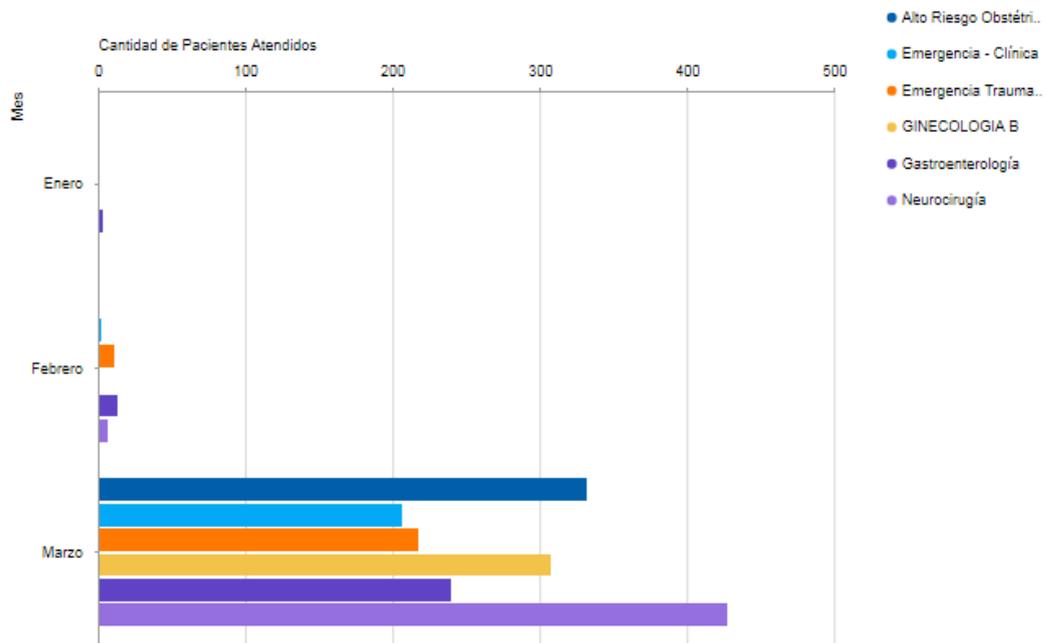
**Figura N° 49:** La tabla de total de pacientes atendidos, es la cantidad de atenciones a pacientes (nuevos y reingresos) por servicio en un determinado periodo de tiempo.



**Figura N° 50:** La gráfica de total de pacientes atendidos por servicio, permite identificar el servicio con mayor afluencia en hospital en un determinado periodo.

Servicio	Mes		
	Enero	Febrero	Marzo
	Cantidad de Pacientes Atendidos	Cantidad de Pacientes Atendidos	Cantidad de Pacientes Atendidos
Alto Riesgo Obstétrico II	-	-	331
Emergencia - Clínica	-	2	206
Emergencia Traumatolo...	-	11	217
GINECOLOGIA B	-	-	307
Gastroenterología	3	13	239
Neurocirugía	-	6	427

**Figura N° 51:** La tabla de pacientes atendidos detallado por servicio y mes, es la cantidad de atenciones a pacientes (nuevos y reingresos) por servicio mensualmente en un determinado periodo de tiempo.



**Figura N° 52:** La gráfica de pacientes atendidos detallado por servicio y mes, muestra en que mes hubo más atenciones así como cuál es el servicio que tuvo mayor afluencia en dicho mes.

### Promedio de consultas

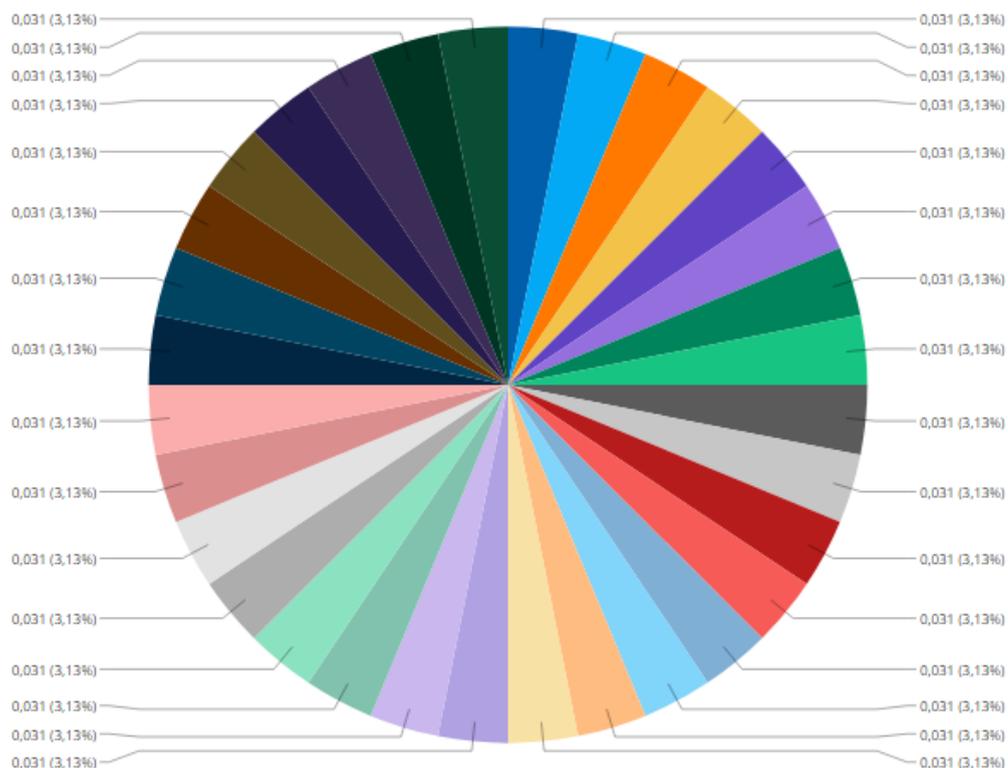
- Promedio de consultas que se han otorgado a cada paciente que acudió al hospital durante un período de tiempo.

Perspectivas:

- **Paciente:** Todos los Pacientes
- **Periodo:** Febrero del 2015

Idpacier...	Nombres	Apellidoater...	Apellidoater...	Promedio consulta
4276	EMILIO	SANDOVAL	INCIO	0,031
4957	RAFAEL	PATIÑO	ESPINOZA	0,031
11186	INOCENTA Olga	ALVAREZ	LUJAN	0,031
14535	ZULLY LILIANA	ORDOÑEZ	REYNA	0,031
20543	FERNANDO	LIÑAN	FLORES	0,031
26477	INOCENTA TEOFILA	GOMEZ	TREJO	0,031
30875	KARINA LIZET	OROSCO	ARTEAGA	0,031
34794	BETHY JANETH	MENDEZ	PEREDA	0,031
43425	JOSE SANTOS	FIESTAS	GALAN	0,031
45708	EDWAR MANUEL	CASTILLO	CUSTODIO	0,031
93005	GEAN CARLOS	BOM	CALDERON	0,031
98061	MAXIMINA	MARTINEZ	FLORES	0,031
125332	PERCY	GARCIA	BASURTO	0,031
134722	MANUEL	ROJAS	RODRIGUEZ	0,031
144128	RODRIGO Oquendo	PAREDES	QUISPE	0,031
187618	LUIS Santiago	FLORES	VASQUEZ	0,031
227789	MIGUEL	VILCHEZ	MARGARITO	0,031
238578	SANTOS	HERRERA	HOYOS	0,031
242661	JUVITZA	LEON	ALVA	0,031
243389	ELISABETH	HOYOS	VASQUEZ	0,031
259416	RRNN	PONCE	ASIAN	0,031
260845	IGNACIA GUADALUPE	PEÑA	DE FERNANDEZ	0,031
262983	EDY YOVANNI	OQUEÑA	BALLADARES	0,031
263519	MIGUEL ANGEL	BLAS	CASTILLO	0,031
264899	ROMMEL WILLIAM	QUEZADA	MEJIA	0,031
266351	ARNULFO JEAN PAUL	ALEJOS	CALIXTRO	0,031
277754	RN	CABREJOS	DE LA CRUZ	0,031
280198	RN	BARRERA	SUMARY	0,031
280852	RICARDO David	RODRIGUEZ	FIERRO	0,031
281848	AMELIA	MANRIQUE	REMIGIO	0,031
284794	STEFANY Rubi	CASTILLO	BENITES	0,031
284926	APOLONIA Alejandri...	ACOSTA	VALLE	0,031

**Figura N° 53:** La tabla de concentración de consultas, indica el número de consultas en promedio que se le ha otorgado a cada paciente en un periodo de tiempo.



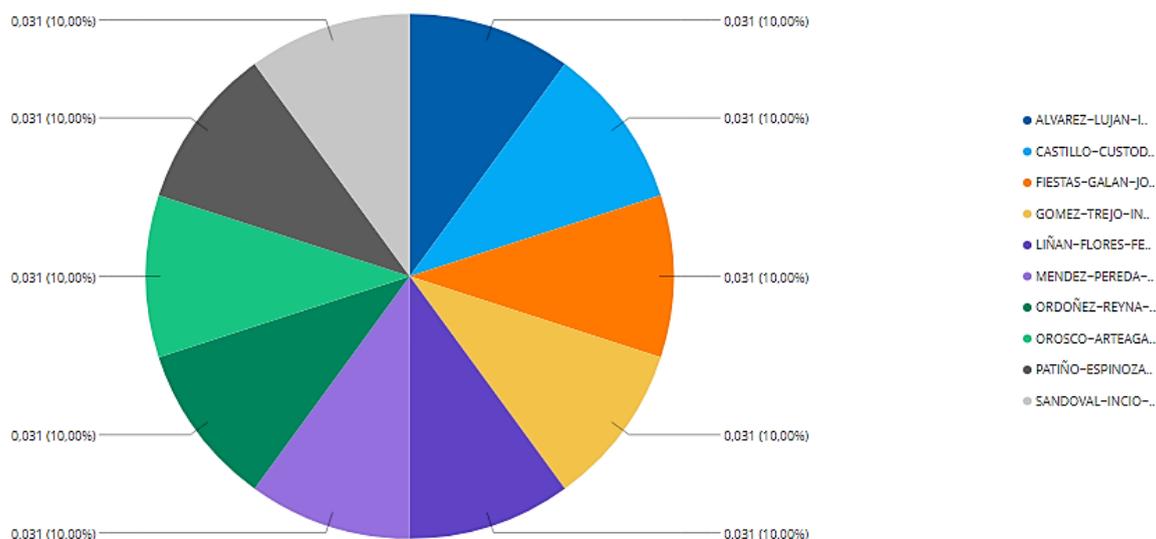
**Figura N° 54:** La gráfica de concentración de consultas, muestra el promedio de atenciones por cada paciente que acudió al hospital durante un periodo.

Perspectivas:

- **Paciente:** Los Primeros 10 pacientes atendidos en este periodo.
- **Periodo:** Febrero del 2015

Idpacien...	Nombres	Apellidoater...	Apellidoater...	Promedio consulta
4276	EMILIO	SANDOVAL	INCIO	0,031
4957	RAFAEL	PATIÑO	ESPINOZA	0,031
11186	INOCENTA Olga	ALVAREZ	LUJAN	0,031
14535	ZULLY LILIANA	ORDOÑEZ	REYNA	0,031
20543	FERNANDO	LIÑAN	FLORES	0,031
26477	INOCENTA TEOFIL...	GOMEZ	TREJO	0,031
30875	KARINA LIZET	OROSCO	ARTEAGA	0,031
34794	BETHY JANETH	MENDEZ	PEREDA	0,031
43425	JOSE SANTOS	FIESTAS	GALAN	0,031
45708	EDWAR MANUEL	CASTILLO	CUSTODIO	0,031

**Figura N° 55:** La tabla de concentración de consultas de los 10 primeros pacientes, indica el número de consultas en promedio que se le ha otorgado a cada paciente en el periodo indicado.



**Figura N° 56:** La gráfica de concentración de consultas de los 10 primeros pacientes, muestra el promedio de atenciones que cada uno durante un periodo de tiempo.

**Grado de cumplimiento de la programación**

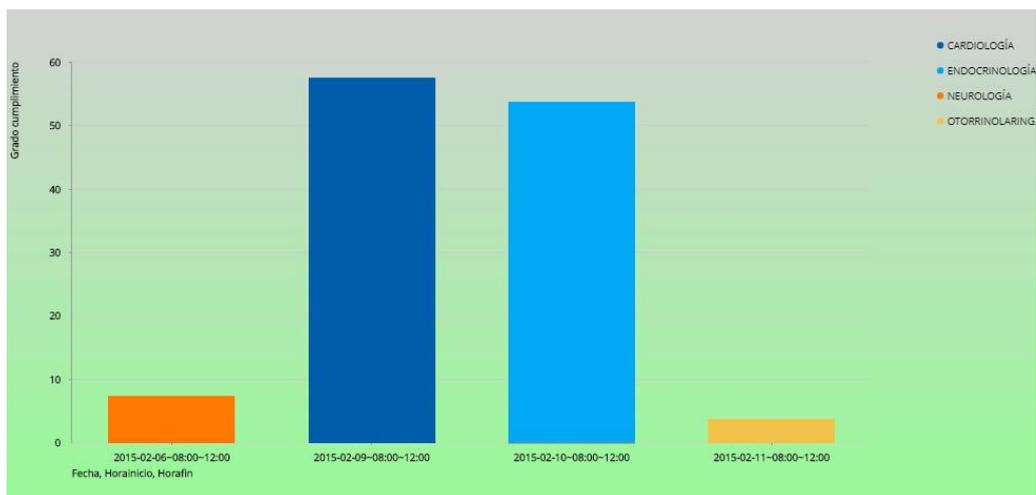
- Grado de cumplimiento de la programación de las horas médico por servicio o especialidad en un periodo de tiempo.

Perspectivas:

- **Programación:** Programación de las horas médico en el periodo.
- **Servicio:** Todos los servicios.
- **Periodo:** Febrero del 2015

Fecha	Horainicio	Horafin	Servicio			
			CARDIOLOGÍA	ENDOCRINOLOGÍA	NEUROLOGÍA	OTORRINOLARIN...
			Grado cumplimiento	Grado cumplimiento	Grado cumplimiento	Grado cumplimiento
2015-02-06	08:00	12:00	-	-	7.5	-
2015-02-09	08:00	12:00	57.5	-	-	-
2015-02-10	08:00	12:00	-	53.75	0	-
2015-02-11	08:00	12:00	-	-	-	3.75
2015-02-12	08:00	12:00	-	0	0	-
2015-02-13	08:00	12:00	0	-	-	-
2015-02-26	08:00	12:00	-	-	0	-

**Figura N° 57:** La tabla de grado de cumplimiento de la programación, muestra porcentaje de cumplimiento de las horas médico programadas por servicio en un periodo de tiempo.



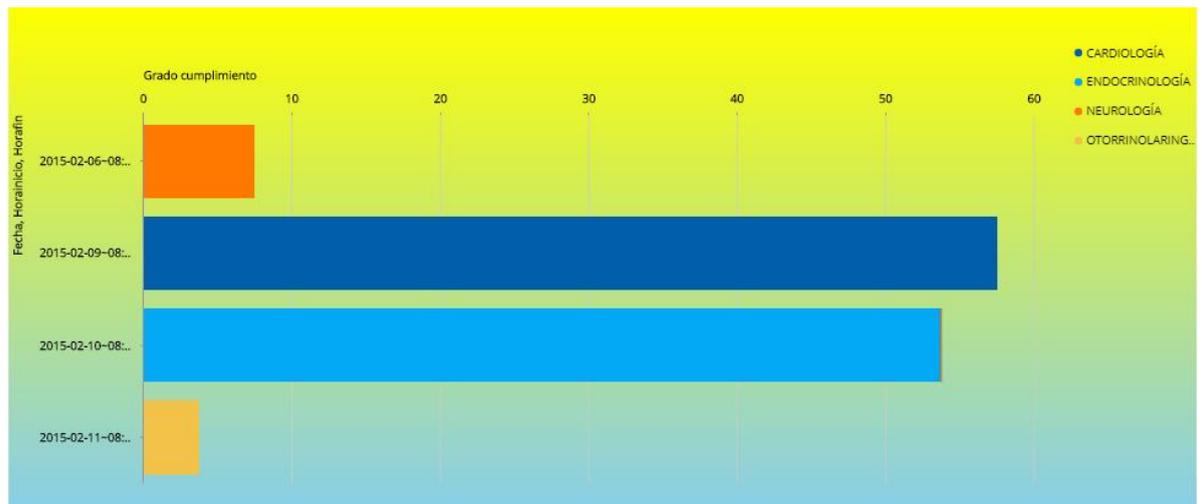
**Figura N° 58:** La gráfica grado de cumplimiento de la programación, permite identificar a simple vista que programación tuvo mayor cumplimiento, en un periodo de tiempo y a qué servicio corresponde.

Perspectivas:

- **Programación:** Programación de las horas médico en el periodo.
- **Especialidad:** Todas las especialidades.
- **Periodo:** Febrero del 2015

Fecha	Horainicio	Horafin	Especialidad			
			CARDIOLOGÍA	ENDOCRINOLOGÍA	NEUROLOGÍA	OTORRINOLARIN...
			Grado cumplimiento	Grado cumplimiento	Grado cumplimiento	Grado cumplimiento
2015-02-06	08:00	12:00	-	-	7.5	-
2015-02-09	08:00	12:00	57.5	-	-	-
2015-02-10	08:00	12:00	-	53.75	0	-
2015-02-11	08:00	12:00	-	-	-	3.75
2015-02-12	08:00	12:00	-	0	0	-
2015-02-13	08:00	12:00	0	-	-	-
2015-02-26	08:00	12:00	-	-	0	-

**Figura N° 59:** La tabla de grado de cumplimiento de la programación, muestra porcentaje de cumplimiento de las horas médico programadas por especialidad en un periodo de tiempo.



**Figura N° 60:** La gráfica grado de cumplimiento de la programación, permite identificar a simple vista que programación tuvo mayor cumplimiento y a qué especialidad corresponde todo esto en un periodo de tiempo.

#### 4.1.5 Difusión

Una vez que se ha diseñado los cubos OLAP (Schema Worbench), y los reportes (Analisys Report), Pentaho ofrece un módulo adicional denominado Pentaho BI Server, el mismo que mediante un entorno web, nos permitió mostrar la información al usuario final, a través de la visualización de los cubos y reportes generados con las herramientas antes mencionadas.

Para la visualización de cubos OLAP, Pentaho BI Server posee complementos(plugins) como el visor denominado JPivot y Saiku Analisis, pero la visualización que estos generan no es muy amigable al usuario, por esa razón preferimos trabajar con el Analysis Report , el mismo que nos proporcionó una visión bastante mejorada de los cubos OLAP. Adicional a esto, se realizó la conexión a los Data Marts creados anteriormente, para posteriormente importar los cubos y reportes creados.

## 4.2 Discusión

### 4.2.1 Contrastación de la hipótesis.

La contrastación de la hipótesis se ha realizado de acuerdo al método propuesto Pretest-Postest para poder aceptar o rechazar la hipótesis. Así mismo para la realización de este diseño se identificaron indicadores cualitativos y cuantitativos para evaluar el rendimiento del sistema actual y de sistema propuesto mediante la inteligencia de negocios.

**Tabla N° 12:** Determinación de pruebas estadísticas de los indicadores.

Ítem	Indicadores	n	Criterios de decisión	Prueba Estadística
1	Nivel de satisfacción de los usuarios	6	$n \leq 30$	Prueba estadística T de Student, diferencia de medias.
2	Calidad de Información	6	$n \leq 30$	Prueba estadística T de Student, diferencia de medias.
3	Tiempo de generación de reportes que soportan la toma de decisiones	64	$n > 30$	Prueba estadística Distribución Normal Z, diferencias de medias.
4	Costo por la generación de reportes	64	$n > 30$	Prueba estadística Distribución Normal Z, diferencias de medias
5	Tiempo para analizar reportes y formular alternativas	36	$n > 30$	Prueba estadística Distribución Normal Z, diferencias de medias

**Fuente:** Cálculo de muestra.

Para asegurar la consistencia de las encuesta se realizó la confiabilidad con **Alfa de Cronbach** a una prueba piloto, obteniendo resultados dentro del rango de fiabilidad (Anexo). Ya que entre más cerca de 1 este  $\alpha$ , mas alto el grado de confiabilidad.

Indicador 1:  $\alpha = 0.94$

Indicador 2:  $\alpha = 0.84$

#### 4.2.1.1 Indicadores cualitativos.

Para contrastar la hipótesis se aplicó una encuesta a los usuarios internos del sistema. Han sido tabuladas, de manera que se calculen los resultados obtenidos de acuerdo a los rangos que se presentan a continuación. En la Tabla podemos ver el rango de valores para evaluar el nivel de satisfacción y la calidad en los usuarios internos.

**Tabla N° 13:** Rango de valores.

<b>Rango</b>	<b>Nivel de Aprobación</b>	<b>Peso</b>
AP	Aprobación Plena	5
AS	Aprobación Simple	4
IN	Indecisión o Indiferencia	3
DS	Desaprobación Simple	2
DP	Desaprobación Plena	1

A continuación tenemos a los usuarios involucrados en la emisión de reportes analíticos.

**Tabla N° 14:** Lista usuarios.

Usuarios Internos	
Dirección Ejecutiva	U1
Oficina de administración	U2
Oficina de Planeamiento Estratégico	U3
Área de Estadística e Informática	U4
Oficina de Gestión de la Calidad	U5
Órgano de Control Institucional	U6

Los valores se calcularon en base a las respuestas proporcionadas por los usuarios mencionados anteriormente, ya que ellos se encuentran inmersos en el manejo del sistema actual. Para realizar la ponderación correspondiente de las preguntas aplicadas en las encuestas se tomó como base la escala de Likert (rango de ponderación: [1-5]).

Para cada pregunta se contabilizó la frecuencia de ocurrencia para cada una de los posibles tipos de respuestas (5) por cada entrevistado (6), luego se calculó el puntaje total y puntaje promedio, como se detalla:

Se tiene:

$$PT_i = \sum_{j=1}^5 (F_{ij} \times P_j) \dots\dots\dots (1)$$

Donde:

$PT_j =$  Puntaje Total de la pregunta  $i -$  esima

$F_{ij} =$  Frecuencia  $j -$  esima de la Pregunta  $i -$  esima

$P_j =$  Peso  $j -$  esima

El cálculo del promedio ponderado por cada pregunta sería:

$$\overline{PP}_i = \frac{PT_i}{n} \dots\dots\dots (2)$$

$\overline{PP}_i =$  Promedio de Puntaje Total de la pregunta  $i$ -esima.

$n = 6$  número de usuarios del sistema.

**4.2.1.1.1** *Calculo para hallar el Nivel De Satisfacción De Los Usuarios.*

En la tabla N°15 podemos ver la ponderación de los criterios de evaluación del indicador cualitativo Nivel de Satisfacción del Usuario del Sistema.

**Tabla N° 15:** Ponderación del Usuario del Sistema - Nivel de Satisfacción.

N°.	Pregunta	PESO					Puntaje Total	Puntaje Promedio
		AP	AS	IN	DS	DP		
		5	4	3	2	1		
1	¿El sistema es de fácil uso?	P1,5	P1,4	P1,3	P1,2	P1,1	$\sum_{j=1}^5 (F_{1j} x P_j)$	$\frac{PT_1}{n}$
2	¿Está usted satisfecho con la variedad de reportes que ofrece el sistema?	P2,5	P2,4	P2,3	P2,2	P2,1	$\sum_{j=1}^5 (F_{2j} x P_j)$	$\frac{PT_2}{n}$
3	¿La interfaz del sistema es amigable para usted?	P3,5	P3,4	P3,3	P3,2	P3,1	$\sum_{j=1}^5 (F_{3j} x P_j)$	$\frac{PT_3}{n}$
4	¿La información que le brinda la solución de BI le ayuda a tener una mejor visión de la situación de la institución?	P4,5	P4,4	P4,3	P4,2	P4,1	$\sum_{j=1}^5 (F_{4j} x P_j)$	$\frac{PT_4}{n}$
5	¿Cómo califica la flexibilidad de la solución de BI al momento de realizar reportes analíticos?	P5,5	P5,4	P5,3	P5,2	P5,1	$\sum_{j=1}^5 (F_{5j} x P_j)$	$\frac{PT_5}{n}$
6	¿Los reportes solicitados al área de estadística se entregan a tiempo?	P6,5	P6,4	P6,3	P6,2	P6,1	$\sum_{j=1}^5 (F_{6j} x P_j)$	$\frac{PT_6}{n}$
7	¿Cómo considera la manera de presentar sus reportes analíticos?	P7,5	P7,4	P7,3	P7,2	P7,1	$\sum_{j=1}^5 (F_{7j} x P_j)$	$\frac{PT_7}{n}$
8	¿Cómo considera el tiempo de elaboración de reportes personalizados por parte del área de estadística?	P8,5	P8,4	P8,3	P8,2	P8,1	$\sum_{j=1}^5 (F_{8j} x P_j)$	$\frac{PT_8}{n}$
9	¿Cómo califica usted la calidad de los reportes entregados a la Dirección Ejecutiva?	P9,5	P9,4	P9,3	P9,2	P9,1	$\sum_{j=1}^5 (F_{9j} x P_j)$	$\frac{PT_9}{n}$
10	¿Cómo califica la flexibilidad de la solución de BI al momento de realizar reportes analíticos?	P10,5	P10,4	P10,3	P10,2	P10,1	$\sum_{j=1}^5 (F_{10j} x P_j)$	$\frac{PT_{10}}{n}$

En la tabla N° 16 podemos ver la ponderación de los criterios de evaluación del indicador cualitativo nivel de satisfacción del usuario interno con los valores obtenidos en las encuestas realizadas.

**Tabla N° 16:** Tabulación usuario del sistema Pre Test - Nivel de Satisfacción

	N°	Pregunta	AP	AS	IN	DS	DP	Puntaje	Puntaje
			5	4	3	2	1	Total	Promedio
FUNCIONALIDAD	1	¿El sistema es de fácil uso?			2	2	2	12	2.00
	2	¿Está usted satisfecho con la variedad de reportes que ofrece el sistema?			2	3	1	13	2.17
	3	¿La interfaz del sistema es amigable para usted?				4	2	10	1.67
	4	¿La información que le brinda el sistema le ayuda a tener una mejor visión de la situación de la institución?			1	3	2	11	1.83
	5	¿Cómo califica la flexibilidad del sistema al momento de realizar reportes analíticos?			2	2	2	12	2.00
SERVICIO Y CALIDAD	6	¿Los reportes solicitados al área de estadística se entregan a tiempo?			2	2	2	12	2.00
	7	¿Cómo considera la manera de presentar sus reportes analíticos?		1	2	2	1	15	2.50
	8	¿Cómo considera el tiempo de elaboración de reportes personalizados por parte del área de estadística?			2	1	3	11	1.83
	9	¿Cómo califica usted la calidad de los reportes entregados a la Dirección Ejecutiva?		1	2	3		16	2.67
	10	¿Los reportes emitidos son fáciles de analizar?			3	2	1	14	2.33

**Calculo para hallar el nivel de satisfacción de los usuarios con la Solución de BI Propuesta.**

Las encuestas mostradas fueron aplicadas a los usuarios del sistema, han sido tabuladas de manera que se calculen los resultados obtenidos de acuerdo a la tabla N° 15.

Cada tipo de respuesta de las encuestas aplicadas tiene un peso. Luego se procede a hallar el puntaje promedio de cada criterio usado por cada indicador. Para finalmente hallar el puntaje total por cada indicador con las formulas empleadas anteriormente,

A continuación en la tabla N° 17 se muestra los resultados de la encuesta aplicada para conocer el nivel de satisfacción de los usuarios con la solución de BI propuesta.

**Tabla N° 17:** Tabulación usuario del sistema Post Test - Nivel de Satisfacción

N°	Pregunta	AP	AS	IN	DS	DP	Puntaje	Puntaje	
		5	4	3	2	1	Total	Promedio	
FUNCIONALIDAD	1	¿La solución de BI es de fácil uso?	5	1				29	4.83
	2	¿Está usted satisfecho con la variedad de reportes que ofrece la solución de BI?	5	1				29	4.83
	3	¿La interfaz de la solución de BI es amigable para usted?	4	2				28	4.67
	4	¿La información que le brinda el sistema le ayuda a tener una mejor visión de la situación de la institución?	5	1				29	4.83
	5	¿Cómo califica la flexibilidad del sistema al momento de realizar reportes analíticos?	4	2				28	4.67
SERVICIO Y CALIDAD	6	¿Los reportes solicitados al área de estadística se entregan a tiempo?	4	2				28	4.67
	7	¿Cómo considera la manera de presentar sus reportes analíticos?	4	1	1			27	4.50
	8	¿Cómo considera el tiempo de elaboración de reportes personalizados por parte del área de estadística?	4	2				28	4.67
	9	¿Cómo califica usted la calidad de los reportes entregados a la Dirección Ejecutiva?	5	1				29	4.83
	10	¿Los reportes emitidos son fáciles de analizar?	5	1				29	4.83

Podemos ver en la tabla N° 18 la contrastación de los resultados de las pruebas realizadas Pre y Post Test.

**Tabla N° 18:** Contrastación Pre & Post Test - Nivel de Satisfacción.

	Pregunta	PRE TEST	POST TEST	$D_i$	$D_i^2$
		$NSUA_i$	$NSUP_i$		
FUNCIONALIDAD	1	2.00	4.83	-2.83	8.01
	2	2.17	4.83	-2.66	7.08
	3	1.67	4.67	-3.00	9.00
	4	1.83	4.83	-3.00	9.00
	5	2.00	4.67	-2.67	7.13
SERVICIO Y CALIDAD	6	2.00	4.67	-2.67	7.13
	7	2.50	4.50	-2.00	4.00
	8	1.83	4.67	-2.84	8.07
	9	2.67	4.83	-2.16	4.67
	10	2.33	4.83	-2.50	6.25
			$\Sigma$	<b>-26.33</b>	<b>70.32</b>

Donde:

$NSU_A$ =Nivel de satisfacción del usuario con el sistema actual

$NSU_P$ =Nivel de satisfacción del usuario con la solución de BI propuesta.

## Prueba de la Hipótesis para el Indicador Cualitativo:

### Nivel de Satisfacción del Usuario del Sistema

#### A. Definición de Variables

$NSU_A$ =Nivel de satisfacción del usuario con el sistema actual

$NSU_P$ =Nivel de satisfacción del usuario con la solución de BI propuesta.

#### B. Hipótesis Estadísticas

- **Hipótesis  $H_0$ :** El nivel de satisfacción del usuario con el Sistema de Trabajo Actual es mayor o igual que el Nivel de satisfacción de los usuarios finales con la Solución de BI Propuesto.

$$H_0: NSU_A - NSU_P \geq 0$$

- **Hipótesis  $H_a$ :** El Nivel de satisfacción del usuario con el Sistema de Trabajo Actual es menor que el Nivel de satisfacción de los usuarios finales con la Solución de BI Propuesto.

$$H_a: NSU_A - NSU_P < 0$$

### C. Nivel de Significancia

El Nivel de significancia ( $\alpha$ ) escogido para la prueba de la hipótesis es del 5%.

Siendo ( $\alpha$ )=0.05 (nivel de significancia) y  $n-1(10-1)=9$  grados de libertad, se tiene el valor crítico de T de Student (Ver tabla T Student en Anexo):

**Valor Crítico:**

$$t_{\alpha} = 0.05 = -1.833$$

Como  $\alpha = 0.05$  y  $n-1 = 10-1 = 9$  grados de libertad, la región de rechazo consiste en aquellos valores de  $t$  menores que:

$$-t_{0.05} = -1.833$$

### D. Resultados de la Hipótesis Estadística

a) Diferencia Promedio

$$\bar{D} = \frac{\sum_{i=1}^n D_i}{n} \dots \dots \dots (3)$$

$$\bar{D} = \frac{-26.33}{10}$$

$$\bar{D} = -2.63$$

**b) Desviación Estándar**

$$S_{D^2} = \frac{n \sum_{i=1}^n D_i^2 - [\sum_{i=1}^n D_i]^2}{n(n-1)} \dots(4)$$

$$S_{D^2} = \frac{10(70.32) - (-26.33)^2}{10(10 - 1)}$$

$$S_{D^2} = 0.11$$

**c) Calculo de T:**

$$t = \frac{\bar{D}\sqrt{n}}{\sqrt{S_D}}$$

$$t = \frac{(-2.63)\sqrt{10}}{\sqrt{0.33}}$$

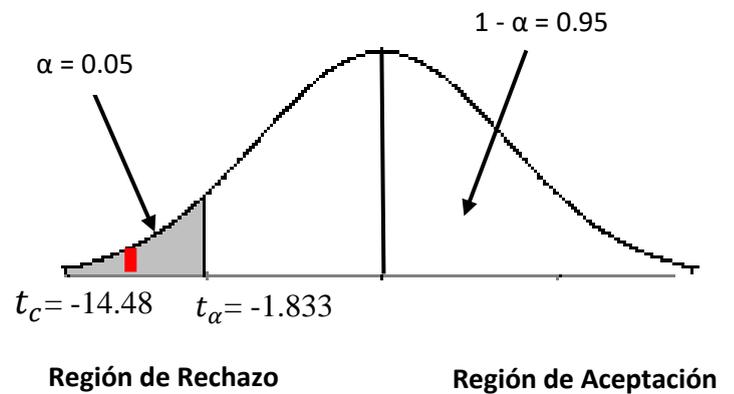
$$t = -14.48$$

**d) Conclusión:**

Puesto que:  $t_c = -14.48$  ( $t_{calculado}$ ) <  $t_\alpha = -1.833$  ( $t_{tabular}$ ), y estando este valor dentro de la región de rechazo, se concluye que  $NSU_A - NSU_P < 0$ , se rechaza  $H_0$  y  $H_a$  es aceptada, por lo tanto se prueba la validez de la hipótesis con un nivel de error de 5% ( $\alpha=0.05$ ), siendo la Solución de BI es

una alternativa de solución para el problema de investigación.

En la Figura podemos ver la región de aceptación y rechazo para la prueba de la hipótesis Nivel de satisfacción del usuario final.



**Figura N° 61:** Zona de Aceptación y Rechazo – Nivel de Aceptación

#### 4.2.1.1.2 Cálculo para medir la Calidad de la información.

En la tabla N° 19 podemos ver la ponderación de los criterios de evaluación del indicador cualitativo Calidad de la Información.

**Tabla N° 19:** Ponderación del Usuario del Sistema – Calidad de la Información.

N°.	Pregunta	PESO					Puntaje Total	Puntaje Promedio
		AP	AS	IN	DS	DP		
		5	4	3	2	1		
1	¿Se le ofrece datos relevantes para la toma de decisiones?	P1,5	P1,4	P1,3	P1,2	P1,1	$\sum_{j=1}^5 (F_{1j} \times P_j)$	$\frac{PT_1}{n}$
2	¿Considera que los datos brindados cuentan con el nivel de detalle que requiere?	P2,5	P2,4	P2,3	P2,2	P2,1	$\sum_{j=1}^5 (F_{2j} \times P_j)$	$\frac{PT_2}{n}$
3	¿Hay exactitud en los datos?	P3,5	P3,4	P3,3	P3,2	P3,1	$\sum_{j=1}^5 (F_{3j} \times P_j)$	$\frac{PT_3}{n}$
4	¿La información recibida tiene datos actualizados?	P4,5	P4,4	P4,3	P4,2	P4,1	$\sum_{j=1}^5 (F_{4j} \times P_j)$	$\frac{PT_4}{n}$
5	¿Es fácil la comprensión de los datos?	P5,5	P5,4	P5,3	P5,2	P5,1	$\sum_{j=1}^5 (F_{5j} \times P_j)$	$\frac{PT_5}{n}$
6	¿La información brindada para la toma de decisiones se le brinda completa?	P6,5	P6,4	P6,3	P6,2	P6,1	$\sum_{j=1}^5 (F_{6j} \times P_j)$	$\frac{PT_6}{n}$

En la tabla N° 20 podemos ver la ponderación de los criterios de evaluación del indicador cualitativo Calidad de la Información interno con los valores obtenidos en las encuestas realizadas.

**Tabla N° 20:** Tabulación usuario del sistema Pre Test - Calidad de la Información

N°	Pregunta	AP	AS	IN	DS	DP	Puntaje	Puntaje
		5	4	3	2	1	Total	Promedio
1	¿Se le ofrece datos relevantes para la toma de decisiones?				4	2	9	1.50
2	¿Considera que los datos brindados cuentan con el nivel de detalle que requiere?			1	3	2	11	1.83
3	¿Hay exactitud en los datos?				4	2	10	1.67
4	¿La información recibida tiene datos actualizados?			1	3	2	11	1.83
5	¿Es fácil la comprensión de los datos?			2	2	2	12	2.00
6	¿La información brindada para la toma de decisiones se le brinda completa?				4	2	10	1.67

**Calculo para medir la Calidad de la Información con la Solución de BI Propuesta.**

Las encuestas mostradas que fueron aplicadas a los usuarios del sistema, han sido tabuladas de manera que se calculen los resultados obtenidos de acuerdo a la tabla N° 19.

Cada tipo de respuesta de las encuestas aplicadas tiene un peso. Luego se procede a hallar el puntaje promedio de cada criterio usado por cada indicador. Para finalmente hallar el puntaje total por cada indicador con las formulas empleadas anteriormente.

A continuación en la tabla N° 21 se muestra los resultados de la encuesta aplicada para conocer la calidad de la información con la solución de BI propuesta.

**Tabla N° 21:** Tabulación usuario del sistema Post Test – Calidad de la Información.

N°	Pregunta	AP	AS	IN	DS	DP	Puntaje	Puntaje
		5	4	3	2	1	Total	Promedio
1	¿Se le ofrece datos relevantes para la toma de decisiones?	5	1				29	4.83
2	¿Considera que los datos brindados cuentan con el nivel de detalle que requiere?	4	2				28	4.67
3	¿Hay exactitud en los datos?	4	2				28	4.67
4	¿La información recibida tiene datos actualizados?	5	1				29	4.83
5	¿Es fácil la comprensión de los datos?	5	1				29	4.83
6	¿La información brindada para la toma de decisiones se le brinda completa?	4	2				28	4.67

Podemos ver en la tabla N° 22 la contrastación de los resultados de las pruebas realizadas Pre y Post Test.

**Tabla N° 22:** Contrastación Pre & Post Test – Calidad de la Información.

Pregunta	PRE TEST	POST TEST	$D_i$	$D_i^2$
	$CIA_i$	$CIP_i$		
1	1.50	4.83	-3.33	11.09
2	1.83	4.67	-2.84	8.07
3	1.67	4.67	-3.00	9.00
4	1.83	4.83	-3.00	9.00
5	2.00	4.83	-2.83	8.01
6	1.67	4.67	-3.00	9.00
		$\Sigma$	<b>-18.00</b>	<b>54.16</b>

Donde:

$CIA$ =Calidad de la Información con el sistema actual

$CIP$ = Calidad de la Información con la solución de BI propuesta.

**Prueba de la Hipótesis para el Indicador  
Cualitativo:**

**Calidad de la Información**

**A. Definición de Variables**

$CI_A$ =Calidad de la Información con el sistema actual

$CI_P$ = Calidad de la Información con la solución de BI propuesta.

**B. Hipótesis Estadísticas**

- **Hipótesis  $H_0$ :** La Calidad de la Información con el Sistema de Trabajo Actual es mayor o igual que la Calidad de la Información con la Solución de BI Propuesto.

$$H_0: CI_A - CI_P \geq 0$$

- **Hipótesis  $H_a$ :** La Calidad de la Información con el Sistema de Trabajo Actual es menor que la Calidad de la Información con la Solución de BI Propuesto.

$$H_a: CI_A - CI_P < 0$$

### C. Nivel de Significancia

El Nivel de significancia ( $\alpha$ ) escogido para la prueba de la hipótesis es del 5%.

Siendo ( $\alpha$ )=0.05 (nivel de significancia) y  $n-1(6-1)=5$  grados de libertad, se tiene el valor crítico de T de Student (Ver tabla T Student en Anexo):

**Valor Crítico:**

$$t_{\alpha} = 0.05 = -2.0150$$

Como  $\alpha = 0.05$  y  $n-1 = 6-1 = 5$  grados de libertad, la región de rechazo consiste en aquellos valores de  $t$  menores que:

$$-t_{0.05} = -2.0150$$

### D. Resultados de la Hipótesis Estadística

a) **Diferencia Promedio**

$$\bar{D} = \frac{\sum_{i=1}^n D_i}{n} \dots \dots \dots (3)$$

$$\bar{D} = \frac{-18.00}{6}$$

$$\bar{D} = -3.00$$

**b) Desviación Estándar**

$$S_{D^2} = \frac{n \sum_{i=1}^n D_i^2 - [\sum_{i=1}^n D_i]^2}{n(n-1)} \dots (4)$$

$$S_{D^2} = \frac{6(54.16) - (-18.00)^2}{6(6-1)}$$

$$S_{D^2} = 0.032$$

**c) Calculo de T:**

$$t = \frac{\bar{D}\sqrt{n}}{\sqrt{S_D}}$$

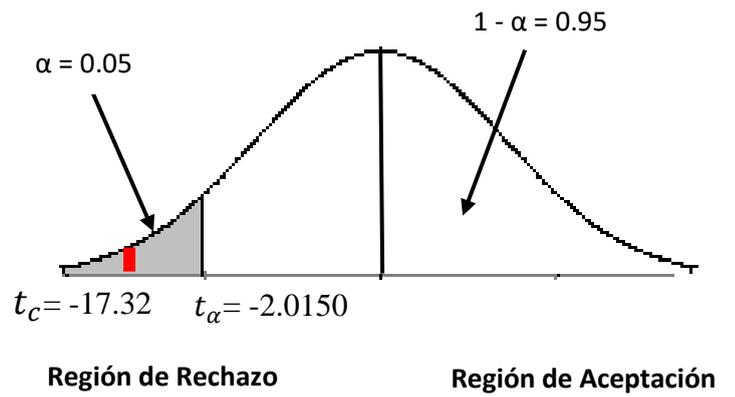
$$t = \frac{(-3.00)\sqrt{6}}{\sqrt{0.18}}$$

$$t = -17.32$$

**d) Conclusión:**

Puesto que:  $t_c = -17.32 (t_{calculado}) < t_{\alpha} = -2.0150 (t_{tabular})$ , y estando este valor dentro de la región de rechazo, se concluye que  $CI_A - CI_P < 0$ , se rechaza  $H_0$  y  $H_a$  es aceptada, por lo tanto se prueba la validez de la hipótesis con un nivel de error de 5% ( $\alpha=0.05$ ), siendo la Solución de BI es una alternativa de solución para el problema de investigación.

En la Figura podemos ver la región de aceptación y rechazo para la prueba de la hipótesis Calidad de Información que se le brinda al usuario final.



**Figura N° 62:** Zona de Aceptación y Rechazo – Calidad de la Información.

## 4.2.1.2 Indicadores cuantitativos.

### 4.2.1.2.1 Prueba de la Hipótesis para el indicador cuantitativo.

**Tiempo de generación de reportes que soportan la toma de decisiones.**

#### A. Definición de Variables

$TGR_A$ : Tiempo de generación de reportes que soportan la toma de decisiones con el Sistema de Trabajo Actual (minutos).

$TGR_P$ : Tiempo de generación de reportes que soportan la toma de decisiones con la Solución de BI propuesta (minutos).

#### B. Hipótesis Estadísticas

- **Hipótesis  $H_0$** : El tiempo de generación de reportes que soportan la toma de decisiones con el Sistema de Trabajo Actual es menor que el Tiempo de generación de reportes con la Solución de BI propuesta.

$$H_0: TGR_A - TGR_P \leq 0$$

- **Hipótesis  $H_a$** : El tiempo de generación de reportes que soportan la toma de decisiones con el Sistema de Trabajo Actual es mayor que el Tiempo de generación de reportes con la Solución de BI propuesta.

$$H_a: TGR_A - TGR_P > 0$$

### C. Nivel de Significancia

Usando un nivel de significancia ( $\alpha = 0.05$ ) del 5%. Por lo tanto el nivel de confianza ( $1 - \alpha = 0.95$ ) será del 95%.

### D. Estadígrafo de contraste

Puesto que  $n=64$  es grande, usaremos la distribución normal (Z):

#### a) Promedio:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \dots\dots\dots 6$$

#### b) Varianza:

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} \dots\dots\dots 7$$

#### c) Calculo de Z

$$Z_C = \frac{(\bar{X}_A - \bar{X}_P)}{\sqrt{\left[ \frac{\sigma_A^2}{n_A} + \frac{\sigma_P^2}{n_P} \right]}}$$

### E. Resultados de la Estadística:

Para calcular El tiempo Promedio en trazar el corte se ha estimado de una muestra de 64 tiempos para elaborar reportes.

**Tabla N° 23:** Toma de Tiempo de Generación de Reportes.

N°	PRE TEST (minutos)	POST TEST (minutos)	PRE TEST (minutos)	POST TEST (minutos)	PRE TEST (minutos)	POST TEST (minutos)
	$TGR_A$	$TGR_P$	$TGR_A - \overline{TGR_A}$	$TGR_P - \overline{TGR_P}$	$(TGR_A - \overline{TGR_A})^2$	$(TGR_P - \overline{TGR_P})^2$
1	245	5	-29	0	841	0
2	160	4	-114	-1	12996	1
3	278	5	4	0	16	0
4	156	4	-118	-1	13924	1
5	195	4	-79	-1	6241	1
6	360	6	86	1	7396	1
7	245	6	-29	1	841	1
8	165	4	-109	-1	11881	1
9	256	5	-18	0	324	0
10	278	5	4	0	16	0
11	390	6	116	1	13456	1
12	178	4	-96	-1	9216	1
13	450	7	176	2	30976	4
14	295	5	21	0	441	0
15	284	5	10	0	100	0
16	370	6	96	1	9216	1
17	465	7	191	2	36481	4
18	267	5	-7	0	49	0
19	271	5	-3	0	9	0
20	189	4	-85	-1	7225	1
21	298	5	24	0	576	0
22	289	5	15	0	225	0
23	185	4	-89	-1	7921	1
24	398	6	124	1	15376	1
25	289	5	15	0	225	0
26	278	5	4	0	16	0
27	296	5	22	0	484	0
28	178	4	-96	-1	9216	1
29	186	4	-88	-1	7744	1
30	268	5	-6	0	36	0
31	368	6	94	1	8836	1

32	376	6	102	1	10404	1
33	355	6	81	1	6561	1
34	345	6	71	1	5041	1
35	199	4	-75	-1	5625	1
36	182	4	-92	-1	8464	1
37	176	4	-98	-1	9604	1
38	278	5	4	0	16	0
39	277	5	3	0	9	0
40	182	4	-92	-1	8464	1
41	274	5	0	0	0	0
42	396	6	122	1	14884	1
43	375	6	101	1	10201	1
44	168	4	-106	-1	11236	1
45	295	5	21	0	441	0
46	173	4	-101	-1	10201	1
47	187	4	-87	-1	7569	1
48	146	4	-128	-1	16384	1
49	279	5	5	0	25	0
50	390	6	116	1	13456	1
51	372	6	98	1	9604	1
52	369	6	95	1	9025	1
53	156	4	-118	-1	13924	1
54	278	5	4	0	16	0
55	289	5	15	0	225	0
56	396	6	122	1	14884	1
57	173	4	-101	-1	10201	1
58	287	5	13	0	169	0
59	132	4	-142	-1	20164	1
60	156	4	-118	-1	13924	1
61	278	5	4	0	16	0
62	391	6	117	1	13689	1
63	387	6	113	1	12769	1
64	289	5	15	0	225	0
<b>Total</b>	<b>17536.00</b>	<b>320</b>			<b>469720</b>	<b>46</b>

a) **Promedio:**

$$\overline{TGR_A} = \frac{17536}{64}$$

$$\overline{TGR_A} = 274$$

$$\overline{TGR_P} = \frac{320}{64}$$

$$\overline{TGR_P} = 5$$

b) **Varianza:**

$$\sigma^2_A = \frac{469720}{64}$$

$$\sigma^2_A = 7339.375$$

$$\sigma^2_P = \frac{46}{64}$$

$$\sigma^2_P = 0.71875$$

c) **Calculo de Z**

$$Z_C = \frac{(274 - 5)}{\sqrt{\left[\frac{7339.38}{64} + \frac{0.72}{64}\right]}}$$

$$Z_C = \frac{269}{\sqrt{[114.678 + 0.011]}}$$

$$Z_c = \frac{269}{\sqrt{114.689}}$$

$$Z_c = \frac{269}{10.71}$$

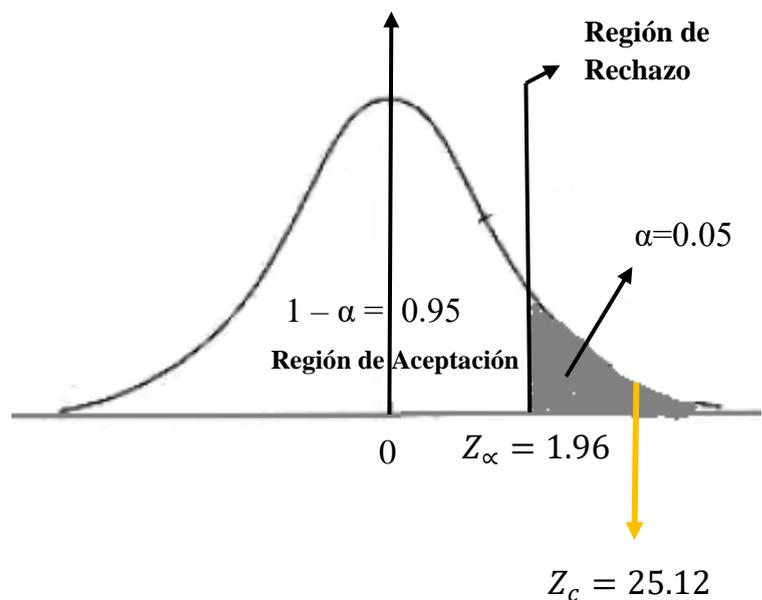
$$Z_c = 25.12$$

**d) Región Crítica:**

Para  $\alpha=0.05$ , en la tabla de Distribución Normal encontramos  $Z_\alpha = 1.96$ . Entonces la región crítica de la prueba es  $Z_c = < 1.96, \infty >$ .

**e) Conclusiones:**

En la Figura podemos ver la región de aceptación y rechazo para la prueba de la hipótesis Tiempo para generar los reportes.



**Figura N° 63:** Zona de Aceptación y Rechazo – Tiempo para generar los reportes.

Puesto que:  $Z_c = 25.12$  calculado es mayor que  $Z_\alpha = 1.96$  y estando este valor dentro de la región de rechazo  $< 1.96, \infty >$ , entonces se rechaza  $H_0$  y por consiguiente se acepta  $H_a$ .

Se concluye entonces que el Tiempo para generar los reportes con la Solución de BI es menor que con el sistema de trabajo actual con un nivel de error de 5% y un nivel de confianza del 95%.

#### 4.2.1.2.2 Prueba de la Hipótesis para el indicador cuantitativo.

##### **Costo por la generación de reportes.**

##### **A. Definición de Variables**

$CGR_A$ : Costo de generación de reportes con el Sistema de Trabajo Actual (soles).

$CGR_P$ : Costo de generación de reportes con la Solución de BI propuesta (Soles).

##### **B. Hipótesis Estadísticas**

- **Hipótesis  $H_0$ :** El Costo de generación de reportes que soportan la toma de decisiones con el Sistema de Trabajo Actual es menor que el Costo de generación de reportes con la Solución de BI propuesta.

$$H_0: CGR_A - CGR_P \leq 0$$

- **Hipótesis  $H_a$ :** El Costo de generación de reportes que soportan la toma de decisiones con el Sistema de Trabajo Actual es mayor que el Costo de generación de reportes con la Solución de BI propuesta.

$$H_a: CGR_A - CGR_P > 0$$

### C. Nivel de Significancia

Usando un nivel de significancia ( $\alpha = 0.05$ ) del 5%. Por lo tanto el nivel de confianza ( $1 - \alpha = 0.95$ ) será del 95%.

### D. Estadígrafo de contraste

Puesto que  $n=64$  es grande, usaremos la distribución normal (Z):

#### a) Promedio:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \dots\dots\dots 6$$

#### b) Varianza:

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} \dots\dots\dots 7$$

c) **Calculo de Z**

$$Z_C = \frac{(\overline{X}_A - \overline{X}_P)}{\sqrt{\left[\frac{\sigma_A^2}{n_A} + \frac{\sigma_P^2}{n_P}\right]}}$$

**E. Resultados de la Estadística:**

Para calcular El Costo Promedio en trazar el corte se ha estimado de una muestra de 64 costos para generar reportes.

**Tabla N° 24:** Toma de Costos de Generación de Reportes.

N°	PRE TEST (soles)	POST TEST (soles)	PRE TEST (soles)	POST TEST (soles)	PRE TEST (soles)	POST TEST (soles)
	$CGR_A$	$CGR_P$	$CGR_A - \overline{CGR}_A$	$CGR_P - \overline{CGR}_P$	$(CGR_A - \overline{CGR}_A)^2$	$(CGR_P - \overline{CGR}_P)^2$
1	107.19	2.19	-12.69	0.00	160.97	0.00
2	70.00	1.75	-49.88	-0.44	2487.52	0.19
3	121.63	2.19	1.75	0.00	3.06	0.00
4	68.25	1.75	-51.63	-0.44	2665.14	0.19
5	85.31	1.75	-34.56	-0.44	1194.57	0.19
6	157.50	2.63	37.63	0.44	1415.64	0.19
7	107.19	2.63	-12.69	0.44	160.97	0.19
8	72.19	1.75	-47.69	-0.44	2274.10	0.19
9	112.00	2.19	-7.88	0.00	62.02	0.00
10	121.63	2.19	1.75	0.00	3.06	0.00
11	170.63	2.63	50.75	0.44	2575.56	0.19
12	77.88	1.75	-42.00	-0.44	1764.00	0.19
13	196.88	3.06	77.00	0.88	5929.00	0.77
14	129.06	2.19	9.19	0.00	84.41	0.00
15	124.25	2.19	4.38	0.00	19.14	0.00
16	161.88	2.63	42.00	0.44	1764.00	0.19
17	203.44	3.06	83.56	0.88	6982.69	0.77
18	116.81	2.19	-3.06	0.00	9.38	0.00
19	118.56	2.19	-1.31	0.00	1.72	0.00

20	82.69	1.75	-37.19	-0.44	1382.91	0.19
21	130.38	2.19	10.50	0.00	110.25	0.00
22	126.44	2.19	6.56	0.00	43.07	0.00
23	80.94	1.75	-38.94	-0.44	1516.13	0.19
24	174.13	2.63	54.25	0.44	2943.06	0.19
25	126.44	2.19	6.56	0.00	43.07	0.00
26	121.63	2.19	1.75	0.00	3.06	0.00
27	129.50	2.19	9.63	0.00	92.64	0.00
28	77.88	1.75	-42.00	-0.44	1764.00	0.19
29	81.38	1.75	-38.50	-0.44	1482.25	0.19
30	117.25	2.19	-2.63	0.00	6.89	0.00
31	161.00	2.63	41.13	0.44	1691.27	0.19
32	164.50	2.63	44.63	0.44	1991.39	0.19
33	155.31	2.63	35.44	0.44	1255.82	0.19
34	150.94	2.63	31.06	0.44	964.88	0.19
35	87.06	1.75	-32.81	-0.44	1076.66	0.19
36	79.63	1.75	-40.25	-0.44	1620.06	0.19
37	77.00	1.75	-42.88	-0.44	1838.27	0.19
38	121.63	2.19	1.75	0.00	3.06	0.00
39	121.19	2.19	1.31	0.00	1.72	0.00
40	79.63	1.75	-40.25	-0.44	1620.06	0.19
41	119.88	2.19	0.00	0.00	0.00	0.00
42	173.25	2.63	53.38	0.44	2848.89	0.19
43	164.06	2.63	44.19	0.44	1952.54	0.19
44	73.50	1.75	-46.38	-0.44	2150.64	0.19
45	129.06	2.19	9.19	0.00	84.41	0.00
46	75.69	1.75	-44.19	-0.44	1952.54	0.19
47	81.81	1.75	-38.06	-0.44	1448.75	0.19
48	63.88	1.75	-56.00	-0.44	3136.00	0.19
49	122.06	2.19	2.19	0.00	4.79	0.00
50	170.63	2.63	50.75	0.44	2575.56	0.19
51	162.75	2.63	42.88	0.44	1838.27	0.19
52	161.44	2.63	41.56	0.44	1727.44	0.19
53	68.25	1.75	-51.63	-0.44	2665.14	0.19
54	121.63	2.19	1.75	0.00	3.06	0.00
55	126.44	2.19	6.56	0.00	43.07	0.00
56	173.25	2.63	53.38	0.44	2848.89	0.19
57	75.69	1.75	-44.19	-0.44	1952.54	0.19
58	125.56	2.19	5.69	0.00	32.35	0.00
59	57.75	1.75	-62.13	-0.44	3859.52	0.19
60	68.25	1.75	-51.63	-0.44	2665.14	0.19
61	121.63	2.19	1.75	0.00	3.06	0.00
62	171.06	2.63	51.19	0.44	2620.16	0.19
63	169.31	2.63	49.44	0.44	2444.07	0.19
64	126.44	2.19	6.56	2.19	43.07	4.79
<b>Total</b>	<b>7672.00</b>	<b>140.00</b>			<b>89907.34</b>	<b>13.59</b>

a) **Promedio:**

$$\overline{CGR}_A = \frac{7672}{64}$$

$$\overline{CGR}_A = \mathbf{119.88}$$

$$\overline{CGR}_P = \frac{140}{64}$$

$$\overline{CGR}_P = \mathbf{2.19}$$

b) **Varianza:**

$$\sigma^2_A = \frac{89907.34}{64}$$

$$\sigma^2_A = \mathbf{1404.80}$$

$$\sigma^2_P = \frac{13.59}{64}$$

$$\sigma^2_P = \mathbf{0.21}$$

c) **Calculo de Z**

$$Z_C = \frac{(119.88 - 2.19)}{\sqrt{\left[\frac{1404.80}{64} + \frac{0.21}{64}\right]}}$$

$$Z_C = \frac{117.69}{\sqrt{[21.95 + 0.003]}}$$

$$Z_C = \frac{117.69}{\sqrt{21.953}}$$

$$Z_c = \frac{117.69}{4.6854}$$

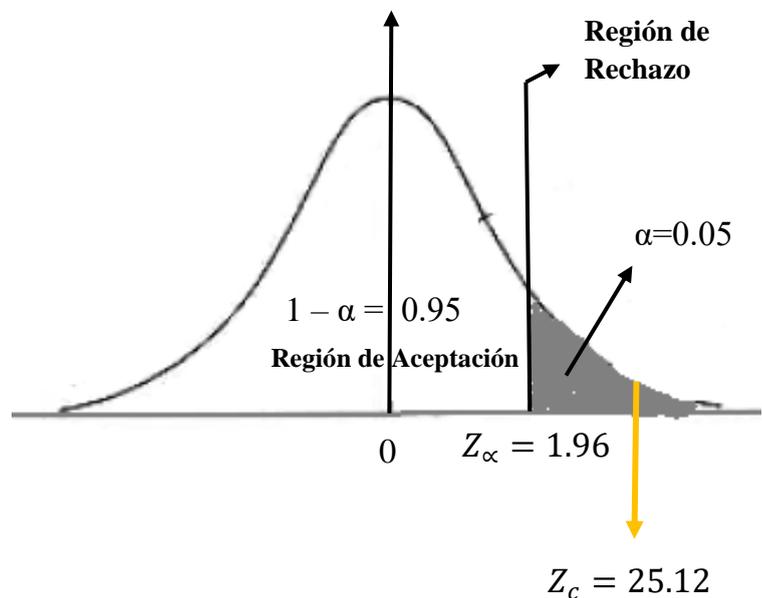
$$Z_c = 25.12$$

**d) Región Crítica:**

Para  $\alpha=0.05$ , en la tabla de Distribución Normal encontramos  $Z_\alpha = 1.96$ . Entonces la región crítica de la prueba es  $Z_c = < 1.96, \infty >$ .

**e) Conclusiones:**

En la Figura podemos ver la región de aceptación y rechazo para la prueba de la hipótesis Costo por la generación de reportes.



**Figura N° 64:** Zona de Aceptación y Rechazo – Costo de generación de reportes.

Puesto que:  $Z_c = 25.12$  calculado es mayor que  $Z_\alpha = 1.96$  y estando este valor dentro de la región de rechazo  $< 1.96, \infty >$ , entonces se rechaza  $H_0$  y por consiguiente se acepta  $H_a$ .

Se concluye entonces que el Costo de generación de reportes con la Solución de BI es menor que con el sistema de trabajo actual con un nivel de error de 5% y un nivel de confianza del 95%.

#### **4.2.1.2.3** *Prueba de la Hipótesis para el indicador cuantitativo.*

### **Tiempo para analizar información y formular alternativas.**

#### **A. Definición de Variables**

$TAIFA_A$ : Tiempo para analizar información y formular alternativas con el Sistema de Trabajo Actual (minutos).

$TAIFA_P$ : Tiempo para analizar información y formular alternativas con la Solución de BI propuesta (minutos).

## B. Hipótesis Estadísticas

- **Hipótesis  $H_0$ :** El Tiempo para analizar información y formular alternativas que soportan la toma de decisiones con el Sistema de Trabajo Actual es menor que el Tiempo para analizar información y formular alternativas con la Solución de BI propuesta.

$$H_0: TAIFA_A - TAIFA_P \leq 0$$

- **Hipótesis  $H_a$ :** El Tiempo para analizar información y formular alternativas que soportan la toma de decisiones con el Sistema de Trabajo Actual es mayor que el Tiempo para analizar información y formular alternativas con la Solución de BI propuesta.

$$H_a: TAIFA_A - TAIFA_P > 0$$

## C. Nivel de Significancia

Usando un nivel de significancia ( $\alpha = 0.05$ ) del 5%. Por lo tanto el nivel de confianza ( $1 - \alpha = 0.95$ ) será del 95%.

## D. Estadígrafo de contraste

Puesto que  $n=36$  es grande, usaremos la distribución normal ( $Z$ ):

a) **Promedio:**

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \dots\dots\dots 6$$

b) **Varianza:**

$$\sigma^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} \dots\dots\dots 7$$

c) **Calculo de Z**

$$Z_C = \frac{(\bar{X}_A - \bar{X}_P)}{\sqrt{\left[ \frac{\sigma_A^2}{n_A} + \frac{\sigma_P^2}{n_P} \right]}}$$

**E. Resultados de la Estadística:**

Para calcular el Tiempo Promedio en trazar el corte se ha estimado de una muestra de 36 tiempos para analizar información y formular alternativas.

**Tabla N° 25:** Toma de Tiempos para analizar información y formular alternativas.

N°	PRE TEST (minutos)	POST TEST (minutos)	PRE TEST (minutos)	POST TEST (minutos)	PRE TEST (minutos)	POST TEST (minutos)
	$TAIFA_A$	$TAIFA_P$	$TAIFA_A - \overline{TAIFA_A}$	$TAIFA_P - \overline{TAIFA_P}$	$(TAIFA_A - \overline{TAIFA_A})^2$	$(TAIFA_P - \overline{TAIFA_P})^2$
1	31	6	-29.19	-3.61	852.32	13.04
2	56	10	-4.19	0.39	17.59	0.15
3	43	9	-17.19	-0.61	295.65	0.37
4	56	10	-4.19	0.39	17.59	0.15
5	45	9	-15.19	-0.61	230.87	0.37
6	23	5	-37.19	-4.61	1383.43	21.26
7	56	10	-4.19	0.39	17.59	0.15
8	87	12	26.81	2.39	718.54	5.71
9	42	9	-18.19	-0.61	331.04	0.37
10	47	9	-13.19	-0.61	174.09	0.37
11	54	10	-6.19	0.39	38.37	0.15
12	68	11	7.81	1.39	60.93	1.93
13	81	12	20.81	2.39	432.87	5.71
14	80	12	19.81	2.39	392.26	5.71
15	78	12	17.81	2.39	317.04	5.71
16	69	11	8.81	1.39	77.54	1.93
17	34	6	-26.19	-3.61	686.15	13.04
18	44	7	-16.19	-2.61	262.26	6.82
19	67	8	6.81	-1.61	46.32	2.60
20	84	11	23.81	1.39	566.70	1.93
21	39	7	-21.19	-2.61	449.20	6.82
22	41	7	-19.19	-2.61	368.43	6.82
23	56	10	-4.19	0.39	17.59	0.15
24	67	11	6.81	1.39	46.32	1.93
25	73	12	12.81	2.39	163.98	5.71
26	47	7	-13.19	-2.61	174.09	6.82
27	82	12	21.81	2.39	475.48	5.71
28	45	7	-15.19	-2.61	230.87	6.82
29	92	14	31.81	4.39	1011.59	19.26
30	87	13	26.81	3.39	718.54	11.48
31	90	14	29.81	4.39	888.37	19.26
32	89	13	28.81	3.39	829.76	11.48
33	68	8	7.81	-1.61	60.93	2.60
34	59	8	-1.19	-1.61	1.43	2.60
35	48	7	-12.19	-2.61	148.70	6.82
36	39	7	-21.19	-2.61	449.20	6.82
<b>Total</b>	<b>2167</b>	<b>346</b>			<b>12953.64</b>	<b>208.56</b>

a) **Promedio:**

$$\overline{TAIFA_A} = \frac{2167}{36}$$

$$\overline{TAIFA_A} = \mathbf{60.19}$$

$$\overline{TAIFA_P} = \frac{346}{36}$$

$$\overline{TAIFA_P} = \mathbf{9.61}$$

b) **Varianza:**

$$\sigma^2_A = \frac{12953.64}{36}$$

$$\sigma^2_A = \mathbf{359.82}$$

$$\sigma^2_P = \frac{208.56}{36}$$

$$\sigma^2_P = \mathbf{5.79}$$

c) **Calculo de Z**

$$Z_C = \frac{(60.19 - 9.61)}{\sqrt{\left[\frac{359.82}{36} + \frac{5.79}{36}\right]}}$$

$$Z_C = \frac{50.58}{\sqrt{[9.995 + 0.161]}}$$

$$Z_C = \frac{50.58}{\sqrt{10.16}}$$

$$Z_C = \frac{50.58}{3.19}$$

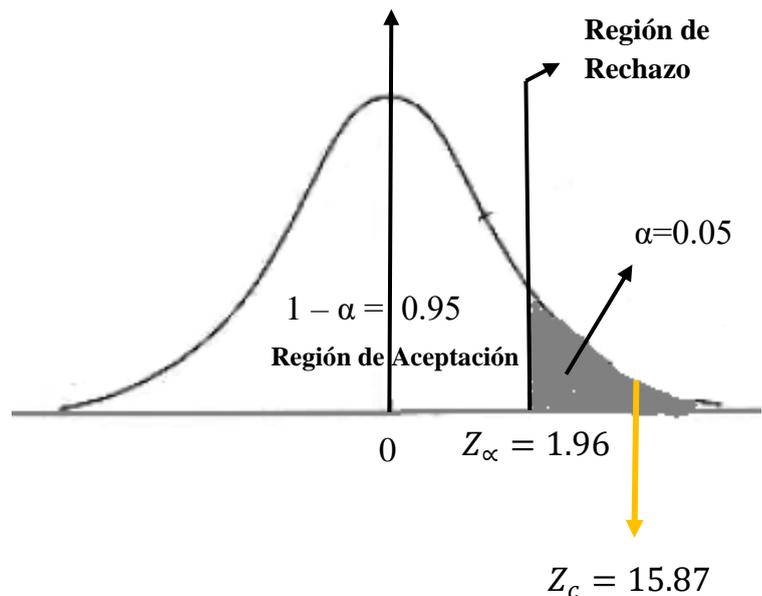
$$Z_C = \mathbf{15.87}$$

**d) Región Crítica:**

Para  $\alpha=0.05$ , en la tabla de Distribución Normal encontramos  $Z_{\alpha} = 1.96$ . Entonces la región crítica de la prueba es  $Z_c = < 1.96, \infty >$ .

**e) Conclusiones:**

En la Figura podemos ver la región de aceptación y rechazo para la prueba de la hipótesis Tiempo para analizar información y formular alternativas.



**Figura N° 65:** Zona de Aceptación y Rechazo – Tiempo para analizar información y formular alternativas.

Puesto que:  $Z_c = 15.87$  calculado es mayor que  $Z_{\alpha} = 1.96$  y estando este valor dentro de la región de rechazo  $< 1.96, \infty >$ , entonces se rechaza  $H_0$  y por consiguiente se acepta  $H_a$ .

Se concluye entonces que el Tiempo para analizar información y formular alternativas con la Solución de BI es menor que con el sistema de trabajo actual con un nivel de error de 5% y un nivel de confianza del 95%.

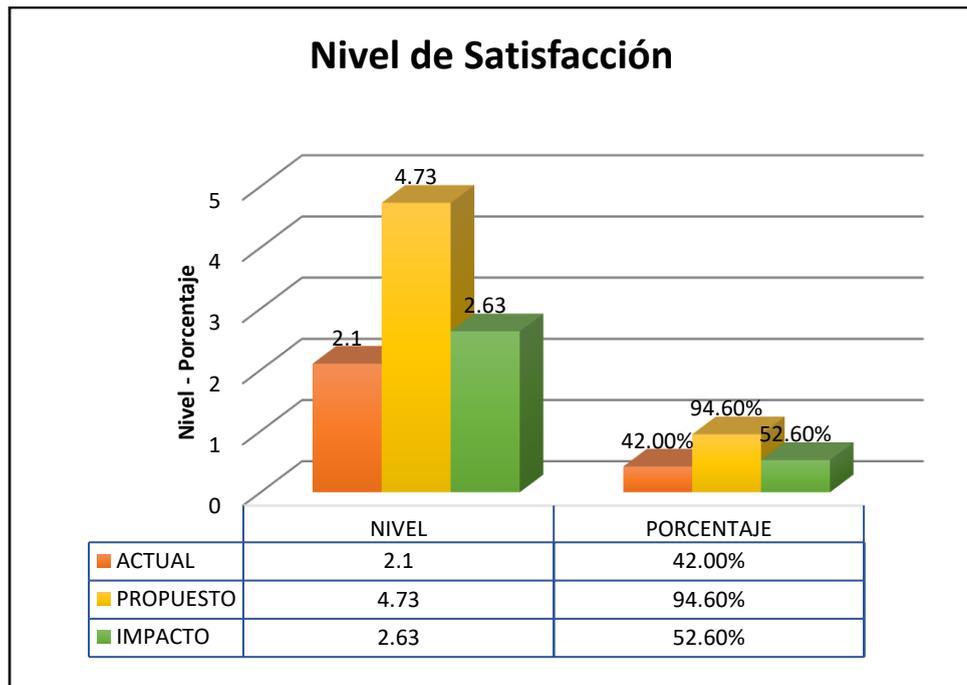
### 4.3 Resultados

#### 4.3.1 Indicador cualitativo: nivel de satisfacción de los usuarios.

Para este indicador se usó la encuesta evaluada con la escala de Likert (1-5), donde 5 es el 100%. (Tabla N°26) Para el nivel de satisfacción de los usuarios con el sistema de trabajo actual ( $NSU_A$ ) y el nivel de satisfacción de los usuarios con la solución de BI propuesta ( $NSU_P$ ) obteniendo los siguientes resultados:

**Tabla N° 26:** Presentación de Resultados Nivel de Satisfacción del Usuario.

Nivel de satisfacción del usuario con el sistema de trabajo actual.		Nivel de satisfacción del usuario con la solución de BI propuesta.		Nivel de Impacto del indicador de satisfacción del usuario (Incremento)	
$NSUT_A$ (1-5)	$NSUT_A$ (%)	$NSUS_P$ (1-5)	$NSUS_P$ (%)	$\Delta NISU$ (1-5)	$\Delta NISU$ (%)
2.10	42.00%	4.73	94.60%	2.63	52.60%



**Figura N° 66:** Gráfico de Resultados Indicador Nivel de Satisfacción del Usuario.

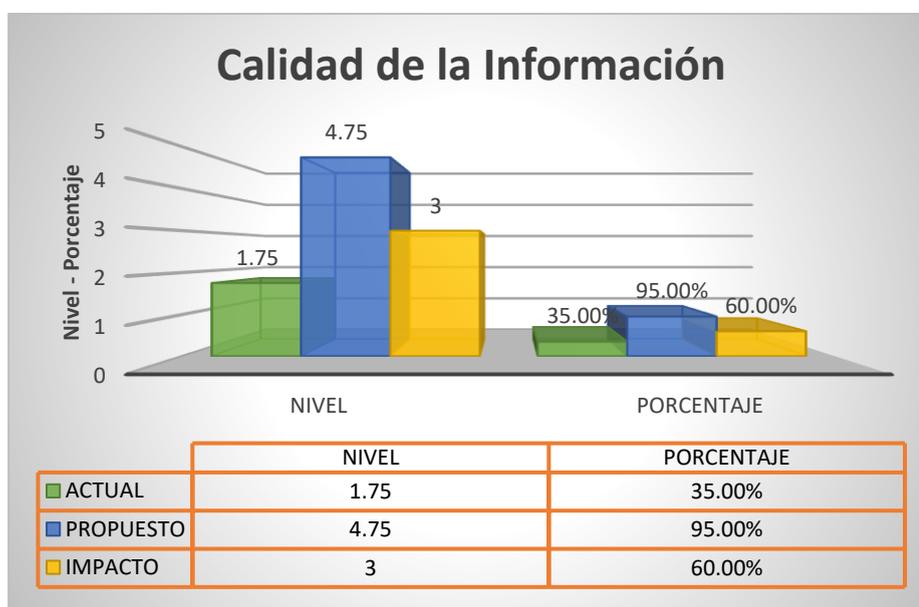
En el cuadro anterior (Tabla N° 26) se observa que el nivel de satisfacción del usuario se ha incrementado en un 52.60% y el objetivo cumple con lo propuesto.

#### 4.3.2 Indicador cualitativo: calidad de la información.

Para este indicador se usó la encuesta evaluada con la escala de Likert (1-5), donde 5 es el 100%. (Tabla N° 27) Para el Calidad de la Información con el sistema de trabajo actual ( $CI_A$ ) y la Calidad de la Información con la solución de BI propuesta ( $CI_P$ ) obteniendo los siguientes resultados:

**Tabla N° 27:** Presentación de Resultados de Calidad de la Información.

Calidad de la Información con el sistema de trabajo actual.		Calidad de la Información con la solución de BI propuesta.		Nivel de Impacto del indicador de Calidad de la Información (Incremento)	
$CI_A$ (1-5)	$CI_A$ (%)	$CI_P$ (1-5)	$CI_P$ (%)	$\Delta CI$ (1-5)	$\Delta CI$ (%)
1.75	35.00%	4.75	95.00%	3.00	60.00%



**Figura N° 67:** Gráfico de Resultados del Indicador Calidad de la Información.

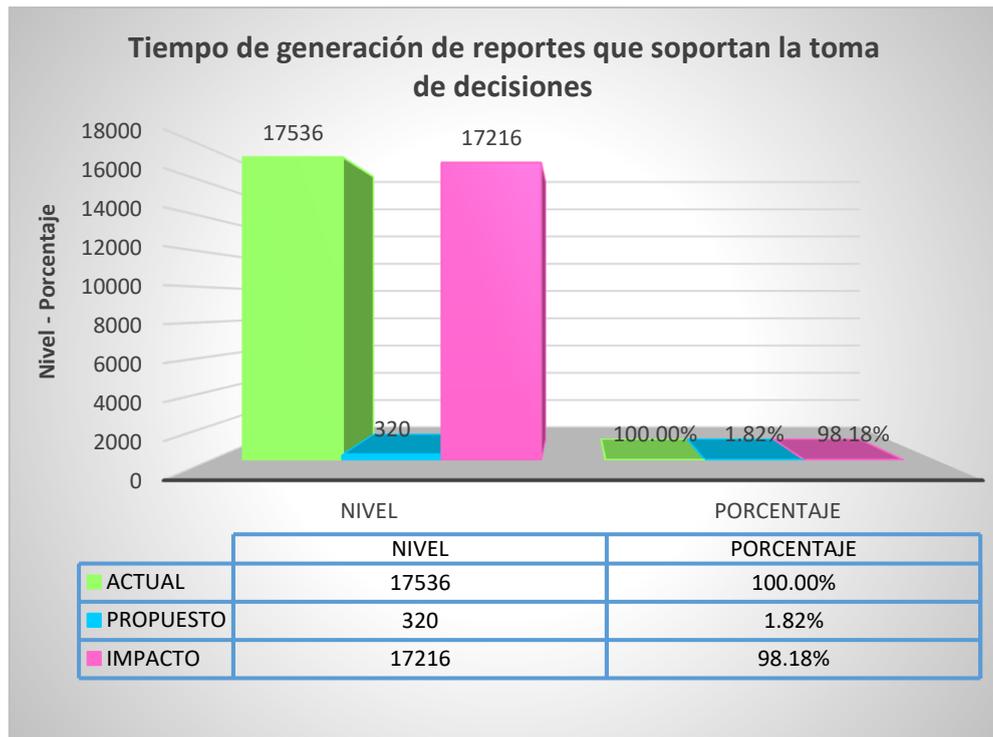
En el cuadro anterior (Tabla N° 27) se observa que la calidad de la información se ha incrementado en un 60.00% y el objetivo cumple con lo propuesto.

### 4.3.3 Indicador cuantitativo: tiempo de generación de reportes que soportan la toma de decisiones.

Para este indicador se usó los datos tomados en campo del cuantitativo tiempo de generación de reportes que soportan la toma de decisiones con el sistema de trabajo actual ( $TGR_A$ ) para compararlos con el cuantitativo tiempo de generación de reportes que soportan la toma de decisiones con la solución de BI propuesta ( $TGR_P$ ).

**Tabla N° 28:** Presentación de Resultados de Tiempo de generación de reportes que soportan la toma de decisiones.

Tiempo de generación de reportes que soportan la toma de decisiones con el sistema de trabajo actual.		Tiempo de generación de reportes que soportan la toma de decisiones con la solución de BI propuesta.		Nivel de Impacto del indicador de Tiempo de generación de reportes que soportan la toma de decisiones (Reducción)	
$TGR_A$ (minutos)	$TGR_A$ (%)	$TGR_P$ (minutos)	$TGR_P$ (%)	$\Delta TGR$ (minutos)	$\Delta TGR$ (%)
17536	100.00%	320	1.82%	17216	98.18%



**Figura N° 68:** Gráfico de Resultados del Indicador Tiempo de generación de reportes que soportan la toma de decisiones.

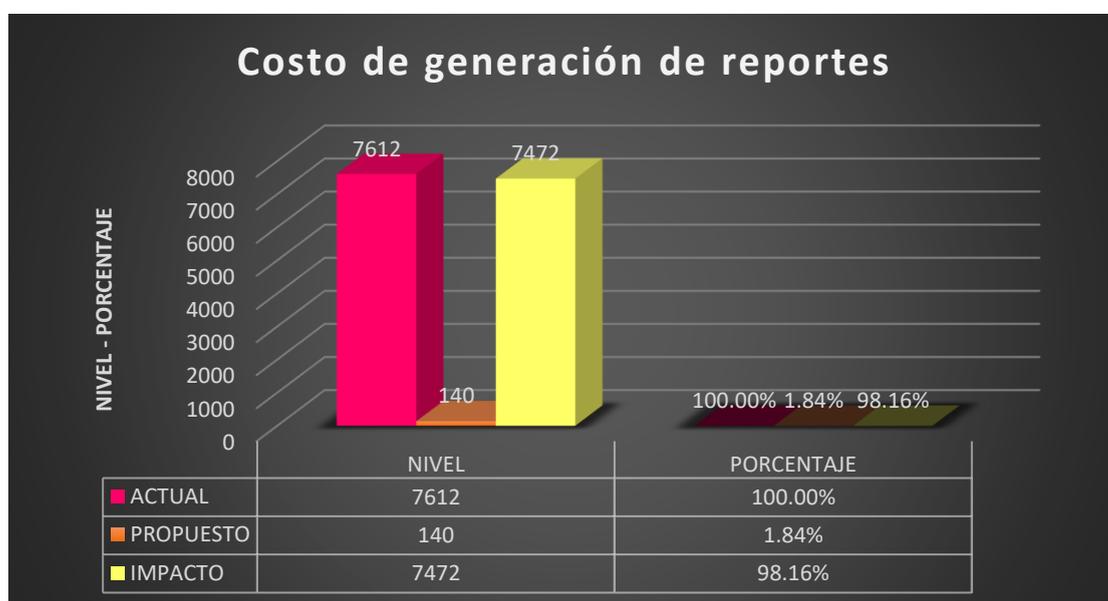
Del cuadro anterior (Tabla N° 28) se observa que el Tiempo de generación de reportes que soportan la toma de decisiones con la solución de BI propuesto se ha reducido en un 98.18%.

#### **4.3.4 Indicador cuantitativo: costo de generación de reportes.**

Para este indicador se usó los datos tomados en campo del cuantitativo costo de generación de reportes con el sistema de trabajo actual ( $CGR_A$ ) para compararlos con el cuantitativo costo de generación de reportes con la solución de BI propuesta ( $CGR_P$ ).

**Tabla N° 29:** Presentación de Resultados de Costo de generación de reportes.

Costo de generación de reportes con el sistema de trabajo actual.		Costo de generación de reportes con la solución de BI propuesta.		Nivel de Impacto del indicador de Costo de generación de reportes (Reducción)	
$CGR_A$ (soles)	$CGR_A$ (%)	$CGR_P$ (soles)	$CGR_P$ (%)	$\Delta CGR$ (soles)	$\Delta CGR$ (%)
7612	100.00%	140	1.84%	7472	98.16%



**Figura N° 69:** Gráfico de Resultados del Indicador Costo de generación de reportes.

Del cuadro anterior (Tabla N° 29) se observa que el Costo de generación de reportes con la solución de BI propuesto se ha reducido en un 98.16%.

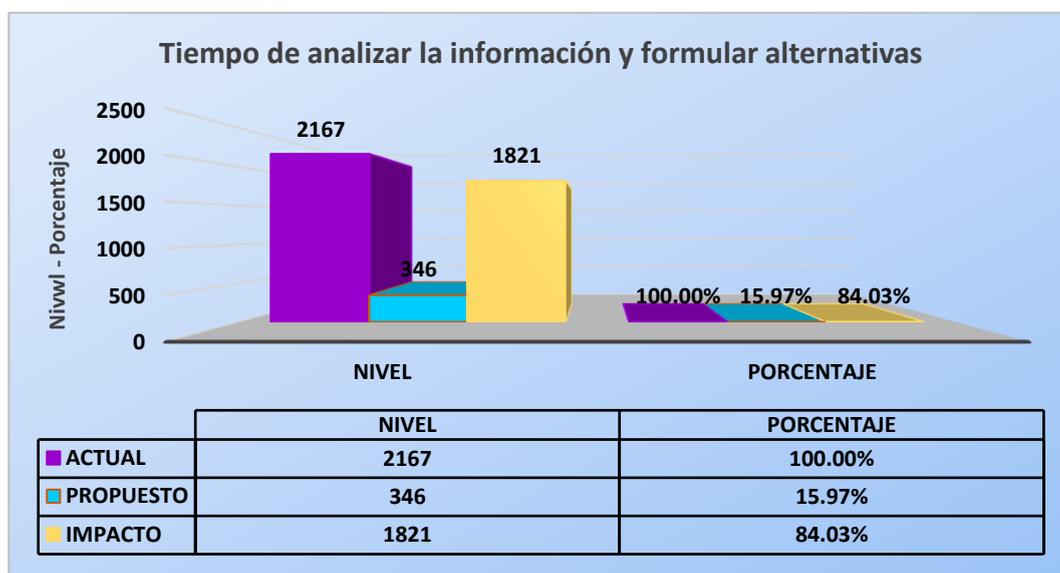
#### 4.3.5 Indicador cuantitativo: tiempo para analizar información y formular alternativas.

Para este indicador se usó los datos tomados en campo del cuantitativo Tiempo de analizar la información y formular alternativas con el sistema de trabajo actual ( $TAIFA_A$ ) para compararlos con el cuantitativo Tiempo

de analizar la información y formular alternativas con la solución de BI propuesta ( $TAIFA_P$ ).

**Tabla N° 30:** Presentación de Resultados del Tiempo de analizar la información y formular alternativas.

Tiempo para analizar información y formular alternativas con el sistema de trabajo actual.		Tiempo para analizar información y formular alternativas con la solución de BI propuesta.		Nivel de Impacto del indicador de Tiempo para analizar información y formular alternativas (Reducción)	
$TAIFA_A$ (minutos)	$TAIFA_A$ (%)	$TAIFA_P$ (minutos)	$TAIFA_P$ (%)	$\Delta TAIFA$ (minutos)	$\Delta TAIFA$ (%)
2167	100.00%	346	15.97%	1821	84.03%



**Figura N° 70:** Gráfico de Resultados del Indicador Tiempo para analizar información y formular alternativas.

Del cuadro anterior (Tabla N° 30) se observa que el Tiempo para analizar información y formular alternativas con la solución de BI propuesto se ha reducido en un 84.03%.

## CAPITULO V

### CONCLUSION Y RECOMENDACIONES

#### 5.1 Conclusiones

Se levantó de los requerimientos del Hospital la se Caleta (usuarios) utilizando la metodología Hefesto, que permitió identificar fácilmente objetivos, las necesidades del negocio a través de preguntas, dando como resultado 5 indicadores y 6 perspectivas que determinan el nivel de granularidad de nuestro Data Warehouse. Por lo que se concluye que esta metodología es eficiente para investigaciones relacionadas a la construcción de Data WareHouse, y flexible ya que es independientemente de las herramientas que se usan para su implementación.

Al analizar la fuente de datos del sistema operacional para confeccionar el Data Warehouse, teniendo como base el modelo conceptual que ya ha sido creado. Se concluyó que el esquema en estrella es el más simple de interpretar y optimiza tiempo en las acciones propias para diseñar las dimensiones, hechos y las uniones pertinentes entre estas tablas.

Para el diseño de la solución de BI se concluye que Pentaho, es una herramienta muy intuitiva, manejable, multiplataforma, sencilla al momento de integrar los datos y lo más importante que al ser Open Source, los componentes requeridos para la generación de información adicional son accesibles para los usuarios.

La implementación de la solución de BI se realizó con la herramienta Pentaho Data Integration dando como resultado que el proceso ETL conste de 3 flujos de trabajo y 10 flujos de carga o transformación. Schema Workbench permitio el diseñado los 4 cubos OLAP. Concluyendo finalmente con los reportes de interés de los usuarios mostrados mediante un entorno web (Pentaho BI Server).

Realizadas las pruebas mediante solución de inteligencia de negocios:

- ✓ Se concluye que el nivel de satisfacción del usuario mejoró por cuanto los estadísticos del pre test arrojaron en base a la encuesta con rango de ponderación: [1-5] con el sistema anterior de trabajo el valor de 2.10 (42.00%) y con los estadísticos de post test con la solución de BI 4.73 (94.60 %), lo que significa que el nivel de satisfacción de los usuarios se incrementó en un 2.63 (52.60%).
- ✓ Se concluye que la calidad de la información mejoró por cuanto los estadísticos del pre test arrojaron en base a la encuesta con rango de ponderación: [1-5] con el sistema anterior de trabajo el valor de 1.75 (35.00%) y con los estadísticos de post test con la solución de BI 4.75 (95.00%), lo que significa que la calidad de la información que se le entrega a los usuarios ha incrementado a 3(60.00%).
- ✓ Se concluye que el tiempo de generación de reportes que soportan la toma de decisiones se redujo por cuanto los estadísticos del pre test arrojaron con el sistema anterior de trabajo el valor de 17536 minutos (100.00%) y con los estadísticos de post test con la solución de BI 320 minutos (1.82%), lo que significa que el tiempo de generación de reportes que soportan la toma de decisiones ha reducido en 17216 minutos (98.18%).
- ✓ Se concluye que el costo de generación de reportes se redujo por cuanto los estadísticos del pre test arrojaron con el sistema anterior de trabajo el valor de 7612 soles (100.00%) y con los estadísticos de post test con la solución de BI 140 soles (1.84%), lo que significa que el costo de generación de reportes ha reducido en 7412 soles (98.16 %).
- ✓ Se concluye que tiempo para analizar la información y formular alternativas se redujo por cuanto los estadísticos del pre test arrojaron con el sistema anterior de trabajo el valor de 2167 minutos (100.00%) y con los estadísticos de post test con la solución de BI 346 minutos (15.97%), lo que significa que tiempo para analizar la información y formular alternativas ha reducido en 1821 minutos (84.03%).

Finalmente queda demostrado que una solución de inteligencia de negocios mejora la toma de decisiones en la Dirección Ejecutiva del Hospital La Caleta, en base a la mejora del nivel de satisfacción de los usuarios en 52.60%, la calidad de la información en 60.00%, el tiempo de generación de reportes que soportan la toma de decisiones en 98.18%, costo de generación de reportes en 98.16 % y el tiempo de análisis de información y formular alternativas en 84.03%. Por lo tanto, nuestra hipótesis de problema quedo demostrada.

## **5.2 Recomendaciones**

Para mejorar la continuidad operacional del sistema propuesto, se presentan las siguientes recomendaciones:

Si el ETL incluye una gran cantidad de datos, planifique la hora de inicio de forma que los demás usuarios resulten mínimamente afectados por la ralentización de la actividad del servidor debido a la carga de proceso.

Se recomienda establecer como política de la institución hacer uso de la solución de inteligencia de negocios ya que se puede conocer muy bien el proceso a desarrollar, pero si las herramientas no son utilizadas entonces llevará al fracaso del producto final.

Es importante que los usuarios finales del sistema, reciban capacitaciones para obtener el máximo rendimiento y un mejor aprovechamiento en cuanto a las diferentes formas de obtener, analizar y exportar la información requerida, para poder tomar las decisiones adecuadas teniendo una buena data previa.

Se deben realizar encuestas periódicas a los usuarios para determinar los niveles de satisfacción con respecto a la aplicación y evaluar posibles mejoras en el sistema.

Realizar un plan de mantenimiento y actualización, para la herramienta de toma de decisiones y la infraestructura tecnológica, procedimiento necesario para el correcto funcionamiento del almacén de datos y la herramienta, debido a su demandante crecimiento en el almacenamiento de datos.

Para continuar con la investigación, es necesario obtener acceso a las fuentes de información tales como el Sistema de Contabilidad, Logística y Recursos Humanos. De esta forma se podrá explotar aún más la información respecto a todos los servicios que se prestan en la institución y así contar con una solución de inteligencia de negocios más completa que permita la generación de un mayor y mejor número de reportes analíticos que ayuden a mejorar la toma de decisiones de la dirección ejecutiva del Hospital La Caleta.

## REFERENCIAS BIBLIOGRÁFICAS

- Conesa Caralt, J., & Curto Díaz, J., (2010, Mayo). Introducción al Business Intelligence. Barcelona: Universitat Oberta De Catalunya.
- Hatch, D. (2009, Marzo). Operational Business Intelligence What you Need to Know About Improving Your Customer-Facing Performance.
- Howson, C. (2009). Business Intelligence: Estrategias para una implementación exitosa. México: McGraw-Hill. Base de datos en línea: E-Libro.
- Kimball, R. y Ross, M. (2010). The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence. USA Wiley Publishing, Inc.
- Kommineni Sivaganesh, S., & Suresh Chandra, S. (2012). Optimization of ETL Work Flow in DataWarehouse. India: Engg Journals Publications. Revista: International Journal on Computer Science and Engineering. Catálogo de Revistas Especializadas en Línea: DOAJ.
- Kroenke, D., (2003). Procesamiento de Bases de Datos: Fundamentos, diseño e implementación. México: Pearson Educación
- Larissa T. Moss, & Shaku Atre. (2003). Business Intelligence Roadmap: The Complete Project Lifecycle for Decision Support Applications (3 ed.). Boston: Pearson Education.
- Maqueda Lafuente, J., (1992, Enero). Dirección Estratégica y Planificación Financiera de la PYME. Madrid: Ediciones Díaz de Santos SA.
- Namakforoosh, M., (2005). Metodología de la Investigación. México: LIMUSA Noriega Editores.
- Oltra Badenes, R. (2012) Sistemas Integrados de Gestión Empresarial : Evolución histórica y tendencias de futuro . Editorial UPV.
- Sabherwal, R. & Becerra-Fernández, I. (2011). Business Intelligence Practices, Technologies, and Management. EUA: John Wiley & Sons, Inc.

Salcedo Parra, O., Milena Galeano, R., & Rodriguez B, L. (2010). En Modelamiento Dimensional de Datos. España: Editorial NETBIBLO. Revista: Red de Revistas Científicas de América Latina, el Caribe, España y Portugal. Base de datos en línea: E-Libro.

Technics Publications,LLC,(2010). Guía de Fundamentos para la Gestión de Datos. USA: DAMA International.

## **REFERENCIAS DE PÁGINAS EN LA WEB**

Bernabeu R. Dario (Mayo de 2009). Proceso de BI Recuperado el 19 de Abril de 2017,<http://www.dataprix.com/data-warehousing-y-metodologia-hefesto/1-business-intelligence/13-proceso-bi>

Gravitar. (2016). Introducción a PENTAHO. Retrieved 02 de diciembre, 2016, from <http://gravitar.biz/bi/introduccion-pentaho-parte-1/>

Sinnexus. (2007, 2016). Business Intelligence - Informática Estratégica. Retrieved 20 de octubre 2016, from [http://www.sinnexus.com/business\\_intelligence/datamart.aspx](http://www.sinnexus.com/business_intelligence/datamart.aspx)

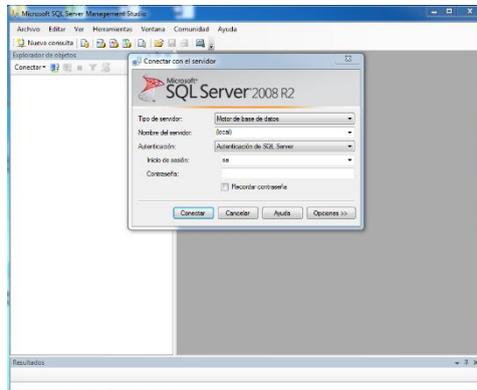
# ANEXO A

## CONEXIÓN SQL SERVER 2008

### 6.1 Configuración de la base de datos en SQL Server 2008

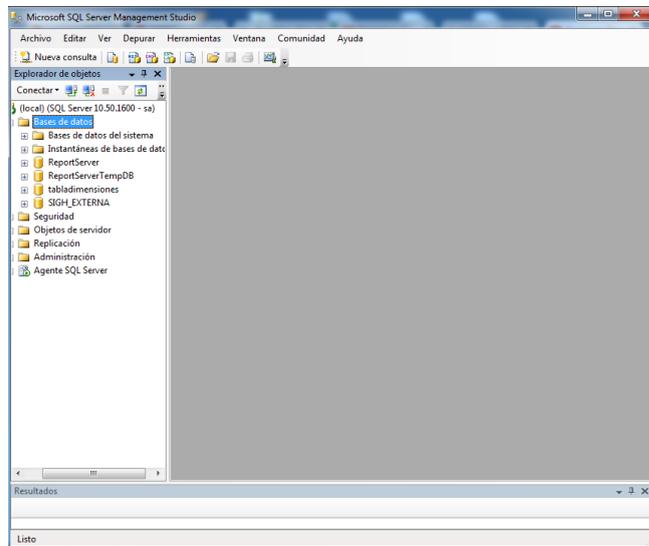
Luego de tener instalado el SQL SERVER, nos dirigimos a **Inicio**, seleccionamos **Todos los programas**, Microsoft SQL Server 2008 R2 y, hacemos clic a continuación en Microsoft AQL Server Management Studio.

En el cuadro de dialogo “Conectar con el Servidor” ingresamos el Usuario de Inicio de Sesión y la Contraseña.



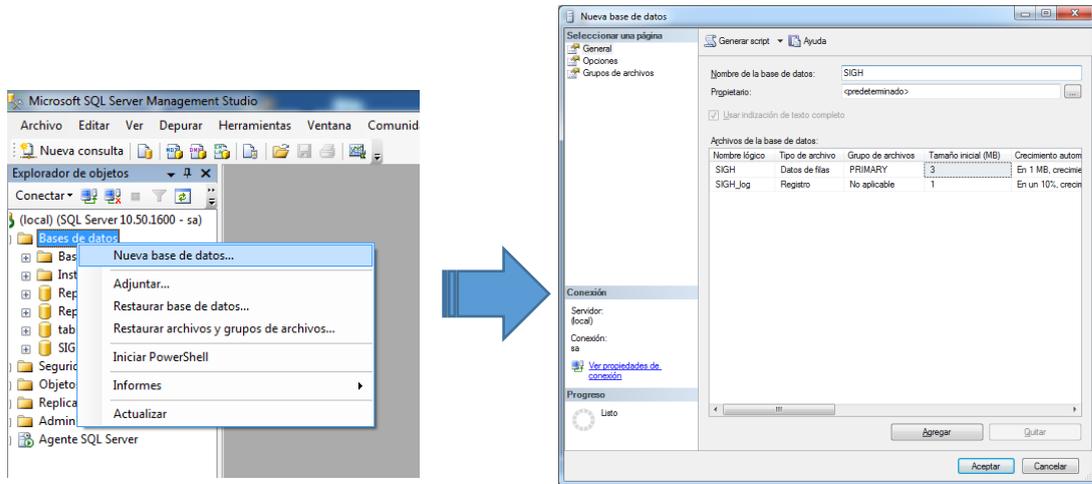
**Figura N° 71:** Pantalla de Inicio y Conexión al Servidor de SQL SERVER 2008.

El sistema nos abrirá la siguiente ventana mostrándonos el Explorador de Objetos.



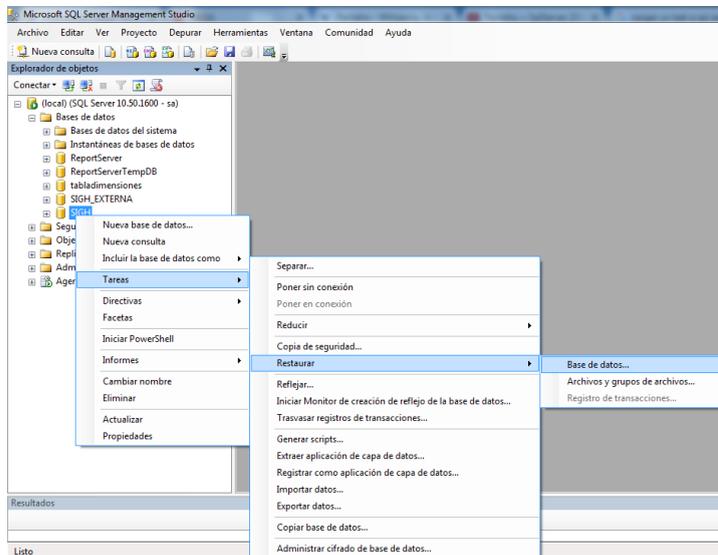
**Figura N° 72:** Explorador de Objetos SQL SERVER 2008.

Hacemos CLIC DERECHO en Bases de datos, y seleccionamos “Nueva base de datos...” e Ingresamos el *Nombre de la base de datos* y “Aceptar”



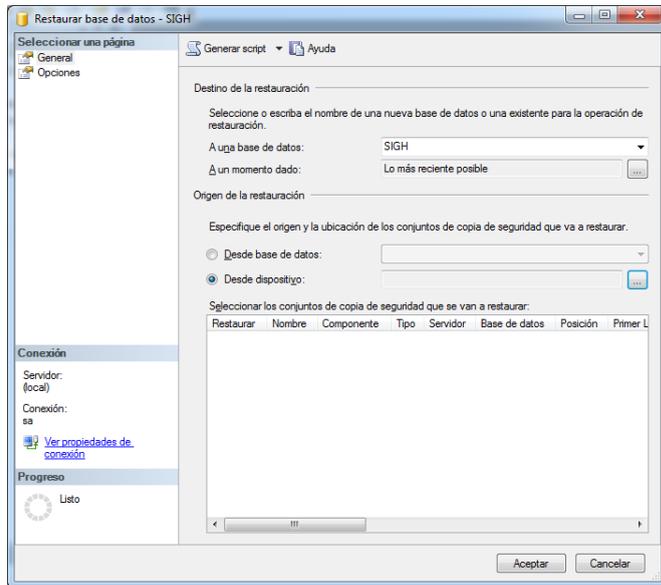
**Figura N° 73:** Opción de Nueva Base de datos.

Una vez creada la base de datos, hacemos *CLIC DERECHO en la Base SIGH*, seleccionamos *Tareas*, luego seleccionamos *Restaurar* y por ultimo seleccionamos *Base de datos...*



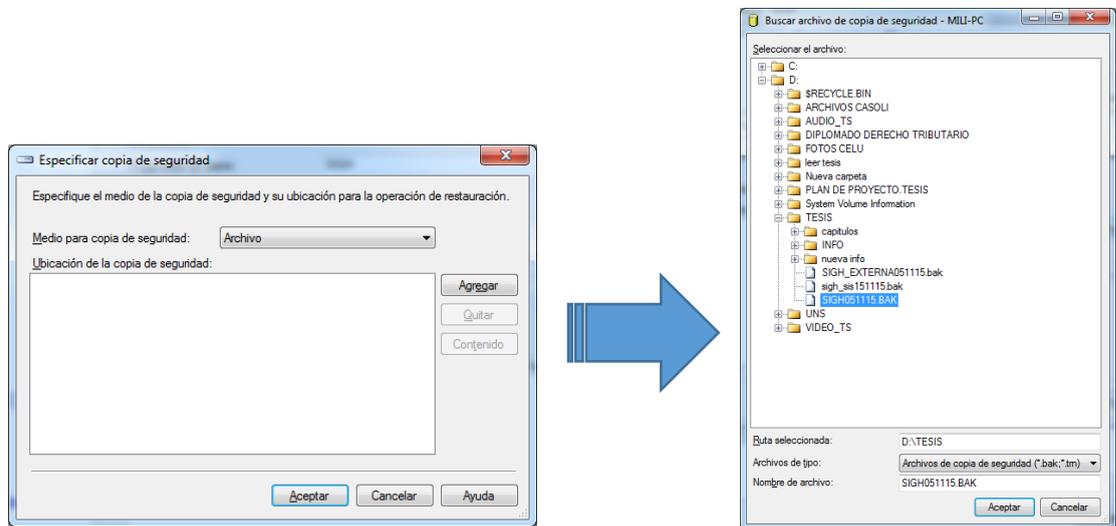
**Figura N° 74:** Opciones para restaurar la base de datos.

En la ventana “Restaurar base de datos” especificaremos el origen y la ubicación de los conjuntos de copia de seguridad que va a restaurar, para ello seleccionamos en “Desde dispositivo” y en hacemos CLIC en  .



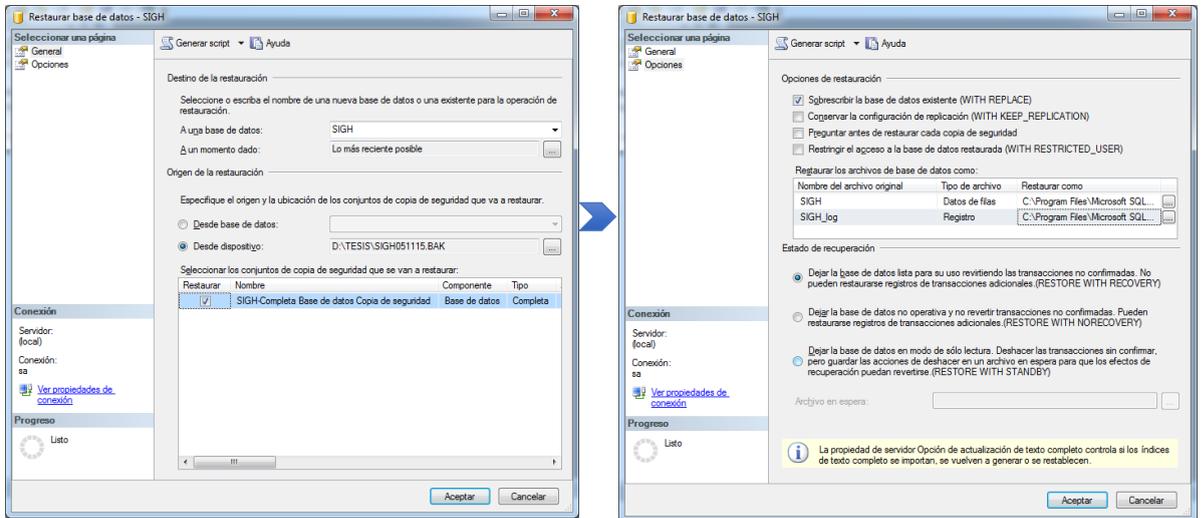
**Figura N° 75:** Restaurar base de datos -SIGH.

En la ventana “Especificar copia de seguridad”, hacemos CLIC en “Agregar” para que se nos muestre el cuadro de diálogo “Buscar archivo de copia de Seguridad”, seleccionamos el archivo .BAK y presionamos “Aceptar”.



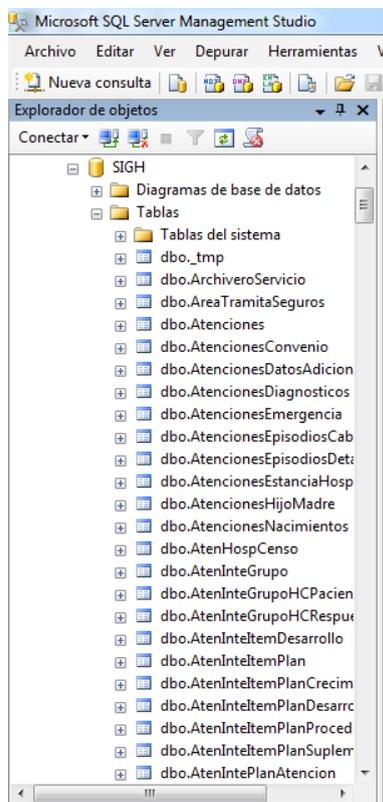
**Figura N° 76:** Especificación de copia de seguridad.

En la siguiente ventana en sección “Seleccionar los conjuntos de copia de seguridad que van a restaurar” debemos activar la casilla restaurar. En la parte Izquierda en el Ítem opciones debemos activar la casilla “**Sobrescribir la base de datos existente (WITH REPLACE)**” y “**ACEPTAR**”



**Figura N° 77:** Datos Generales de la restauración.

Finalmente después de Restaurar la Base de datos, podemos visualizar las tablas con las que vamos a trabajar.



**Figura N° 78:** Base de datos SIGH restaurada.

## 6.2 Creación de la base de datos DW en SQL Server 2008

Crearemos una base de datos para que almacene la Información del Data Warehouse con el que trabajaremos nuestra solución de Inteligencia de negocios.

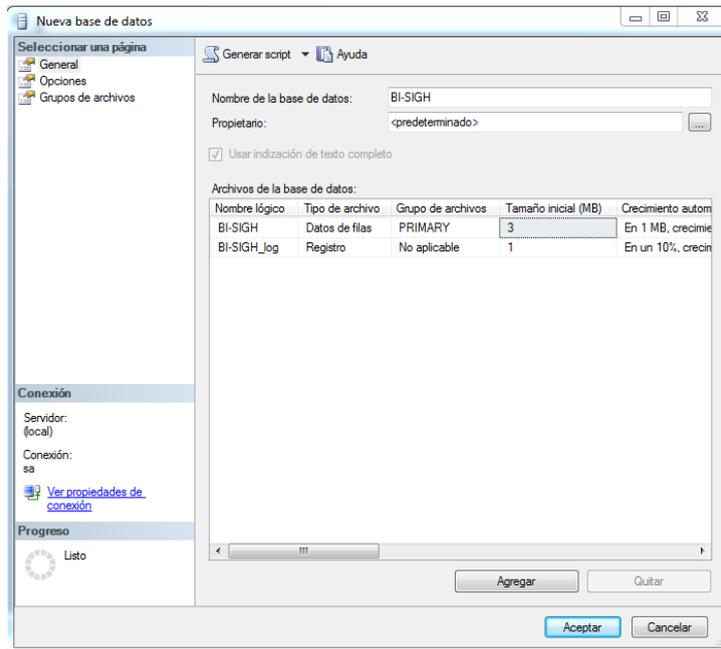


Figura N° 79: Creación del Data Warehouse BI-SIGH.

## ANEXO B

### CONEXIÓN A PENTAHO

#### 7.1 Configuración de la conexión de Pentaho con SQL Server 2008.

En la Carpeta **D:\PENTAHO\data-integration** se encuentra la herramienta Spoon con la cual se realiza los procesos de transformación ETL.

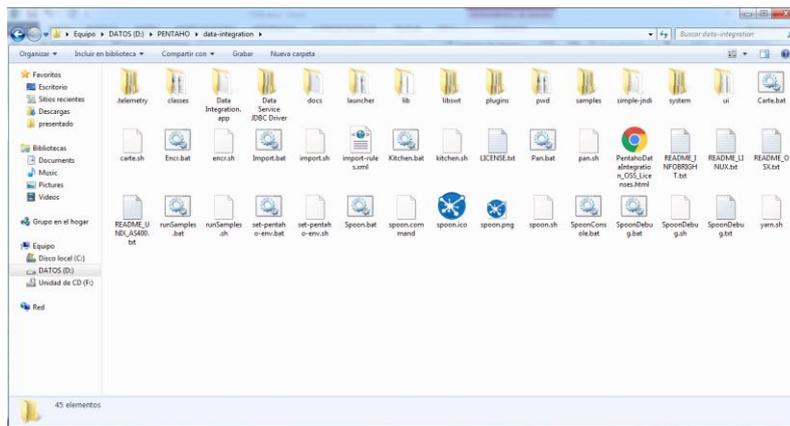


Figura N° 80: Carpeta Utilitarios Pentaho

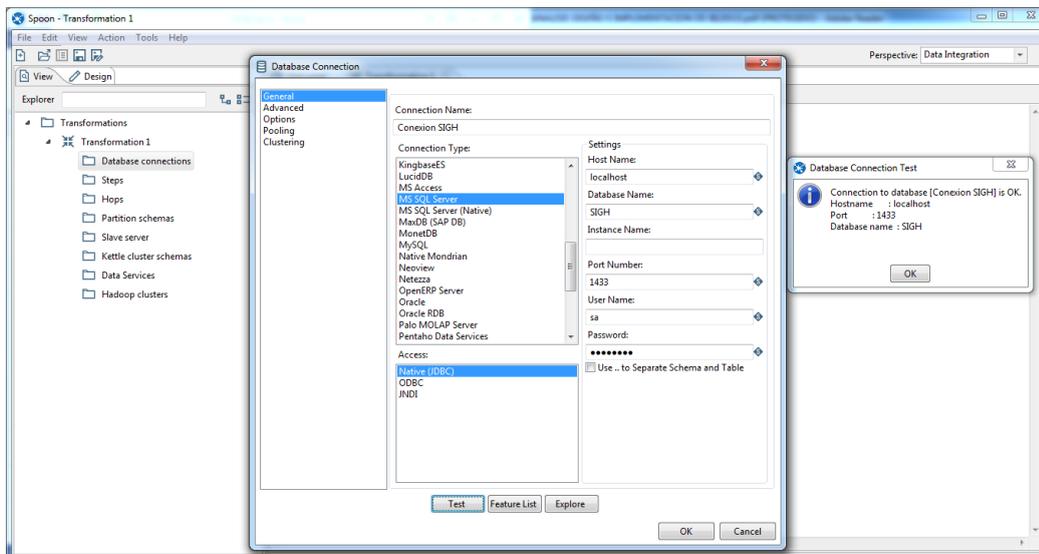
Se debe copiar el driver en la dirección **D:\PENTAHO\data-integration\lib** para la conexión entre Pentaho y la base de datos SQL Server 2008, el driver `sqljdbc4.jar` se descargó de la página oficial de Microsoft <https://docs.microsoft.com/es-es/sql/connect/jdbc/microsoft-jdbc-driver-for-sql-server>.

Pentaho Data Integration (kettle): dentro de la carpeta descomprimida **D:\PENTAHO\data-integration** ejecutar el archivo `Spoon.bat` que se encuentra en la dirección **D:\PENTAHO\data-integration** que iniciará la aplicación para generar gráficamente los procesos ETL.



Figura N° 81: Pentaho Data Integration

Se debe configurar la conexión para poder realizar los procesos de transformación en la herramienta en donde se encontraran los datos procesados.



**Figura N° 82:** Configuración Conexión Base de Datos

## ANEXO C

### CUESTIONARIOS

#### 8.1 Cuestionario para hallar el Nivel de Satisfacción de los Usuarios.

##### CUESTIONARIO DE NIVEL DE SATISFACCIÓN-SISTEMA ACTUAL

El presente cuestionario encontraras una serie de preguntas sobre aspectos referidos a la funcionalidad, servicio y calidad del sistema actual. El objetivo de estas preguntas es conocer el nivel de satisfacción de los usuarios con respecto al sistema actual.

**Indicaciones:** Para contestar marque en cada frase la opción que mejor indique tu grado de acuerdo o desacuerdo mediante la siguiente escala:

1=Desaprobación Plena; 2= Desaprobación Simple; 3= Indecisión o Indiferencia; 4= Aprobación Simple; 5= Aprobación Plena

Dimensión 1: Funcionalidad del Sistema						
Nº	ITEM	PUNTUACIÓN				
1	¿El sistema es de fácil uso?	1	2	3	4	5
2	¿Está usted satisfecho con la variedad de reportes que ofrece el sistema?	1	2	3	4	5
3	¿La interfaz del sistema es amigable para usted?	1	2	3	4	5
4	¿La información que le brinda el sistema le ayuda a tener una mejor visión de la situación de la institución?	1	2	3	4	5
5	¿Cómo califica la flexibilidad del sistema al momento de realizar reportes analíticos?	1	2	3	4	5
Dimensión 2: Servicio y Calidad						
6	¿Los reportes solicitados al área de estadística se entregan a tiempo?	1	2	3	4	5
7	¿Cómo considera la manera de presentar sus reportes analíticos?	1	2	3	4	5
8	¿Cómo considera el tiempo de elaboración de reportes personalizados por parte del área de estadística?	1	2	3	4	5
9	¿Cómo califica usted la calidad de los reportes entregados a la Dirección Ejecutiva?	1	2	3	4	5
10	¿Los reportes emitidos son fáciles de analizar?	1	2	3	4	5

  
Dr. Juan Miguel Ponce Josa Ph.D.

Figura N° 83: Cuestionario de Nivel de Satisfacción - Sistema Actual

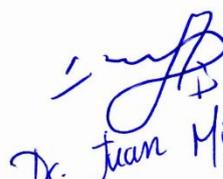
## CUESTIONARIO DE NIVEL DE SATISFACCIÓN-SOLUCION DE INTELIGENCIA DE NEGOCIOS

El presente cuestionario encontraras una serie de preguntas sobre aspectos referidos a la funcionalidad, servicio y calidad de la solución de inteligencia de negocios. El objetivo de estas preguntas es conocer el nivel de satisfacción de los usuarios con respecto a la solución de inteligencia de negocios.

**Indicaciones:** Para contestar marque en cada frase la opción que mejor indique tu grado de acuerdo o desacuerdo mediante la siguiente escala:

1=Desaprobación Plena; 2= Desaprobación Simple; 3= Indecisión o Indiferencia; 4= Aprobación Simple; 5= Aprobación Plena

Dimensión 1: Funcionalidad del Sistema						
N°	ITEM	PUNTUACIÓN				
1	¿La solución de BI es de fácil uso?	1	2	3	4	5
2	¿Está usted satisfecho con la variedad de reportes que ofrece la solución de BI?	1	2	3	4	5
3	¿La interfaz de la solución de BI es amigable para usted?	1	2	3	4	5
4	¿La información que le brinda el sistema le ayuda a tener una mejor visión de la situación de la institución?	1	2	3	4	5
5	¿Cómo califica la flexibilidad del sistema al momento de realizar reportes analíticos?	1	2	3	4	5
Dimensión 2: Servicio y Calidad						
6	¿Los reportes solicitados al área de estadística se entregan a tiempo?	1	2	3	4	5
7	¿Cómo considera la manera de presentar sus reportes analíticos?	1	2	3	4	5
8	¿Cómo considera el tiempo de elaboración de reportes personalizados por parte del área de estadística?	1	2	3	4	5
9	¿Cómo califica usted la calidad de los reportes entregados a la Dirección Ejecutiva?	1	2	3	4	5
10	¿Los reportes emitidos son fáciles de analizar?	1	2	3	4	5

  
 Dr. Juan Miguel Ponce Josa Ph.D.

**Figura N° 84:** Cuestionario de Nivel de Satisfacción - Solución BI

## 8.2 Cuestionarios para hallar la Calidad de Información.

### CUESTIONARIO DE CALIDAD DE INFORMACION-SISTEMA ACTUAL

El presente cuestionario encontraras una serie de preguntas sobre aspectos referidos a la calidad información que brinda el sistema actual. El objetivo de estas preguntas es conocer el nivel de calidad de la información que obtiene con ayuda de este sistema.

**Indicaciones:** Para contestar marque en cada frase la opción que mejor indique tu grado de acuerdo o desacuerdo mediante la siguiente escala:

1=Desaprobación Plena; 2= Desaprobación Simple; 3= Indecisión o Indiferencia; 4= Aprobación Simple; 5= Aprobación Plena

N°	ITEM	PUNTUACIÓN				
		1	2	3	4	5
1	¿Se le ofrece datos relevantes para la toma de decisiones?	1	2	3	4	5
2	¿Considera que los datos brindados cuentan con el nivel de detalle que requiere?	1	2	3	4	5
3	¿Hay exactitud en los datos?	1	2	3	4	5
4	¿La información recibida tiene datos actualizados?	1	2	3	4	5
5	¿Es fácil la comprensión de los datos?	1	2	3	4	5
6	¿La información brindada para la toma de decisiones se le brinda completa?	1	2	3	4	5

  
Dr. Juan Miguel Percejoza Ph.D.

**Figura N° 85:** Cuestionario de Calidad de Información - Sistema Actual

**CUESTIONARIO DE CALIDAD DE INFORMACION- SOLUCION DE INTELIGENCIA DE NEGOCIOS**

El presente cuestionario encontraras una serie de preguntas sobre aspectos referidos a la calidad información que brinda la solución de inteligencia de negocios. El objetivo de estas preguntas es conocer el nivel de calidad de la información que obtiene con ayuda de la solución de inteligencia de negocios implementada.

**Indicaciones:** Para contestar marque en cada frase la opción que mejor indique tu grado de acuerdo o desacuerdo mediante la siguiente escala:

1=Desaprobación Plena; 2= Desaprobación Simple; 3= Indecisión o Indiferencia; 4=Aprobación Simple; 5= Aprobación Plena

N°	ITEM	PUNTUACIÓN				
		1	2	3	4	5
1	¿Se le ofrece datos relevantes para la toma de decisiones?	1	2	3	4	5
2	¿Considera que los datos brindados cuentan con el nivel de detalle que requiere?	1	2	3	4	5
3	¿Hay exactitud en los datos?	1	2	3	4	5
4	¿La información recibida tiene datos actualizados?	1	2	3	4	5
5	¿Es fácil la comprensión de los datos?	1	2	3	4	5
6	¿La información brindada para la toma de decisiones se le brinda completa?	1	2	3	4	5

  
 Dr. Juan Miguel Ponce Loza Ph.D.

**Figura N° 86:** Cuestionario de Calidad de Información - Solución BI

## ANEXO D

### CONFIABILIDAD POR ALFA DE CRONBACH

#### 9.1 Encuesta para indicador: nivel de satisfacción de los usuarios

Para asegurar la consistencia de los datos de la encuesta realizada a los usuarios (dirección ejecutiva), se calculó el coeficiente de Alfa de Cronbach.

**Tabla N° 31:** Estadísticos de la encuesta para el Indicador Nivel de Satisfacción.

Directivos	Preguntas										Suma de Ítems
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	
U1	3	2	2	3	3	3	4	2	3	3	28
U2	2	2	2	2	2	2	2	3	3	3	23
U3	1	1	1	1	1	1	2	1	2	1	12
U4	3	3	2	2	3	3	3	3	2	2	26
U5	1	2	1	1	1	1	1	1	2	2	13
U6	2	3	2	2	2	2	3	1	4	3	24
Varianza Típica	0.67	0.47	0.22	0.47	0.67	0.67	0.92	0.81	0.56	0.56	38.67

A partir de las varianzas, el alfa de Cronbach se calcula así:

$$\alpha = \left[ \frac{k}{k-1} \right] \left[ 1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right],$$

Donde:

$k$ : El número de ítems=10

$\sum_{i=1}^k S_i^2$ : Sumatoria de las Varianzas de los Items= 6.00

$S_t^2$ : La Varianza de la suma de los Items=38.67

$\alpha$ : El número de ítems= 0.938

Teniendo en cuenta el valor hallado de  $\alpha= 0.938$  y de acuerdo a la escala Alfa de Cronbach (Coeficiente alfa  $> 0.9$  es excelente). Se concluye que el instrumento y por consiguiente los datos, son fiables.

## 9.2 Encuesta para indicador: calidad de la información

Para asegurar la consistencia de los datos de la encuesta realizada a los usuarios (dirección ejecutiva), se calculó el coeficiente de Alfa de Cronbach.

**Tabla N° 32:** Estadísticos de la encuesta para el Indicador Calidad de la Información.

Directivos	Preguntas						Suma de Ítems
	P1	P2	P3	P4	P5	P6	
U1	2	3	2	3	3	2	15
U2	2	2	2	2	3	2	13
U3	1	1	2	1	1	1	7
U4	2	2	1	2	2	2	11
U5	2	2	2	1	2	1	10
U6	1	1	1	2	1	2	8
Varianza Típica	0.22	0.47	0.22	0.47	0.67	0.22	7.56

A partir de las varianzas, el alfa de Cronbach se calcula así:

$$\alpha = \left[ \frac{k}{k-1} \right] \left[ 1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right],$$

Donde:

$k$ : El número de ítems=6

$\sum_{i=1}^k S_i^2$ : Sumatoria de las Varianzas de los Items= 2.28

$S_t^2$ : La Varianza de la suma de los Items=7.56

$\alpha$ : El número de ítems= 0.838

Teniendo en cuenta el valor hallado de  $\alpha= 0.838$  y de acuerdo a la escala Alfa de Cronbach (Coeficiente alfa  $> 0.8$  es bueno). Se concluye que el instrumento y por consiguiente los datos, son fiables.