



UNIDAD ACADÉMICA:

DEPARTAMENTO DE INVESTIGACIÓN Y POSTGRADOS

TEMA:

DESARROLLO DE UN BUSINESS INTELLIGENCE EN SOFTWARE LIBRE, BASADO
EN INDICADORES DE GESTIÓN, PARA UNA COORDINACIÓN DE SALUD.

**Proyecto de Investigación y Desarrollo previo a la obtención del título de
Magíster en Gerencia Informática**

Línea de Investigación, Innovación y Desarrollo principal:

Ingeniería de Software y/o Plataformas Educativas

Caracterización técnica del trabajo:

Desarrollo

Autor:

Fabián Patricio Londo Yachambáy

Director:

Enrique Xavier Garcés Freire Ing. Master

Ambato – Ecuador

Junio 2015

DESARROLLO DE UN BUSINESS INTELLIGENCE EN SOFTWARE LIBRE, BASADO
EN INDICADORES DE GESTIÓN, PARA UNA COORDINACIÓN DE SALUD.

Guía para la Preparación de Informes de Trabajos de Titulación para Postgrados (Tema del Trabajo)

Informe de Trabajo de Titulación

Presentado ante la

Pontificia Universidad Católica del Ecuador
Sede Ambato

Por:

Fabián Patricio Londo Yachambáy Ing.

En cumplimiento parcial de los
requisitos para el Grado de
Magister en Gerencia Informática



Departamento de Investigación y Postgrados

Junio 2015

DESARROLLO DE UN BUSINESS INTELLIGENCE EN SOFTWARE LIBRE, BASADO
EN INDICADORES DE GESTIÓN, PARA UNA COORDINACIÓN DE SALUD.

Guía para la Preparación de Informes de Trabajos de Titulación para Postgrados (Tema del Trabajo)

Aprobado por:

Varna Hernández Junco, PhD
Presidente del Comité Calificador
Director DIP

Darío Robayo Jácome, Mg
Miembro Calificador

Enrique Xavier Garcés Freire, Mg
Miembro Calificador
Director de Proyecto

Dr. Hugo Altamirano Villaroel
Secretario General

Liliana Mena H., Mg
Miembro Calificador

Fecha de aprobación:
Junio 2015

Ficha Técnica

Programa: Magister en Gerencia Informática

Tema: Desarrollo de un Business Intelligence en software libre, basado en indicadores de gestión, para una Coordinación de Salud.

Tipo de trabajo: Tesis (alternativamente Propuesta Metodológica y Tecnológica Avanzada o Diseño de Modelo Complejo)

Clasificación técnica del trabajo: Desarrollo (alternativamente Investigación o Innovación)

Autor: Fabián Patricio Londo Yachambáy

Director: Enrique Xavier Garcés Freire, Ing. Magister

Líneas de Investigación, Innovación y Desarrollo

Principal: Ingeniería de Software y/o Plataformas Educativas

Resumen Ejecutivo

La nueva estructuración de Zonas, Distritos y Circuitos que lleva adelante el estado ecuatoriano ha generado una centralización de información hacia las 9 zonas que tiene el país, las cuales se están fortaleciendo y consolidando en sus diferentes ámbitos. Sin embargo es notoria la falta de aplicativos que ayuden a analizar y procesar información para tomar decisiones oportunas y adecuadas en el momento indicado. El desarrollo de un BI (Inteligencia de Negocios) usando software libre como estipula la política de gobierno, permite realizar el procesamiento y análisis de la información de manera rápida, fiable y oportuna mediante la utilización de técnicas estadísticas, para ello se cuenta con una base de indicadores que permitirán tomar decisiones en la parte administrativa y gerencial, el sistema de BI (Inteligencia de Negocios) está dirigido a los Directores de procesos de las Coordinaciones Zonales de Salud quienes son los encargados de tomar decisiones para que la Salud Pública sea de calidad y llegue a todos los ecuatorianos. El Producto final del trabajo son reportes realizados en Pentaho que permiten visualizar estadísticas de salud, extraídos a través de una concentración, procesamiento y análisis de información de todas las unidades operativas de primero, segundo y tercer nivel que tiene cada Coordinación Zonal de Salud.

Declaración de Originalidad y Responsabilidad

Yo, Fabián Patricio Londo Yachambáy, portador de la cédula de ciudadanía y/o pasaporte No. 0603016114, declaro que los resultados obtenidos en el proyecto de titulación y presentados en el informe final, previo a la obtención del título de Magister en Gerencia Informática, son absolutamente originales y personales. En tal virtud, declaro que el contenido, las conclusiones y los efectos legales y académicos que se desprenden del trabajo propuesto, y luego de la redacción de este documento, son y serán de mi sola y exclusiva responsabilidad legal y académica.

Fabián Patricio Londo Yachambáy
CI. 060301611-4

Dedicatoria

A Dios, por haberme dado la posibilidad de existir y descubrir su infinita bondad. A mí querida esposa por tener siempre su apoyo incondicional, y a Steven mi adorado hijo inspiración de mis sueños y metas, quienes además fueron la motivación y el empuje para culminar este trabajo.

Fabián Londo

Reconocimientos

A todo el personal que conforman la PONTIFICIA UNIVERSIDAD CATÓLICA SEDE AMBATO, que de una u otra forma me extendieron su mano para enriquecer mi conocimiento y calidad humana, en especial al Ingeniero Enrique Xavier Garcés Freire, Mg. por dirigir éste proyecto de investigación.

Resumen

En la actualidad tanto empresas públicas como privadas, presentan dificultades a la hora de analizar sus datos. El avance de las tecnologías de la información, el desarrollo acelerado de los sistemas de comunicación y el desarrollo socioeconómico, ha determinado que las empresas vean la imperiosa necesidad de mejorar cada día su competitividad para no quedarse relegadas en caminos secundarios. Este horizonte se refleja, en determinaciones estratégicas de un crecimiento substancial de la información y su capacidad de almacenar, procesar y analizar. Este incremento involucra que en las diferentes organizaciones sus sistemas de almacenamiento contengan grandes cantidades de datos, pero con un déficit total de valor informativo. Los datos por sí solos carecen de valor, estos deben ser depurados y procesados para que puedan tener una consistencia y dar el valor agregado que es necesario para cualquier institución o compañía. Ante esta problemática se ha construido soluciones de BI (Inteligencia de Negocios) bajo el software libre Pentaho, que a su vez utiliza una base de datos elaborada bajo PostgreSQL, la base está conformada con datos de atención en el sector salud que originara un almacén de datos DW (Data Warehouse). Pentaho permite realizar consultas de información y análisis según los requerimientos que se necesite, el BI construido bajo plataforma libre ayuda a fortalecer el sistema de información del Departamento de Estadística de una Coordinación Zonal de Salud.

Palabras Clave: Inteligencia de Negocios, PostgreSQL, Data Warehouse, Pentaho.

Abstract

At present both public and private companies are having difficulties in analyzing their data. The advancement of information technology, the rapid development of communication systems and socioeconomic development, has determined that companies see the urgent need to improve their competitiveness to not be pushed aside every day not to stay competitive relegated to secondary roads. This horizon is reflected in strategic determinations of substantial growth of information and its ability to store, process and analyze. This increase means that the storage systems in different organizations contain large amounts of data, but with a total deficit of informative value. The data alone are worthless, they must be refined and processed so that they can have a consistency and provide the added value that is necessary for any institution or company. In light of this problem, BI (Business Intelligence) solutions have been created under the open source software Pentaho, which in turn uses a database created under PostgreSQL, the base of data about the service in the health care sector would lead to a data storage called DW (Data Warehouse). Pentaho makes queries of information and analysis possible according to the requirements that are needed. The BI platform built on a free platform will help to strengthen the information system of the Department of Statistics of a health management zone.

Keywords: business intelligence, PostgreSQL, Data Warehouse, Pentaho.

Tabla de Contenidos

Ficha Técnica	iii
Declaración de Originalidad y Responsabilidad	iv
Dedicatoria	v
Reconocimientos	vi
Resumen	vii
Abstract	viii
Tabla de Contenidos	ix
Lista de Tablas	xi
Lista de Figuras	xii
CAPÍTULOS	
1. Introducción	14
1.1. Presentación del trabajo	14
1.2. Descripción del documento	15
2. Planteamiento de la Propuesta de Trabajo	16
2.1. Información técnica básica	16
2.2. Descripción del problema	16
2.3. Preguntas básicas	16
2.4. Formulación de meta	17
2.5. Objetivos	17
2.6. Delimitación funcional	17
3. Marco Teórico	18
3.1. Definiciones y Conceptos	18
3.1.1 Coordinación Zonal de Salud	18
3.1.2 Distribución de las Zonas de Salud	18
3.1.3 Indicadores	20
3.1.4 Business Intelligence	22
3.1.5 Tecnologías que forman parte del BI	22
3.1.6 Software Libre	26
3.1.7 Pentaho	27
3.1.8 Muestreo	31
3.2. Estado del Arte	33
4. Metodología	36
4.1 Diagnóstico	36

4.2	Método aplicado.....	36
4.3	Materiales y herramientas	55
5.	Resultados	57
5.1.	Producto final del Proyecto de titulación	57
5.2.	Evaluación preliminar	64
	Introducción.....	65
	- Propósito del Plan.....	65
	- Alcance.....	65
	Requerimientos de Pruebas.....	65
	Introducción.....	65
	Filosofía de la prueba.....	65
	- Generalidades	65
	- Áreas funcionales.....	65
	- Categorías de resultados de prueba	66
	Entorno de la prueba	66
	- Generalidades	66
	- Hardware.....	66
	- Software	67
	Datos de prueba	67
	Roles y responsabilidades del equipo de pruebas	67
	Identificación de la prueba	67
	- Documentación de la prueba	67
	Requerimientos del Ambiente de Pruebas.....	70
	Hardware	70
	Sistema Operativo	70
5.3.	Análisis de resultados.....	75
6.	Conclusiones y Recomendaciones	77
6.1.	Conclusiones.....	77
6.2.	Recomendaciones.....	78
	Apéndice A	79
	Apéndice B	81
	Apéndice C	87
	Apéndice D	91
	REFERENCIAS	92

Lista de Tablas

1. Historial	64
2. Recursos	66
3. Roles y Responsabilidades	67
4. Criterios	69
5. Importancia de un BI	70
6. Tiempo de Respuesta a Consultas.....	71
7. Respuesta a Consultas	72
8. Facilidad de Uso	73
9. Reportes Entendibles.....	74

Lista de Figuras

1. Zonas de Planificación del Ecuador.....	19
2. Sistema de Indicadores.....	20
3. Generación de Conocimiento un BI.....	22
4. Origen de datos en un Sistema BI.....	25
5. Libertades del Software Libre.....	26
6. Logotipo de Pentaho (PDI).....	27
7. Módulos de Pentaho.....	29
8. Esquema de un Data Warehouse.....	30
9. Esquema de la Metodología Hefesto.....	37
10. Esquema del Modelo Conceptual.....	39
11. Diagrama Entidad de Relación.....	40
12. Modelo Conceptual Ampliado.....	41
13. Tabla Dimensión "ZONA".....	42
14. Tabla Dimensión "DISTRITO".....	42
15. Tabla Dimensión "PROVINCIA".....	43
16. Tabla Dimensión "TIEMPO".....	43
17. Tablas de Hechos.....	44
18. Caso uniones.....	45
19. Carga Inicial.....	46
20. Ventana de PostgreSQL.....	47
21. Creación de la Base de Datos.....	47
22. Selección de nombre para la Base de Datos.....	48
23. Creación de las tablas de entrada.....	49
24. Creación de las Steps de entrada y salida.....	49
25. Conexión de Steps.....	50
26. Entrada de datos al Input.....	50
27. Creación de Campos.....	51
28. Llenado de datos en los campos creados.....	51
29. Conexión a la Base de Datos.....	52
30. Esquema para escoger el tipo de Base de Datos.....	52
31. Nombre de la tabla a crear.....	53
32. Nombre de la tabla a crear.....	53
33. Tabla Generada.....	54
34. Ejecución del ETL.....	54
35. Consola de Pentaho.....	57
36. Ingreso a Pentaho.....	58
37. Menú Create New.....	58
38. Modulo Analysis Reporty.....	59
39. Elaboración del Reporte.....	59
40. Reportes con Gráficos.....	60
41. Total de partos.....	61
42. Total de Altas.....	61
43. Total de Altas (Tabla).....	62
44. Dotación de Camas.....	62
45. Nivel de Ocupación.....	63
46. Interfaz de Información.....	64
47. Consola de Pentaho.....	68

48.	Resultados de la prueba (Partos)	68
49.	Resultados de la Prueba (Ocupación).....	69
50.	Importancia de un BI	71
51.	Tiempo de Respuesta a Consultas.....	72
52.	Respuestas Consistentes	73
53.	Uso del Sistema.....	74
54.	Reportes Importantes.....	75
55.	Dirección web para descargar PostgreSQL.....	81
56.	Instalador de PostgreSQL descargado.....	82
57.	Ventana de Instalación de PostgreSQL.....	82
58.	Ventana para el ingreso de contraseña para PostgreSQL.....	83
59.	Ingreso del puerto a utilizar.	84
60.	Directorio de Instalación.....	84
61.	Finalización de la instalación.	85
62.	Pantalla de inicio de PostgreSQL.	85
63.	Pantalla para trabajar con PostgreSQL.	86
64.	Pantalla de descarga de Pentaho.	87
65.	Ruta de la instalación de Pentaho.	88
66.	Descarga de Pentaho.	88
67.	Proceso de instalación.....	89
68.	Carpeta de Pentaho.....	89
69.	Carpeta para pegar Driver descargado.	90
70.	Consola de Pentaho.....	91

Capítulo 1

Introducción

1.1. Presentación del trabajo

En la actualidad la estructuración de Zonas, Distritos y Circuitos que lleva adelante el Estado ecuatoriano ha generado una centralización de información hacia las 9 zonas que tiene el país en diferentes ámbitos y uno ellos es el sector salud, por lo cual se ha generado la necesidad de contar con un aplicativo que sea capaz de concentrar información que es generada día a día por los miles de pacientes que acuden a recibir atención médica en diferentes unidades y en distintos servicios, por esta razón se ha construido un BI bajo Pentaho que tiene características de software libre.

El diseño del BI ayuda a concentrar la información de diferentes programas y proyectos de manera eficiente en una sola base de datos PostgreSQL para lo cual utilizaron procesos de extracción, transformación y carga, los cuales me permitieron centralizar la información, que servirá de insumo para que Pentaho mediante la utilización de sus herramientas realice el análisis de la información concentrada y se obtengan los indicadores de gestión, que van a ser de mucha utilidad para la toma de decisiones.

La estrategia utilizada para el desarrollo del BI fue primeramente hacerlo en un software que siga las políticas actuales de Gobierno, refiriéndose a la utilización de software libre, de igual forma se debía utilizar una base de datos que tengan similares características, es así que se utilizó *Pentaho* y PostgreSQL respectivamente, fue también determinante escoger indicadores apropiados para Directivos, que permitan tener una visión clara para la toma de decisiones, esto ante el hecho de que existe un sinnúmero de indicadores que se pueden obtener de la información de salud.

Se ha logrado obtener el diseño de un BI construido bajo el software libre *Pentaho* y que tiene toda la información concentrada en una sola base, de esta manera conseguirán cubrir cualquier requerimiento de información e indicadores que sea necesario para la toma de decisiones.

1.2. Descripción del documento

En el capítulo 2 se plantea la propuesta de trabajo, en donde se determina una breve descripción del problema y el planteamiento de objetivos. El marco teórico es abordado en el capítulo 3 enfocándose en la sección 3.1 lo que es: un BI, software libre, Pentaho y sus principales aplicaciones, y los indicadores de salud más importantes, en la sección 3.2 se establece el estado del arte con investigaciones semejantes al tema de estudio. En el capítulo 4 se determina la metodología utilizada para la investigación que tiene como aristas principales los análisis de Síntesis y Métodos Inductivos, también se utilizó en el inciso 4.3 como instrumento de medición encuestas para determinar la aceptación y utilización de Pentaho. En el capítulo 5 se presentan los resultados obtenidos con la utilización de Pentaho, ventanas que muestran cómo se obtuvieron los indicadores de salud planteados. Las conclusiones y recomendaciones forman el capítulo 6 para una coordinación Zonal de Salud, en tanto que la principal recomendación que se realiza es la utilización y capacitación sobre BI en las diferentes empresas del sector público.

El trabajo se lo ha completado con tres Apéndices. En el Apéndice A consta la encuesta realizada a los profesionales de la Coordinación Zonal 3 de salud sobre la aceptabilidad y manejo de Pentaho. En tanto que el apéndice B se evidencia la documentación de las pruebas realizadas a la aplicación, en el apéndice C se documenta direcciones y forma de instalar PostgreSQL y Pentaho.

Capítulo 2

Planteamiento de la Propuesta de Trabajo

2.1. Información técnica básica

Tema: Desarrollo de un Business Intelligence en software libre basado en indicadores de gestión para una Coordinación de Salud

Clasificación técnica del trabajo: Tesis

Líneas de Investigación, Innovación y Desarrollo

Principal: Ingeniería de Software y/o Plataformas Educativas

2.2. Descripción del problema

Un BI ayudará a contar con información de primera mano es decir datos de diferentes fuentes y distintos programas concentrados en un solo lugar, y que al momento de requerirla para hacer un análisis de información o requerimiento de indicadores permita actuar y procesar datos ante las necesidades existentes, evitando causar pérdidas de tiempo y recursos al momento de solicitar indicadores de salud para la toma de decisiones.

2.3. Preguntas básicas

¿Por qué se origina? Por la cantidad de información que acumula en las Coordinaciones de Salud, información generada en las distintas unidades que brindan atención en salud

¿Qué lo origina? La falta de sistemas o aplicativos informáticos que ayude a consolidar y analizar información de salud.

¿Cuándo se origina? Cuando se comienza a registrar información en los formularios físicos de atención a los pacientes en los tres niveles de atención en salud

¿Dónde se origina? En las unidades de salud pertenecientes al Ministerio de Salud

¿Dónde se detecta? En las Coordinaciones de Salud que pertenecen al Ministerio de Salud.

2.4. Formulación de meta

Contar con una base de indicadores sustentable, que permitan tomar decisiones oportunas y adecuadas en periodos de tiempo definidos.

2.5. Objetivos

Objetivo general.-

Desarrollar un Business Intelligence en software libre, basado en indicadores de gestión, para una Coordinación de Salud.

Objetivos específicos.-

- Consolidar y analizar información a través de técnicas estadísticas.
- Obtener una base de indicadores de salud para tomar decisiones.
- Analizar la estructura del Pentaho y su pertinencia con un Business Intelligence.
- Desarrollar una aplicación basada en indicadores de gestión para la toma de decisiones.

2.6. Delimitación funcional

Pregunta 1. ¿Qué será capaz de hacer el producto final del trabajo de titulación?

- Recolección y consolidación de información
- Procesamiento y análisis de la información.
- Reportes de indicadores de salud
- Entrega de reporte en formato PDF al departamento de TIC.
-

Capítulo 3

Marco Teórico

3.1. Definiciones y Conceptos

3.1.1 Coordinación Zonal de Salud

Las Coordinaciones Zonales de Salud son niveles administrativos, que en base a una planificación, permiten conocer necesidades y ayudan a dar soluciones de forma eficaz en prestación de servicios públicos de salud. En la actualidad Ecuador tiene nueve zonas de planificación. Cada Zona está constituida por Distritos y estos a su vez por Circuitos. Desde este nivel se coordina estratégicamente las entidades del sector público, a través de la gestión de la planificación para el diseño de políticas de salud en el área de su jurisdicción. [1]

3.1.2 Distribución de las Zonas de Salud

La Secretaria de Nacional de Planificación y Desarrollo (SENPLADES) ha establecido las nueve zonas de la siguiente forma:

Zona1: Provincia de Esmeraldas, Imbabura, Carchi y Sucumbíos.

Zona2: Provincias de Pichincha (excepto el catón Quito), Napo y Orellana.

Zona3: Provincias de Cotopaxi, Tungurahua, Chimborazo y Pastaza.

Zona4: Provincias de Manabí y Santo Domingo de los Tsáchilas.

Zona5: Provincias de Santa Elena, Guayas (excepto los cantones de Guayaquil y Samborondón y Durán) Bolívar, Los Ríos y Galápagos.

Zona6: Provincias de Cañar, Azuay y Morona Santiago.

Zona7: Provincias de El Oro, Loja y Zamora Chinchipe.

Zona8: Cantones de Guayaquil, Samborondón y Durán.

Zona9: Distrito Metropolitano de Quito.

En el mapa que está a continuación se puede visualizar de mejor manera la distribución de las nueve zonas.

Figura 1: Zonas de Planificación del Ecuador.



Fuente: <https://planificacion.gob.ec>

3.1.2.1 Distrito

Luego de las Zonas, los Distritos son la siguiente unidad básica de planificación que tiene como principal función la administración y prestación de servicios que son dirigidos a toda la ciudadanía perteneciente a esa jurisdicción. Según la SENPLADES se ha conformado 140 Distritos en el país. En promedio cada Distrito abarca 90000 (noventa mil habitantes) aproximadamente. Sin embargo existen cantones con una alta densidad poblacional como son: Quito, Guayaquil, Cuenca, Ambato y Santo Domingo de los Tsáchilas donde se ha considerado establecer distritos dentro de cada uno de estos cantones. [1]

3.1.2.2 Circuito

Posterior a los Distritos de planificación la Secretaria Nacional de Planificación y Desarrollo ha dispuesto la creación de unidades más pequeñas de administración que son los Circuitos, con el propósito que el conjunto de servicios públicos esté más cerca a los ciudadanos. En si un Distrito está formado por parroquias o varias de ellas, en el país existen 1134 circuitos con un promedio de 12000 (doce mil habitantes) aproximadamente. [1]

3.1.3 Indicadores

3.1.3.1 Definición de Indicador

Un indicador es un valor que representa el resultado de hechos referentes a acontecimientos de situaciones o acciones realizadas, y que mediante el tiempo y el espacio permitan establecer comparaciones o cambios de los mismos.

Un indicador debe considerar los siguientes requisitos para ser ideal:

- Tener validez (midan lo que defina el indicador)
- Ser definidos (tener resultados similares, sea que se apliquen a poblaciones con características diferentes)
- Ser específicos (determinen cambios de una situación en estudio) [2]

Se define como indicador a una medida que determina o cuantifica la eficacia o eficiencia de una actividad realizada en un periodo de tiempo. El definir o estimar un indicador no es una tarea sencilla, involucra procedimientos que van a requerir varios recursos como son humanos, tecnológicos y económicos, para que la calidad del indicador sea de entera satisfacción. [3]

En la siguiente grafica se puede observar los pasos que conllevan a la obtención de un indicador.

Figura 2: Sistema de Indicadores



Fuente: Moya Ligia, 2005

Se define como indicador a una proposición capaz de identificar características observables que permita determinar dimensiones, siempre que se fundamente en un análisis teórico que permita realizar un análisis y nos determine descripciones y comparaciones.[4]

3.1.3.2 Indicadores de Gestión

Representan la cuantificación de variables en estudio y permite tener una medición de coberturas de servicios, calidad y satisfacción de productos utilizados y el impacto de solución que se ha realizado pensando en la población y sus necesidades.

Este tipo de indicadores va a permitir medir situaciones que a lo posterior necesiten decisiones que definan cambios necesarios. [5]

3.1.3.2.1 Indicadores de Gestión para Salud

- Altas.- Representa el número total de altas que se dieron a pacientes que hicieron uso del servicio de salud.
- Egresos: Determina el número de pacientes que fueron atendidos en una unidad de salud, pero luego de encontrarse en un estado estable abandonaron la unidad de salud o a su vez salieron por fallecimiento.
- Total días de estadía.-Representa el total de días que los pacientes pernotaron en una unidad operativa de salud haciendo uso de una cama.
- Total días paciente.- determina el número de días que un paciente permaneció atendido en una unidad de salud.
- Giro de camas.- Número de pacientes egresados por cada cama durante un periodo de tiempo
- Diario días paciente.- Número promedio de pacientes hospitalizados que recibieron atención diaria durante un periodo.
- Porcentaje de ocupación.- Porcentaje de camas que estuvieron ocupadas diariamente durante un periodo.
- Porcentaje días de estada.- Representa el promedio de días que los pacientes permanecieron hospitalizados.

- Porcentaje diario de camas disponibles.- Representa las camas disponibles que existen diariamente en términos de porcentaje.
- Partos.- Es el número de partos que se dieron en un periodo de tiempo.[6]

3.1.4 Business Intelligence

Un BI es la técnica que transformar datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios.

En el siguiente gráfico se puede observar que es lo que genera un BI.

Figura 3: Generación de Conocimiento un BI



Fuente: Moya Ligia, 2005

La tecnología BI es una tecnología que permite analizar la información vital de la empresa, a partir de distintas variables y perspectivas, obteniendo una visión global y de conjunto, que de otra forma permanecería oculta por la complejidad de su análisis. [7]

Un BI, son el conjunto de técnicas, métodos y aplicaciones informáticas que se desarrollan de acuerdo a la necesidad que existe por tener una adecuada administración de información y que a su vez permita ser eficientes a la hora de generar reportes para la toma de decisiones. [8]

3.1.5 Tecnologías que forman parte del BI

Algunas tecnologías que forman parte del BI son:

- Data Warehouse

Un DW (Data Warehouse) es la reunión de datos para un determinado negocio, que puede ser: integrada, variante dentro de periodos de tiempo establecido y no volátil dentro de un proceso de toma de decisiones.

- Reporting

Son herramientas que permiten la generación de informes y listados, tanto en detalle como en forma general

- Análisis OLAP (On Line Analytical Processing)

Son aplicaciones que permiten recoger los datos de una organización, almacenarlos y realizar consultas. Para lo cual se crea una capa de negocio con lenguaje funcional por encima de la base de datos.

- Análisis Visual

Son herramientas que permiten a los usuarios realizar búsquedas rápidas de información a través de la interacción con cuadros de control dinámicas y analíticas desde cualquier dispositivo.

- Análisis Predictivo

Es un tipo de análisis descriptivo que relaciona un conjunto de variables, involucra gestión, tecnologías de la información y modelado, involucrando gran cantidad de datos.

- Cuadro de Mando Integral

Son herramientas que reúnen gran cantidad de información en entornos muy llamativos y prácticos, a través del uso de gráficos, mapas entre otros elementos. Los cuadros de mando integral permiten ver fácil complejos modelos de datos.

- Minería de Datos

Son aplicaciones que permiten extraer información de las bases de datos a través de patrones o tendencias, según los requerimientos. Existe una gran variedad de técnicas de minería de datos que determinan distintos tipos de patrones. Algunas técnicas son métodos estadísticos y redes neuronales.

- Gestión del Rendimiento

Son aplicaciones que describen metodología, métricas, procesos y sistemas empleados para supervisar y gestionar el rendimiento de una empresa.

- Previsiones

Las previsiones son aplicaciones que determinan predicciones a través del estudio de series de tiempo.

- Dashboards

Son aplicaciones que permiten obtener resúmenes visuales de información es decir que a una primera visualización tengo una comprensión general del negocio, a través de métricas e indicadores.

- Integración de Datos

Es una herramienta que permite consolidar datos de diferentes fuentes mediante procedimientos de extracción, transformación y carga. [8]

3.1.5.1 Beneficios de un BI

Implementar un BI genera un sinnúmero de beneficios, los más importantes están:

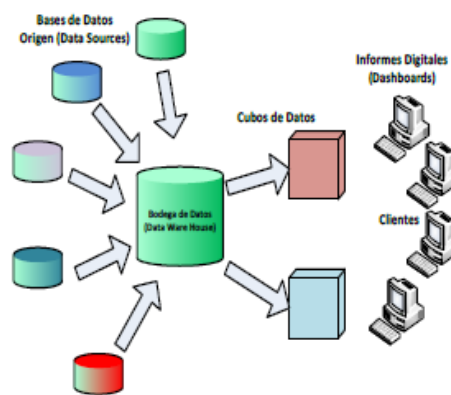
- a) Define un entorno único de información, que permite a los datos transformarse en información de interés, con perspectivas nuevas y mejores, por consiguiente servirá de utilidad para una eficiente toma de decisiones.
- b) Contiene información presente y pasada, siendo determinante en la calidad de los datos que conforman la base.
- c) Establecer y conservar medidas e indicadores claves de rendimiento (KPI, Key Performance Indicator), e indicadores reales de metas (KBI Key Goal Indicator) decisivos para una organización.
- d) Contribuir con información general y desagregada según los requerimientos.
- e) Aumentar la interacción entre el departamento de TIC, y los diferentes departamentos que conforman la empresa.

- f) Desarrollar un mejor entendimiento de los sistemas de información por parte de todo el personal de la empresa
- g) Mejora de la competitividad de la organización como resultado de ser capaces de
- h) Ser competente y efectivo para:
 - 1. Diferenciar lo trascendental de lo general
 - 2. Velocidad de acceso a la información
 - 3. Ser más eficaces en la toma de decisiones.

3.1.5.2 Estrategia para la construcción de un BI

Hacer un proyecto de BI para una empresa no es un proceso fácil de realizar. Hacerlo involucra que, la necesidad exista y se tenga una metodología adecuada para realizar la inteligencia de negocios, dentro de las principales estrategias para la construcción de un BI es conocer la información que se tiene, donde y que lo origina, esto nos permite tener un visión clara sobre que procesos debo hacer y que voy a obtener. En el siguiente gráfico se describe como los datos pasan a ser funcionales para el BI.

Figura 4: Origen de datos en un Sistema BI.



Fuente: Nettleton Hill, 2003

3.1.6 Software Libre

Software libre es determinar una libertad a los usuarios para poder realizar acciones como: ejecutar, copiar, distribuir, estudiar, modificar y realizar contribuciones al software. [9]

Es un programa o software de computación que permite tener ciertas libertades como es su licencia, código fuente y que los usuarios a más de usarlo puedan publicar mejoras al software:

Se considera como software libre al paquete informático que permite determinadas libertades:

- Libertad de utilizar el programa según la necesidad existente.
- Libertad para estudiar cómo fue realizado el programa y modificarlo según la perspectiva de mi proyecto (para esta acción se debe tener el código fuente)
- Tener libertad para distribuir copias del software a diferentes usuarios
- Libertad para hacer cambios al paquete y compartirlo con todos los usuarios, contribuyendo a todos aquellos que utilizan software libre. [10]

En el gráfico se detalla de mejor forma las libertades que tiene un software libre.

Figura 5: Libertades del Software Libre



Fuente: Nettleton Hill, 2003

3.1.7 Pentaho

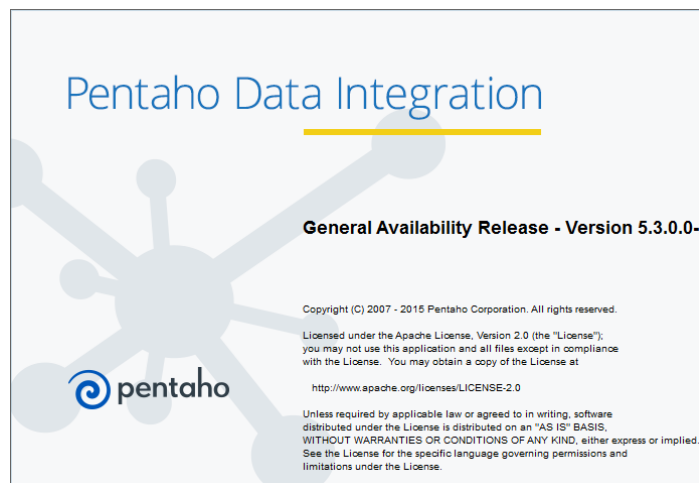
Pentaho pertenece a Pentaho Corp., fundada por el año 2004, sus dueños fueron pioneros en Business Intelligence Open Source de tipo comercial. Pentaho tiene una gama de soluciones BI para diferentes áreas. Las soluciones y su implantación están basados en Java. Pentaho tiene dos versiones: una versión comunitaria gratuita que se orienta específicamente al mundo académico, y una versión de tipo Enterprise comercial que se orienta a empresas del sector público y privado. [11]

Se puede definir a Pentaho como una herramienta que ayuda a construir un BI bajo el contexto de la utilización de software libre y que a nivel empresarial evitara tener gastos innecesarios. En si Pentaho tiene una plataforma flexible que lo permite interactuar con varios software indispensables para la elaboración de un BI. Entre sus diferentes soluciones incluye el análisis multidimensional (OLAP; Procesamiento analítico en línea), construcción de informes, cuadros de mando integral y minería de datos, que son necesarios para la gestión y análisis de información.

Pentaho se ha desarrollado bajo programación en Java al igual que sus componentes y librerías que tienen características y funcionalidades iguales a un software comercial, por lo cual se puede decir que Pentaho cumple con las expectativas de dar solución a requerimientos empresariales. [12]

El siguiente gráfico nos detalla la versión y el logo de Pentaho.

Figura 6: Logotipo de Pentaho (PDI).



FUENTE: <http://pentaho.almacen-datos.com/>

3.1.7.1 La plataforma Pentaho Open Source Business Intelligence

Pentaho por tener características de Software Libre tiene una plataforma totalmente estructurada en Java lo que lo hace ver dúctil, sus herramientas se han enfocado principalmente al Análisis de Datos. De esta forma caracterizamos a Pentaho como un software que tiene una plataforma pensada en el mundo empresarial.

3.1.7.2 Módulos de Pentaho

Pentaho B.I tiene los siguientes módulos:

- a) Reporting – Es un módulo adecuado para la presentación de reportes, estructurado de acuerdo a las necesidades empresariales, este módulo permite obtener reportes rápidos y de gran volumen, sus resultados se representan en una gran variedad de formatos que dan la posibilidad de ser impresos y a su vez exportarlo en formato PDF, XLS, HTML y texto. Pentaho Reporting da la posibilidad de programar de forma automática reportes que sean necesarios obtener periódicamente.
- b) Pentaho Análisis.- Este módulo permite realizar análisis avanzado de manera que la información sea explotada al máximo por parte de los usuarios. Utiliza Tablas Píbot y Crosstabs, generando la factibilidad de hacer una navegación por los datos y observar falencias, puede utilizar filtros para agregar o desagregar datos para el análisis. Estos datos pueden visualizarse en SVG o Flash, Dashboards Widget, pero también tienen la consistencia que se pueden integrar a través de Minería de Datos y portales Web. Tiene flexibilidad para analizar Datos Dinámicos utilizando Microsoft Excel mediante la conexión. OLAP de Mondrian.
- c) Dashboards.- Este módulo involucra la utilización de Pentaho Reporting y Pentaho Analytics. que visualiza informes, gráficos y análisis OLAP. Pentaho Dashboards concentra una diversidad de gráficos, tablas, velocímetros e integrarlos con los Portlets JSP.

- d) Data Mining.- Este módulo permite realizar minería de datos utilizando la herramienta Weka.
- e) Integración de Datos.- Este módulo permite realizar la concentración de información utilizando Kettle ETL en Pentaho Data Integration, el cual realiza procesos de extracción, transformación y carga, lo cual lo hace muy atractivo para la utilización de herramientas comerciales.

El siguiente gráfico resume los módulos que tiene Pentaho.

Figura 7: Módulos de Pentaho



FUENTE: <http://repositorio.utn.edu.ec/bitstream/123456789/2660/1/04%20ISC%20279%20TESIS.pdf>

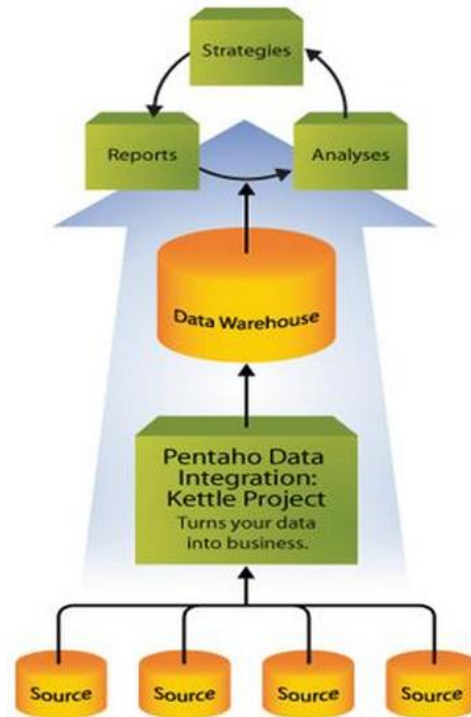
3.1.7.3 Data Warehouse

Un almacén de datos mayormente conocido Data Warehouse en idioma inglés es un conjunto de datos orientado a temas, integrado, no volátil, estable y que se usa para el análisis y toma de decisiones. [8]

Se considera al Data Warehouse como una reunión de datos dirigidos hacia un aspecto de integración fija y que sufre una variabilidad en el tiempo, construidos bajo una estructura de soportar cualquier tipo de requerimiento industrial. [13]

El gráfico que está a continuación nos muestra un esquema de un DW.

Figura 8: Esquema de un Data Warehouse



FUENTE: [http:// pentaho.almacen-datos.com/](http://pentaho.almacen-datos.com/)

3.1.7.4 Características de un Data Warehouse

Las características de un Data Warehouse son:

- a) Ordenar y organizar un conjunto de datos procedentes de diversas fuentes (transaccionales, externas, otras bases de datos).
- b) Los datos presentan cierta estabilidad en el tiempo
- c) Los datos organizados suelen tener relación entre ellos, relaciones causales o de causa efecto (bases de datos relacionales). En ocasiones se estructuran por dimensiones temáticas con más jerarquías (cubos multidimensionales).

- d) En el Data Warehouse no existe volatilidad de los datos son copia transaccional de los datos originales
- e) Permiten tener datos históricos, que puede tener una visión temporal de un hecho, es decir permite consolidar datos actuales y pasados.
- f) Sus fines son específicos y temáticos (análisis de clientes, productos, mercado, sucursales, vendedores, etc.)
- g) El enfoque: los datos son organizados con el fin de obtener información para la toma de decisiones en la empresa u organización a distintos niveles jerárquicos. Por lo tanto se orienta a proporcionar información a los analistas y directores que tiene la empresa, incluso personas que no pertenezcan a la empresa pero requieren de información.

3.1.7.5 PostgreSQL

Es un sistema de Gestión de Base de Datos relacional orientado a objetos que es libre. Entre sus características son de destacar lo siguiente: soporta transacciones, incluye herencia entre tablas, destaca por soportar tipos de datos a parte de los tipos base, como monetarios, elementos gráficos, cadenas de bits, etc.; permite una alta concurrencia ya que mientras un proceso escribe en una tabla, otros pueden acceder a la misma tabla sin necesidad de bloqueos; y está diseñado para su empleo en ambientes con grandes volúmenes de datos. [14]

Se le considera a PostgreSQL como un sistema de gestión de bases de datos objeto - relacional que se orienta a objetos, está publicado con licencia BSD (Distribución de Software Berkeley). Como es libre su código es abierto y puede ser utilizado por toda la comunidad.

3.1.8 Muestreo

Se conoce como muestreo a técnicas que nos permiten establecer un subconjunto de elementos que forman parte de a una población en estudio.

3.1.8.1 Tipos de Muestreo

3.1.8.1.1 Muestreo Aleatorio

Se conoce como Muestreo Aleatorio al proceso mediante el cual se extraen a elementos de una población que van a formar parte de la muestra a través de una selección aleatoria, es decir donde todos los elementos tienen la misma oportunidad de salir. [15]

3.1.8.1.2 Muestreo Estratificado

El muestreo estratificado es aquel método que se basa en escoger a sus elementos según los diferentes estratos que existen en una población, es decir según los niveles que pueden caracterizar a una población en estudio.

3.1.8.1.3 Muestreo por Conglomerados

El muestreo por Conglomerados se caracteriza por escoger a los elementos que van a formar parte de mi muestra, en lugares o sitios donde existe conglomerado humano y que van a cumplir con características que van a dar objetividad a mi investigación.

3.1.8.1.4 Muestreo Mixto

El muestreo Mixto es una consecuencia o una combinación de los muestreos antes mencionados, y que según el tipo de investigación permitirá asociar más de un tipo.

3.1.8.1.5 Software Estadístico

Son paquetes informáticos que están constituidos en forma específica para realizar de forma rápida y eficiente cálculos de Estadística Descriptiva e Inferencial.

3.1.8.2 Statistical Package for the Social Sciences (SPSS)

El software estadístico SPSS es uno de los programas más utilizados para realizar cálculos estadísticos, permite administrar bancos de datos de manera eficiente, realizar proyecciones y tendencias de manera sencilla, cuenta con un interfaz adaptable a usuarios que no necesariamente sean expertos en estadística, lo que facilita de gran manera la obtención de estadísticas descriptivas e inferencias estadísticas.

3.2. Estado del Arte

En la Universidad Técnica de Cotopaxi existe un trabajo de investigación denominado "Implementación de un Data Warehouse, para el departamento de desarrollo integral del Municipio de Latacunga" realizado por Verónica Cuchiparte. Su principal conclusión es que un DW es un conjunto de bases, dicho propósito es obtener información prioritaria según los requerimientos departamentales de la empresa. Esta investigación ayudo a concentrar la información en un Data Warehouse y poner a disposición de los diferentes departamentos del Municipio y puedan utilizar según su necesidad. Esta investigación se relaciona y ayuda a tener conceptos definidos sobre la construcción de un Data Warehouse que es una parte que se realiza en la construcción del BI.

En la Escuela Superior Politécnica de Chimborazo (ESPOCH) existe una tesis cuyo tema es "Estudio de la tecnología Business Intelligence y su aplicación en un modelo de sistema de información gerencial en Petroproducción." realizado por Erika Paola Merino Logroño y Hugo Giovanni Vera Flores. Su principal conclusión es que todas las empresas en su mayoría poseen procesos automatizados y sistemas que ayudan en el día a día a sus empleados, por lo cual se puede aseverar que BI es una de las mejores soluciones para empresas del sector público y privado, en cuanto a integración de datos, análisis, consultas y explotación de los mismos. Esta investigación se relaciona en definiciones y metodologías a seguir para la construcción de un BI, además que realiza fundamentaciones para la consolidación de información, que es muy escaso en instituciones públicas y es el propósito de esta investigación.

La integración, entre soluciones tecnológicas, recurso humano y procesos de negocio en un sistema integrado e interrelacionado ha permitido tomar decisiones a partir de la administración de la información, la investigación aporta conocimientos técnicos sobre el diseño de un BI y la administración de instituciones públicas de salud.

La aplicación de un BI en las instituciones públicas de salud de Colombia (IPS) ha permitido tener una cultura organizacional de información, donde la generación de las mismas permite obtener ventajas competitivas para un mejor servicio de salud. [16]

Los modelos estadísticos de clasificación se emplean cotidianamente en salud como sistemas de evaluación que permiten construir grupos homogéneos en la ayuda de pronósticos, diagnósticos, elección de terapias y cuantas situaciones requieran la discretización como herramienta para la toma de decisiones más acertadas. Su utilidad no solo se evidencia en el ámbito médico sino que muchos de ellos ayudan a mejorar actuaciones y políticas en el sector de la salud pública. Como parte de las tecnologías sanitarias, requieren de una evaluación continua en los diferentes escenarios donde se ejecuten. Por otra parte, emplearlos de manera mecánica, sin una mentalidad crítica, puede traer más riesgos que beneficios, por lo que su poder predictivo o explicativo no justifica que sus resultados se empleen de manera incuestionada. Esta investigación realizada en Colombia para el sector de la Salud Pública de ese país, se relaciona directamente con esta, pues permite considerar parámetros o indicadores que ayudan de gran manera a la toma de decisiones y evaluaciones, que servirá de gran medida en la nueva estructura funcional de salud que tiene Ecuador.

En Cuba se realizó una investigación que consiste en concebir y diseñar una solución matemático computacional para la contextualización de la estrategia sanológica en diferentes escenarios con vistas a la promoción de salud en términos de la incorporación paulatina de herramientas inteligentes, así como implementar un prototipo como una primera aproximación que permita almacenar, consultar y realizar análisis sobre los datos individuales obtenidos a partir de la aplicación del enfoque sanológico con vistas a la toma de decisiones consensuadas.

En el análisis del problema a abordar y su contexto ha resultado crucial el intercambio abierto entre los miembros del equipo sanológico, especialmente, entre los desarrolladores y los sanólogos. El interés informacional se centra en la determinación de las áreas de sanoacción a abordar preferentemente en la ruta de salud de una persona. Otras interrogantes que requieren un mayor grado de elaboración corresponden a discernir las áreas de sanoacción más frecuentemente detectadas en un consultorio en el último mes; encontrar patrones de correlación entre el comportamiento de los miembros de una familia, extrayendo similitudes y diferencias; reconocer las tendencias en la promoción de salud en un territorio.

Durante la etapa de planificación se constató la necesidad de aplicar la estrategia sanológica como punto de partida para contar con datos almacenados que faciliten no solo la interrelación directa del

sanólogo y la persona en función de la ruta de salud individual sino también que favorezcan la obtención de información para el accionar colectivo e, incluso, que en el transcurso del tiempo posibiliten investigaciones en comunidades o territorios que propicien decisiones de amplio alcance. [17]

En Cuba se utilizó información individual para la construcción de un BI que ayudó a obtener indicadores sanológicos, esta investigación se relaciona porque sigue el propósito de obtener indicadores para tomar decisiones y entender que metodología seguir en la construcción de un BI.

Un estudio realizado en Europa por Information Builders Ibéric mostró el costo que tiene la falta de sistemas de toma de decisiones en las organizaciones, según estos datos, el empleado europeo medio pierde una media de 67 minutos diariamente buscando información de la compañía, lo que equivale a un 15,9% de su jornada laboral. Para una organización de 1.000 empleados que gane unos 50.000 euros al día esto equivale a 7,95 millones de euros al año de salario perdido, todo ello por la búsqueda de información para tomar una decisión. [18].

El estudio realizado en Europa sirve para dar estadísticas del tiempo y recursos que se pierden al momento de buscar información no centralizada, lo que permite enrumbarse en la construcción de un BI para sectores sociales. Esta investigación se relaciona con la planteada, porque permite ser claros en la utilidad que va resultar su implementación en el sector público, ya que el principal enfoque es tomar decisiones importantes para el sector salud.

Capítulo 4

Metodología

4.1 Diagnóstico

Se limitó a conocer causas por las cuales era necesaria la construcción de un BI, se conoce que la información generada por sus diferentes programas de atención y concentrados en la Coordinación Zonal de Salud no se lo hace en una sola base, lo que dificulta al momento de requerir información de diferentes programas, además no se tenía definida una base de indicadores que permitan tomar decisiones, ante estas necesidades se hace imprescindible la utilización de un BI el cual será alimentado con información extraída de los tres niveles de atención. A la información extraída se realizó un análisis estadístico descriptivo con el propósito de tener una panorámica clara sobre la información extraída, por la cantidad de información que se tiene se determinó una muestra para realizar el análisis.

Se conoce que dentro de las políticas del actual del Gobierno Ecuatoriano, es la implementación y utilización de software libre en todas las instituciones y empresas del sector público, por esta razón se decide la utilización de un software que sea potencialmente eficiente y tenga las características de libre como es el caso de Pentaho que posee dentro de las características más importantes un potente motor para realizar la extracción, transformación y carga de archivos y que abarca en una sola plataforma todas las soluciones para la construcción de un BI, características que son semejantes a la aplicabilidad en el sector de la Salud.

4.2 Método aplicado

La metodología que se utilizó en el desarrollo del aplicativo es la de la de Hefestos por las características de los datos de Salud que se tiene, las cuales se ajustan a las necesidades de nuestra investigación.

La metodología HEFESTO puede resumirse mediante el siguiente gráfico:

Figura 9: Esquema de la Metodología Hefesto.



Fuente: <http://www.businessintelligence.info/docs/hefesto-v2.pdf>

Paso 1 Análisis de Requerimientos

Se realizó con los usuarios, a través de preguntas que determinaron la identificación de indicadores y perspectivas que serán tomadas en cuenta para la construcción del DW.

a. Identificar Preguntas

Mediante preguntas realizadas al personal de salud, determinando cuales son las necesidades que ellos sobre el requerimiento de información, estas son:

Cuántos partos se atendieron en la Coordinación Zonal de Salud en un periodo de tiempo determinado

Conocer el promedio de días de estadía que tienen los pacientes hospitalizados en un periodo de tiempo

Cuál es el porcentaje de ocupación que tienen las unidades hospitalarias en un periodo de tiempo

Cuántos pacientes son atendidos diariamente en las unidades hospitalarias en un periodo de tiempo.

Cuántos egresos se realizan diariamente en una unidad hospitalaria en un periodo de tiempo determinado

b. Indicadores y Perspectivas

Luego de haber trabajado sobre la necesidad de información que tenían los usuarios, se procedió a determinar los indicadores de las preguntas extraídas anteriormente, así como las perspectivas que intervendrán, según esto se tiene:

Total de partos que se atendieron en las Provincias de la CZS en un periodo de tiempo.

Promedio días de estadía de pacientes en las Provincias de CZS en un periodo de tiempo.

Porcentaje de ocupación de las unidades hospitalarias de las provincias de la CZS en un periodo de tiempo.

Número de pacientes atendidos diariamente en unidad hospitalaria de la CZS en un periodo de tiempo.

Total de egresos realizados por una unidad hospitalaria de la CZS en un periodo de tiempo.

Indicadores

- Total de Partos
- Promedio días de estadía
- Porcentaje de ocupación
- Diario días paciente
- Total de egresos

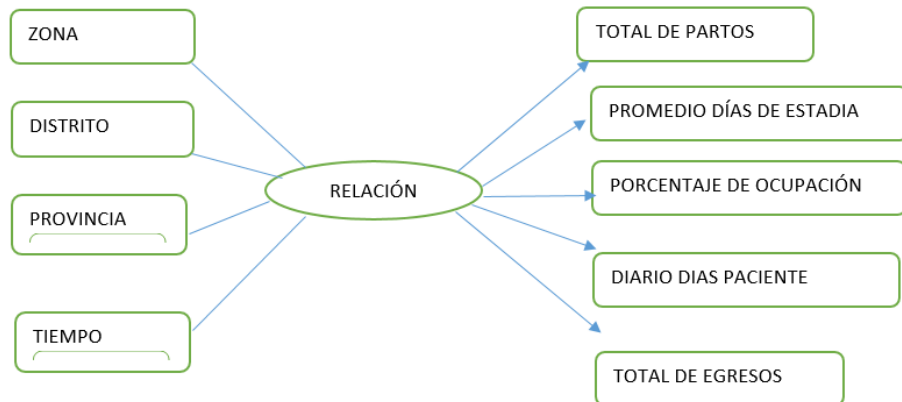
Perspectivas

- Zona
- Distrito
- Circuito
- Provincia
- Tiempo

c. Modelo Conceptual

Se construyó a través de los indicadores y perspectivas establecidos en el paso anterior. El siguiente gráfico visualiza el Modelo Conceptual

Figura 10: Esquema del Modelo Conceptual



Paso 2 Análisis de los OLTP

a. Conformar Indicadores

En esta fase se determinó, las funciones que se van a utilizar para calcular cada uno de indicadores

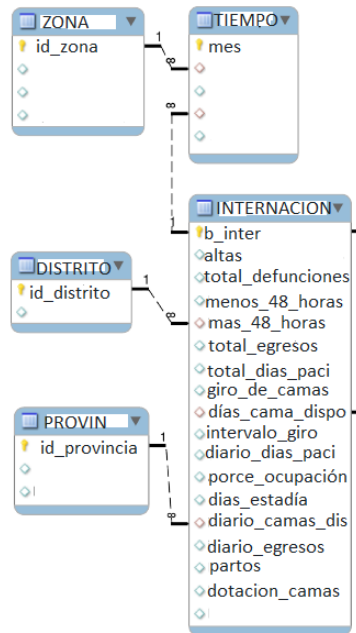
- Total de Partos
Hechos: Total de Partos realizados
Función de Sumarización: SUM
- Promedio días de estadía
Hechos: Promedio de días de estadía
Función de Promedio: AVG
- Porcentaje de Ocupación
Hechos: Porcentaje de ocupación
Función de Porcentaje: SUM
- Diario días Paciente
Hechos: Total días paciente
Función de Sumarización: SUM
- Total de egresos
Hechos: Total de egresos
Función de Sumarización: SUM

b. Establecer Correspondencias

En esta fase se identificó la correspondencia existente entre el modelo conceptual y las fuentes de datos. El siguiente grafico determina la fase de entidad de relación existente.

La tabla de Internación se relaciona con Zona, Distrito, Provincia y Tiempo.

Figura 11: Diagrama Entidad de Relación



c. Nivel de Granularidad

Con respecto a Zona, los datos disponibles son:

- Total de Partos
- Promedio días de estadía
- Porcentaje de Ocupación
- Diario días paciente
- Total de Egresos

Con respecto a Distrito, los datos disponibles son:

- Total de Partos
- Promedio días de estadía
- Porcentaje de Ocupación
- Diario días paciente
- Total de Egresos
-

Con respecto a Provincia, los datos disponibles son:

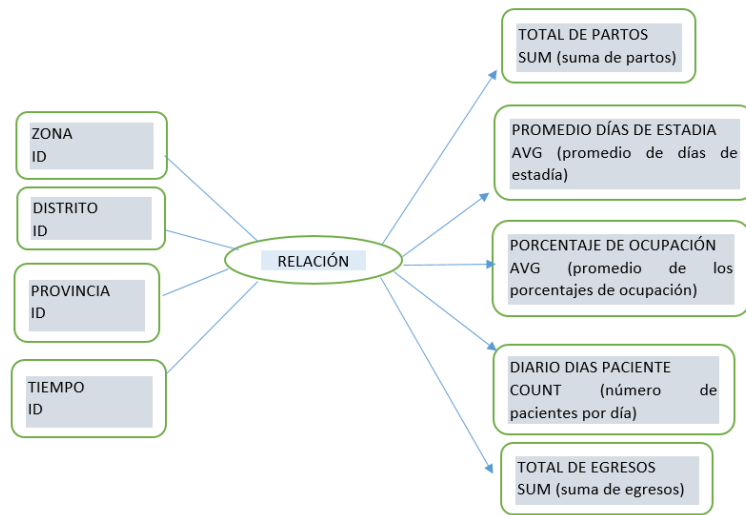
- Total de Partos

- Promedio días de estadía
- Porcentaje de Ocupación
- Diario días paciente
- Total de Egresos

d. Modelo Conceptual Ampliado

En esta fase se realizó el modelo conceptual, colocando bajo cada perspectiva los campos seleccionados y bajo cada indicador su respectiva fórmula de cálculo. La gráfica siguiente muestra de mejor forma este modelo.

Figura 12: Modelo Conceptual Ampliado



Pasó 3 Modelo Lógico del DW

Se realiza el modelo Lógico:

a. Tipo de Modelo Lógico del DW

Según las características del proyecto se utiliza el esquema del modelo estrella.

b. Tablas de dimensiones

Se diseñaron las tablas de dimensiones que van a ser parte del DW

- Perspectiva Zona

La nueva tabla dimensión tendrá el nombre "ZONA"

Clave principal "idzona"

El campo será "Zona"

El siguiente gráfico nos muestra la "Dimensión de Zona"

Figura 13: Tabla Dimensión "ZONA"



- Perspectiva Distrito

La nueva tabla dimensión tendrá el nombre "DISTRITO"

Clave principal "iddistrito"

El campo será "distrito"

El siguiente gráfico nos muestra la "Dimensión de Distrito"

Figura 14: Tabla Dimensión "DISTRITO"



- Perspectiva Provincia

La nueva tabla dimensión tendrá el nombre "PROVINCIA"

Clave principal "idprovincia"

El campo será "provincia"

El siguiente gráfico nos muestra la "Dimensión de Provincia"

Figura 15: Tabla Dimensión “PROVINCIA”.



- Perspectiva Tiempo

La nueva tabla dimensión tendrá el nombre “TIEMPO”

Clave principal “idtiempo”

El campo será “tiempo”.

El siguiente gráfico nos muestra la “Dimensión de Tiempo”

Figura 16: Tabla Dimensión “TIEMPO”



c. Tablas de Hechos

En esta parte se construyen las tablas de hechos, en el siguiente gráfico se detalla la tabla de hechos.

Figura 17: Tablas de Hechos

Total de Partos
idzona
iddistrito
idprovincia

Promedio días de Estadía
idzona
iddistrito
idprovincia

Porcentaje de Ocupación
idzona
iddistrito
Idprovincia

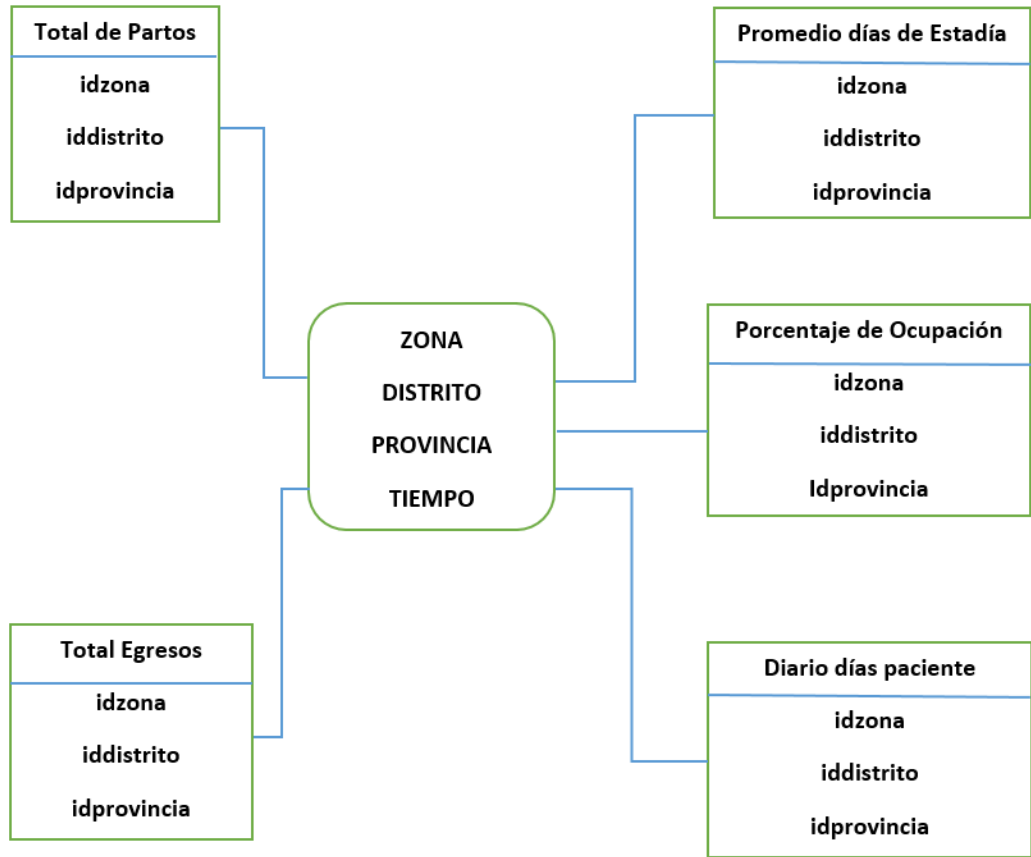
Diario días paciente
idzona
iddistrito
idprovincia

Total Egresos
idzona
iddistrito
idprovincia

d. Uniones

En esta fase se realiza la unión entre las tablas de Dimensiones y de Hechos. En el gráfico se detalla la unión de las tablas.

Figura 18: Caso uniones



Paso 4 Integración de Datos

a. Carga Inicial

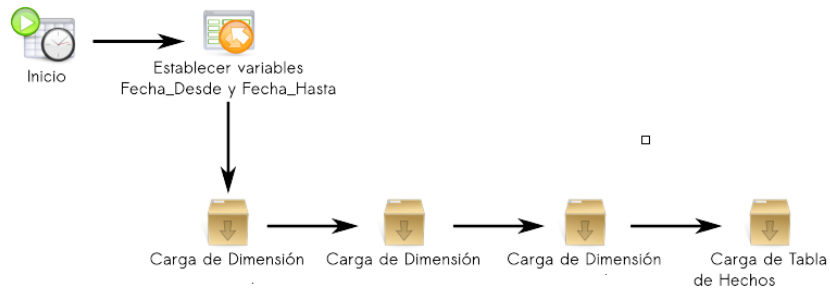
Descripción del proceso ETL (Extracción, transformación y carga), establecido para la carga inicial.

Inicio: inicia la ejecución de los pasos en el momento en que se le indique.

- Establecer variable Fecha: que serán utilizadas posteriormente por algunos pasos.
- Carga de Dimensiones: partos, días de estadía, ocupación, diario días paciente y egresos, ejecuta el contenedor de pasos que cargará la dimensión
- Carga de Tabla de Hechos: Zona

En el gráfico se visualiza la carga inicial realizada.

Figura 19: Carga Inicial



b. Actualización

En esta fase se especifica cómo se actualiza la información del DW:

La información se refresca mensualmente

Los datos de las Dimensiones (Zona, Distrito, Provincia y Fecha) se cargan mensualmente.

Los datos de las tablas de hechos se cargan mensualmente, estos son:

Partos

Días de estadía

Ocupación

Días pacientes

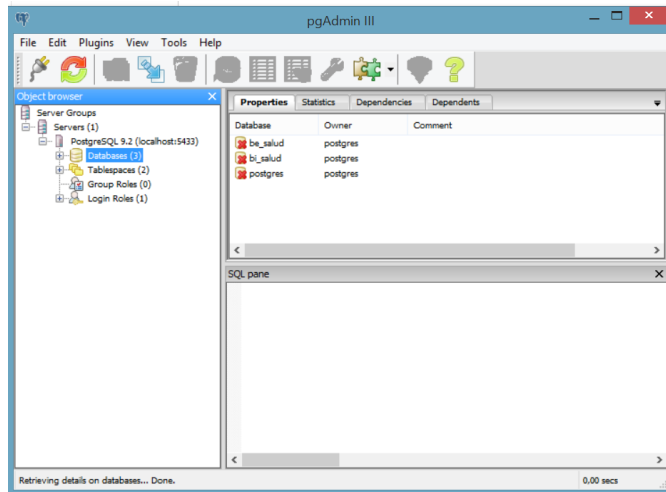
Egresos.

Creación de la Base de Datos en PostgreSQL

Posterior a la instalación de PostgreSQL que mayor comprensión esta un instructivo en Apéndice C, se procedió a la creación de la base de datos. Para lo cual ingresamos a PosgreSql, que en el momento de ingresar me va a pedir una Password que se dio en el momento de la instalación, digitamos el Password y se nos activa la ventana para poder trabajar, el grafico describe de mejor manera como nos queda la ventana luego de haber digitado el Password.

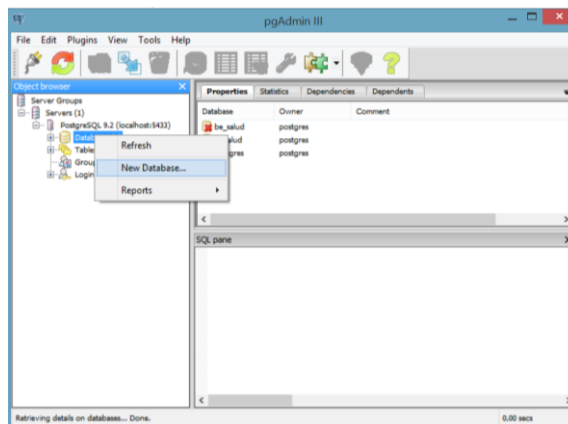
Previo a la creación de la base de datos en PostgreSQL, se debe haber instalado, la explicación de la instalación se encuentra un instructivo en el Apéndice C. Instalado el software se ingresa dando doble clic en el acceso directo, y sale una ventana de dialogo que pide un Password que va a ser el mismo que se digitó en el momento de la instalación. La siguiente gráfica muestra la ventana de PostgreSQL.

Figura 20: Ventana de PostgreSQL.



Luego proceden a la creación de la base de datos, para este proceso realizan un clic derecho sobre el icono Databases que ya viene en PostgreSQL y en el menú desplegado escogen la opción New Database, el gráfico describe de mejor forma este proceso.

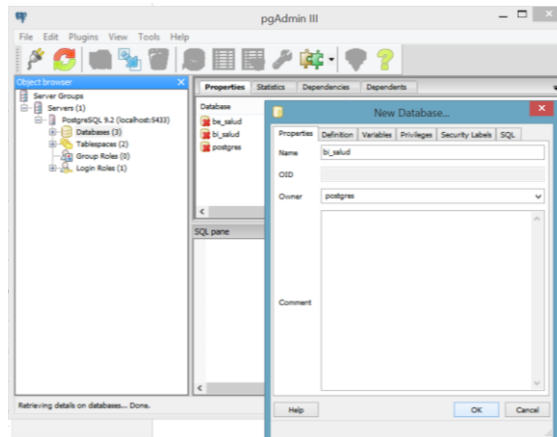
Figura 21: Creación de la Base de Datos.



Al escoger el menú New Dabase (nueva base de datos), se despliega una nueva ventana que pide ingresar el nombre de la nueva base de datos, que la han denominado bi_salud, también piden

ingresar owner (propietario) se escoge Postgres, finalmente dan OK y la base se habrá creado, en la gráfica se detalla este proceso.

Figura 22: Selección de nombre para la Base de Datos.

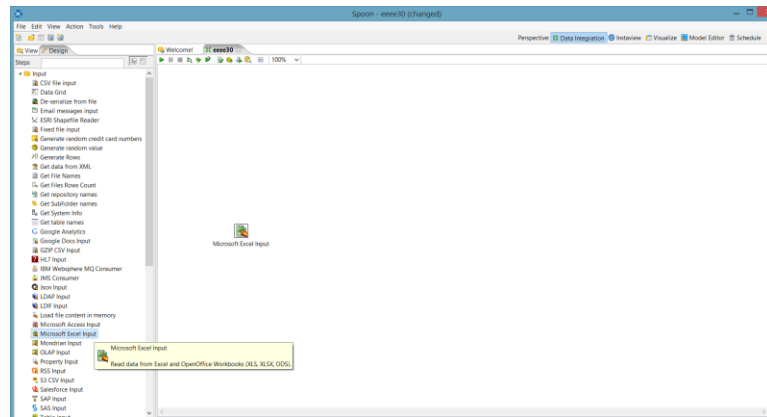


CONEXION A LA BASE DE DATOS

Luego de haber instalado Pentaho (su instalación se detalla en el Apéndice C), se procede a la conexión de la base de datos desde Pentaho con PostgreSQL, este proceso se lo hace de acuerdo a los siguientes pasos:

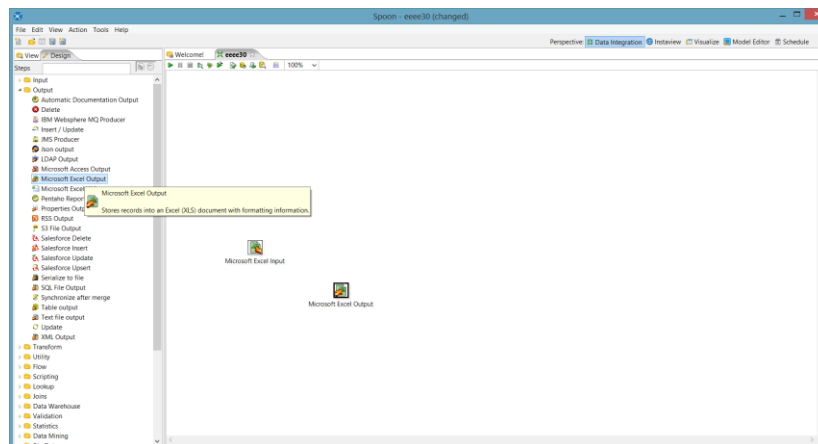
En Pentaho Data Integration (PDI) escogen los steps que nos permitan realizar la conexión, para lo cual arrastran o dan doble clic sobre el step (table input), porque van a ingresar tablas que están en formato xls, en la figura observamos este proceso.

Figura 23: Creación de las tablas de entrada.



Posterior a esto escogen el Step que va ser la salida de datos, para el caso en estudio escogen output table. La imagen muestra este proceso.

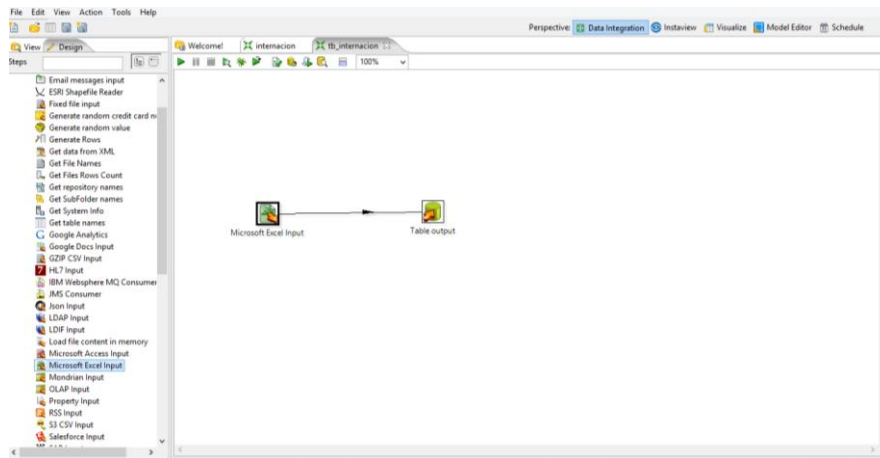
Figura 24: Creación de las Steps de entrada y salida



Elaborado por: Fabián Londo Y.

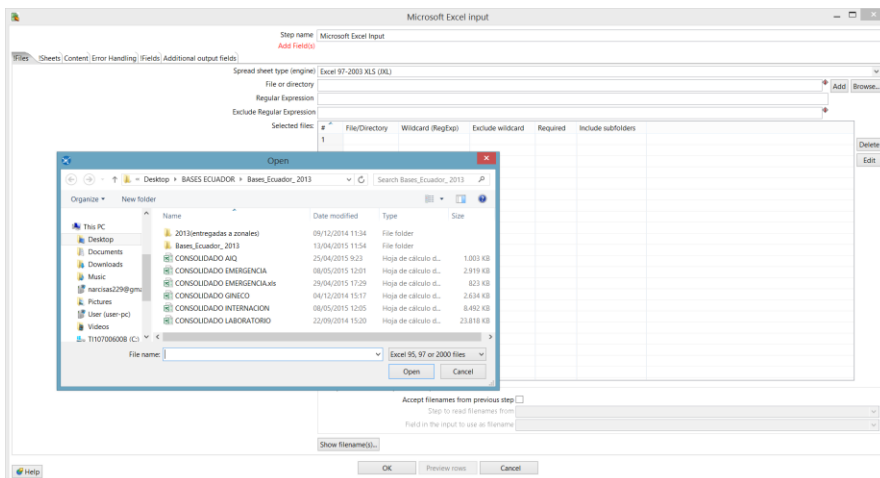
A continuación realizan la conexión entre estos dos Steps, el proceso lo realizan pulsando la tecla shift del teclado y con el mouse arrastran desde la tabla input hacia output. La imagen ilustra cómo les queda la conexión realizada.

Figura 25: Conexión de Steps



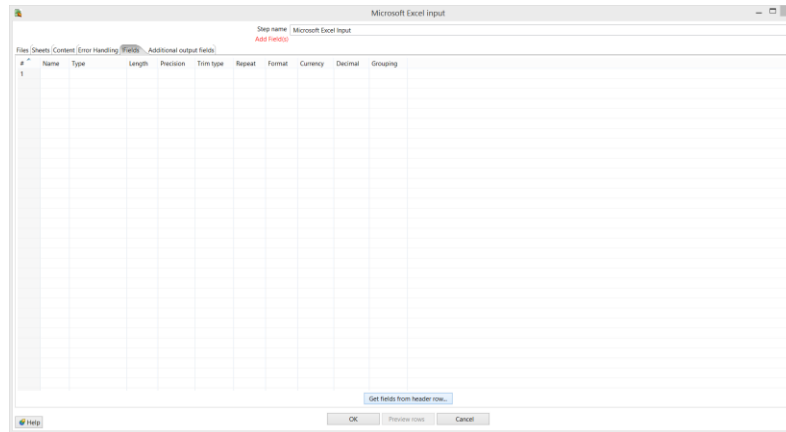
Luego caracterizamos a cada Step, para este proceso hacen doble clic sobre la tabla input y se ilustra una nueva ventana que pide el ingreso de: nombre para la tabla input, dirección del archivo de la información que se va a cargar en el Step. El gráfico visualiza estos pasos.

Figura 26: Entrada de datos al Input



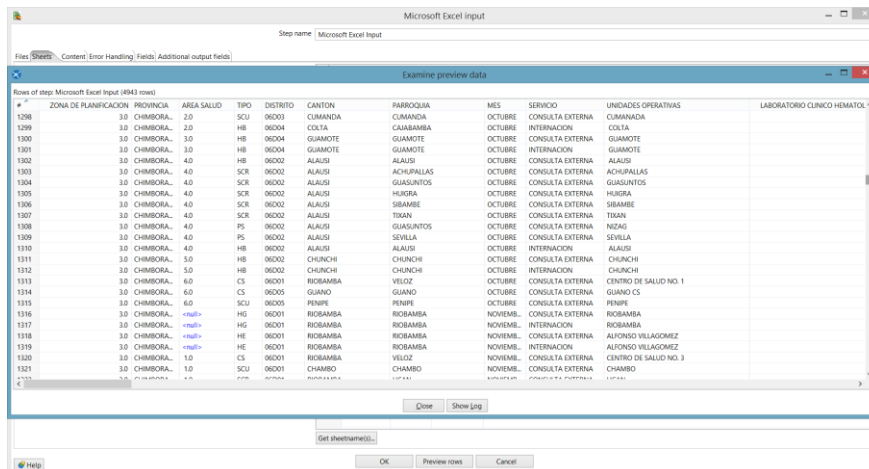
A continuación se ubican en el menú Fields de la ventana de caracterización del step input y en la opción Get fields from header row, dan clic para que se cargue la información del archivo original al step input. El siguiente gráfico visualiza el proceso.

Figura 27: Creación de Campos



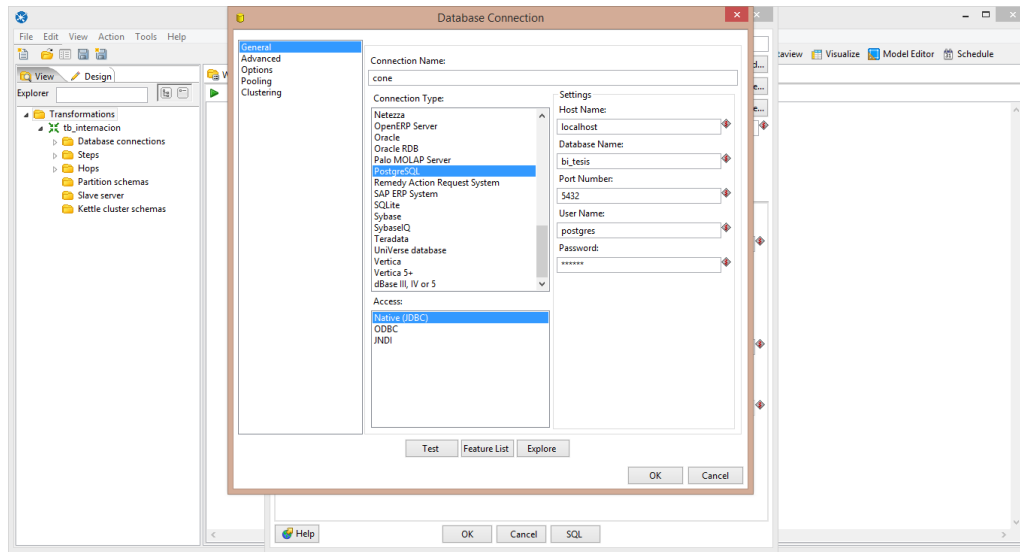
Para visualizar la carga de los datos en el step input, dan clic en preview rows, y se visualiza la información. El siguiente gráfico visualiza lo expuesto.

Figura 28: Llenado de datos en los campos creados



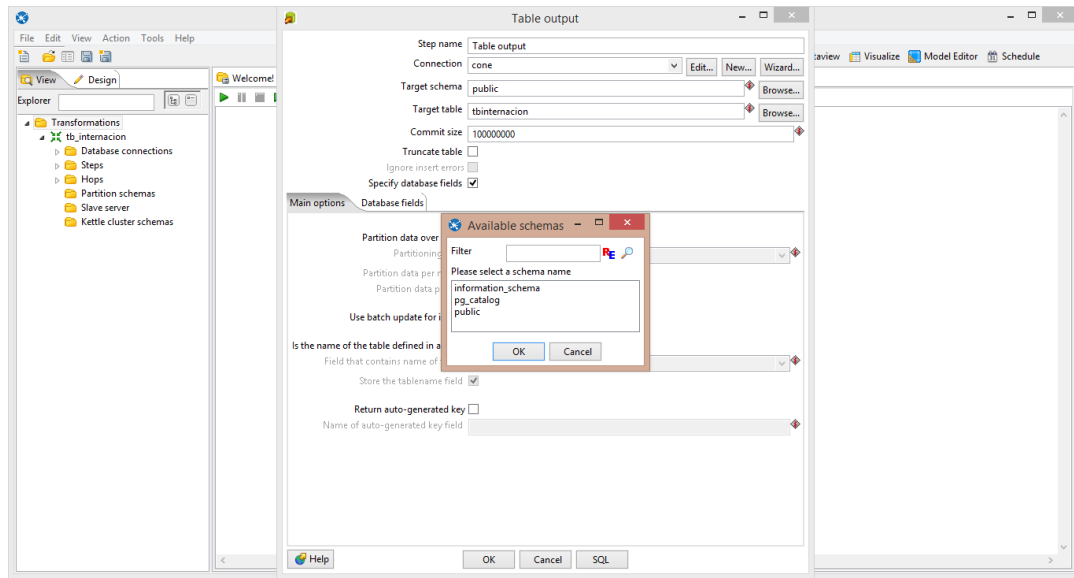
Posteriormente realizan la Conexión a la base de datos PostgreSQL desde Database Connection para generar la salida del proceso ETL (Extracción, transformación y carga). Para esto en el PDI se van a la pestaña View, dan un clic en la carpeta transformations y les sale la pestaña Database connections para conectar la transformación que se va a realizar con la base de postgresQL. La siguiente imagen detalla lo descrito.

Figura 29: Conexión a la Base de Datos



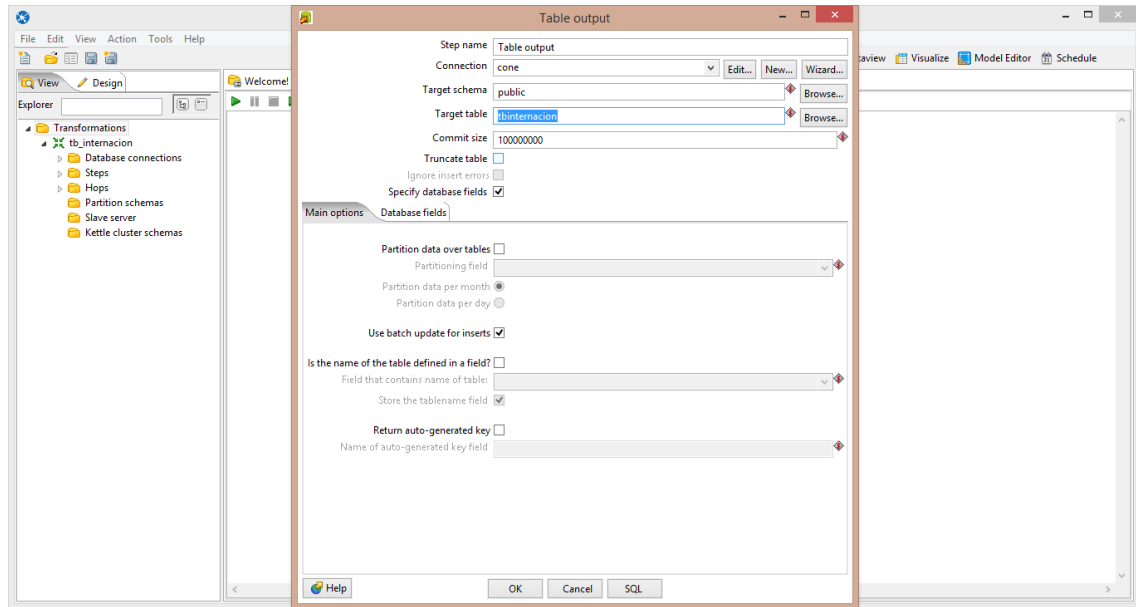
En la siguiente parte seleccionan el esquema para la base de datos, para este proyecto escogen Public. El gráfico visualiza el proceso.

Figura 30: Esquema para escoger el tipo de Base de Datos



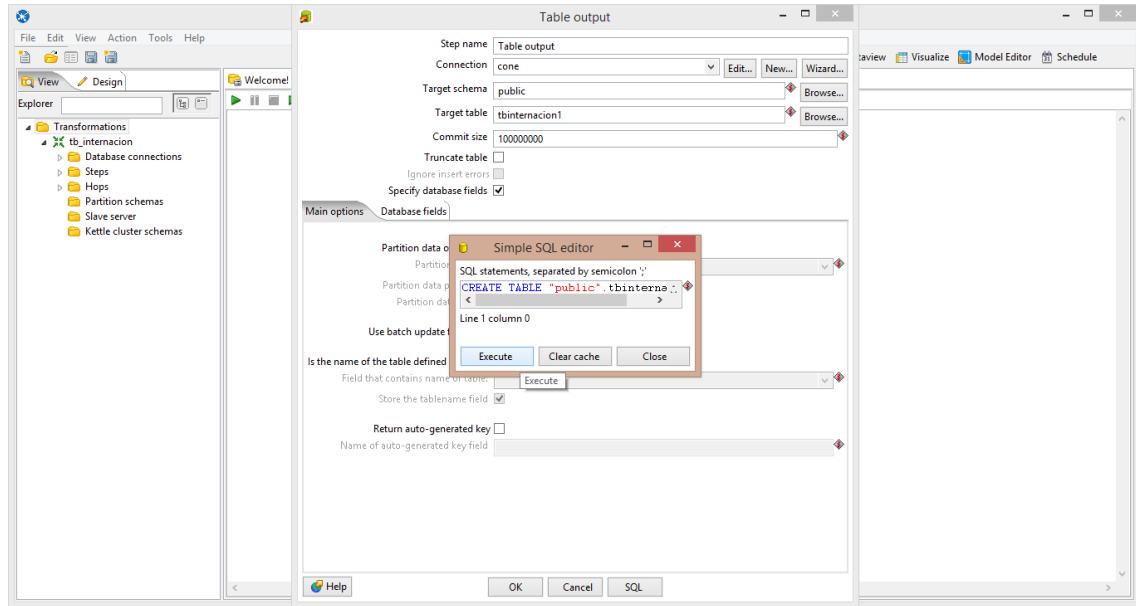
Posteriormente ingresan el nombre de la tabla que se va a crear. El siguiente gráfico nos muestra el proceso.

Figura 31: Nombre de la tabla a crear



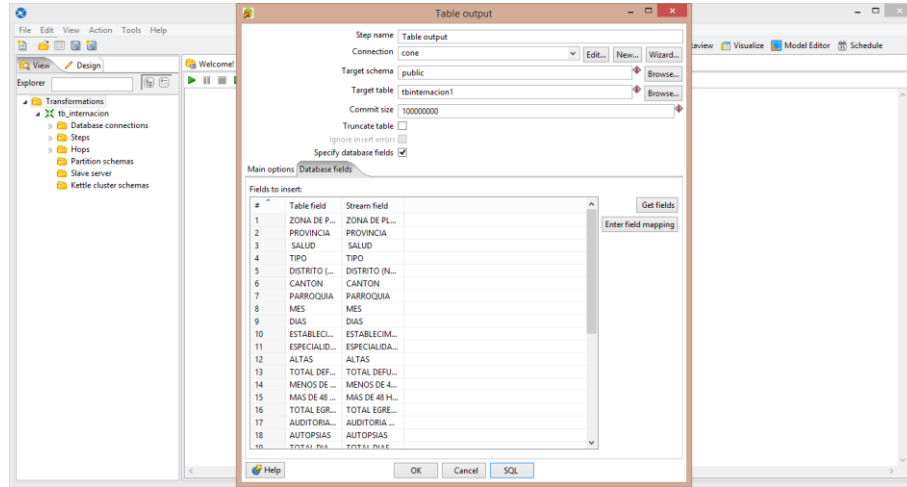
Continuando con el proceso dan un clic en la pestaña SQL, y surge una nueva ventana de dialogo en donde dan un clic en Execute para ejecutar el SQL y se me cree la tabla. El siguiente gráfico visualiza el proceso.

Figura 32: Nombre de la tabla a crear



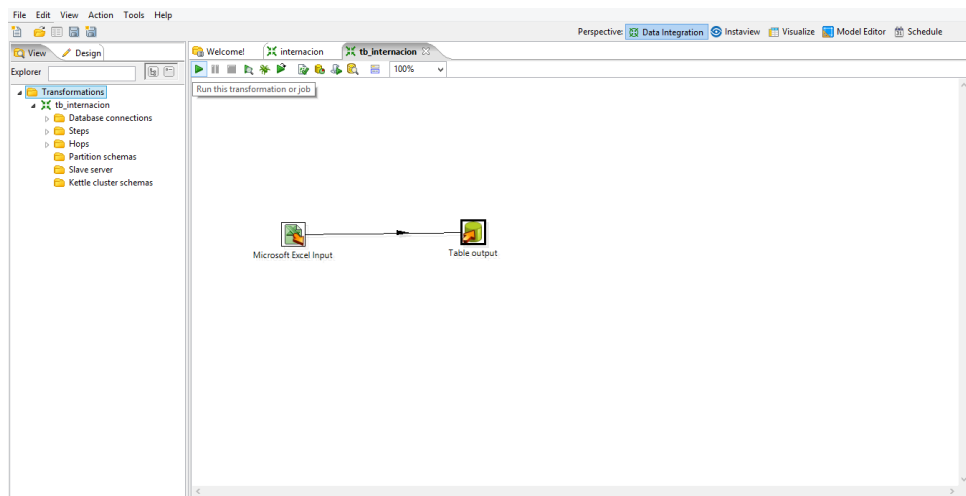
A continuación les muestra la nueva tabla generada con la sentencia SQL ejecutada, en donde se visualizan los campos de toda mi tabla generada. El gráfico les muestra el proceso realizado.

Figura 33: Tabla Generada



Finalmente ejecutan el proceso ETL generado con la opción Run. El siguiente gráfico visualiza el proceso.

Figura 34: Ejecución del ETL



4.3 Materiales y herramientas

La investigación se desarrolló, mediante una modalidad de investigación de tipo Cualitativa, debido a que la primera etapa, es un proceso de recolección y obtención de datos numéricos, mediante la aplicación de fórmulas predeterminadas, para obtener porcentajes relacionados al desarrollo de un Business Intelligence en software libre, basado en indicadores de gestión.

La investigación también es cualitativa, porque examina fenómenos con gran detalle, sin ninguna categoría o hipótesis predeterminada, este enfoque, permitirá determinar las cualidades para una coordinación de salud, que se encuentran sometidas en el objeto de estudio, con la finalidad de entender y comprender la relación que existe entre las dos variables, BI en software libre y los indicadores de gestión.

Para la validación de la información se utilizó un muestreo aleatorio estratificado por la cantidad de información existente y porque se encuentran en diferentes matrices lo cual hace imposible revisar en su totalidad, la validación se aplicó principalmente a las variables que iban hacer analizadas para obtener los indicadores, a las cuales previamente se las hizo un análisis Descriptivo utilizando el software estadístico SPSS (Statistical Package for the Social Sciences) , que va a permitir observar la existencia de datos anómalos o no comunes y verificar su autenticidad.

Para la validación y consistencia de los datos se hizo una revisión directa, es decir que paso por dos filtros de revisión, el primer filtro se lo ejecuto en cada distrito, y el segundo en la Coordinación Zonal, en el segundo filtro no se hizo la validación de todos los datos, pues la cantidad que existe de los mismos demanda gran cantidad de tiempo, razón por la cual se determinó una muestra para revisión, y el tipo de muestreo utilizado fue un muestreo Aleatorio Estratificado.

Para ver la funcionalidad del BI se realizaron pruebas en la Dirección de Planificación, proceso de Estadística, las pruebas se hicieron a profesionales del departamento antes mencionado.

Para determinar el análisis de satisfacción del BI se realizó una encuesta mediante la aplicación de un formulario a los profesionales del proceso de Estadística, la encuesta se la realizo luego de la utilización del BI.

Capítulo 5

Resultados

5.1. Producto final del Proyecto de titulación

Se realizó la recolección y consolidación de datos de las 300 unidades operativas de la Coordinación Zonal de Salud, el proceso se realizó mediante la participación de los 19 distritos que pertenecen a la Coordinación.

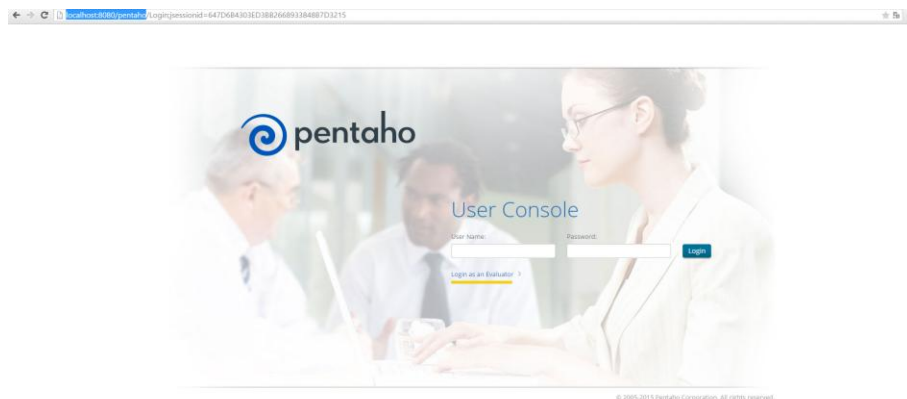
La validación de datos se realizó a través de dos filtros jerárquicos, la revisión en zona se a cien unidades operativas, por lo cual existe un 90% de confiabilidad de la información con un margen de error del 8%.

Para obtener los resultados en el proyecto se realizan los siguientes procesos:

INGRESO A LA CONSOLA DE PENTAHO

Posterior a la instalación de Pentaho, ingresan a la consola a través de la dirección localhost/8080/pentaho por medio del navegador. En la imagen se puede visualizar la acción.

Figura 35: Consola de Pentaho



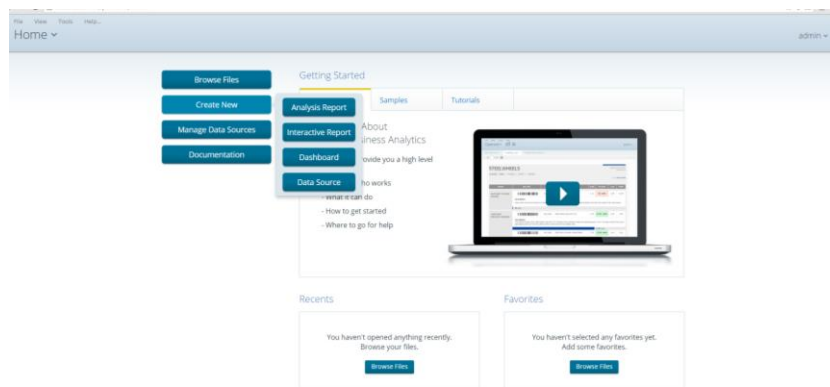
Para el ingreso lo hacen como Evaluador (Login as an Evaluator), que se encuentra en la parte inferior, sin ingresar usuario ni el password, luego hacen clic en Go. En la imagen se muestra el proceso.

Figura 36: Ingreso a Pentaho



Posteriormente les sale otra ventana en donde dan clic en Create New, que les permite escoger con que módulo de Pentaho desean trabajar. El gráfico visualiza lo comentado.

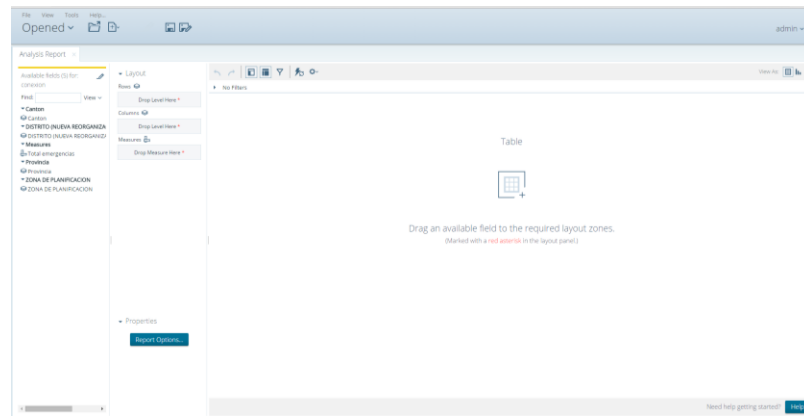
Figura 37: Menú Create New



CREACION DE REPORTES

Para la generación de reportes se ingresa al módulo de Analysis Report, al ingresar a este módulo se visualiza otra pantalla en donde se ingresa el tipo de conexión que se va a trabajar, a continuación se escogerán las métricas (medidas) que serán útiles para la generación de tablas de información que serán parte de la generación de reportes. En la siguiente figura se ilustra la pantalla donde se realiza este proceso.

Figura 38: Modulo Analysis Reporty



En la imagen que está a continuación, se determina Dimensiones y Medidas para la Generación de cubos de información. El siguiente gráfico muestra un ejemplo.

Figura 39: Elaboración del Reporte

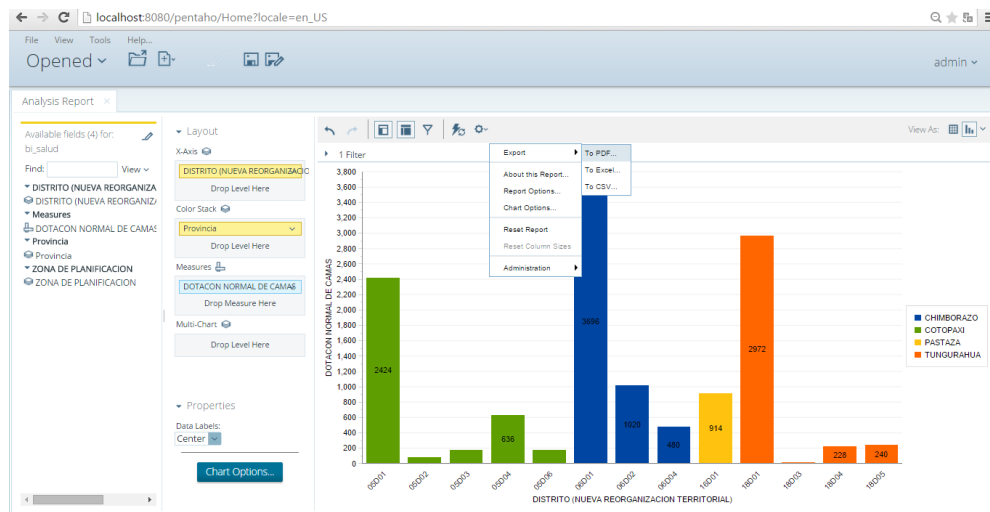
The screenshot shows the 'Analysis Report' interface with a populated data table. The table has a header row for 'ZONA DE PLANIFICACION' with sub-headers for 'Provincia' and 'Municipio'. The 'Provincia' row lists: NAPO, CHIMBORAZO, COTOPAXI, PASTAZA, TUNGURAHUA, MANABI, and SANTO DOMINGO DE LOS TSACHILAS. The 'Municipio' row lists: DISTRITO NUEVA REORGANIZACION TERRITORIAL, DOTACION NORMAL DE CABANAS, DOTACION NORMAL DE CABANAS, DOTACION NORMAL DE CABANAS, DOTACION NORMAL DE CABANAS, DOTACION NORMAL DE CABANAS, and DOTACION NORMAL DE CABANAS. The table contains data for various codes (e.g., 01001, 01002, etc.) across the municipalities. A 'Report Options...' button is visible at the bottom of the sidebar.

	ZONA DE PLANIFICACION						
	Provincia						
	2	3	3	3	4	4	20
	NAPO	CHIMBORAZO	COTOPAXI	PASTAZA	TUNGURAHUA	MANABI	SANTO DOMINGO DE LOS TSACHILAS
DISTRITO NUEVA REORGANIZACION TERRITORIAL	DOTACION NORMAL DE CABANAS	DOTACION NORMAL DE CABANAS	DOTACION NORMAL DE CABANAS	DOTACION NORMAL DE CABANAS	DOTACION NORMAL DE CABANAS	DOTACION NORMAL DE CABANAS	DOTACION NORMAL DE CABANAS
01001	-	-	-	-	-	-	-
01002	-	-	-	-	-	-	-
01003	-	-	-	-	-	-	-
01004	-	-	-	-	-	-	-
01006	-	-	-	-	-	-	-
01008	-	-	-	-	-	-	-
02001	-	-	-	-	-	-	-
02002	-	-	-	-	-	-	-
02003	-	-	-	-	-	-	-
03001	-	-	-	-	-	-	-
03002	-	-	-	-	-	-	-
03003	-	-	-	-	-	-	-
05001	-	-	3424	-	-	-	-
05002	-	-	84	-	-	-	-
05003	-	-	180	-	-	-	-

GENERACION DE TABLAS Y GRÁFICOS

Para la generación de tablas y gráficos, previamente se debe escoger las Dimensiones y Medidas que van a dar lugar a la obtención de consultas realizadas. Para realizar una tabla se ubica en la columna de campos que se encuentra en la parte izquierda y se determinan las variables a ser calculadas, luego se arrastran las variables escogidas hacia donde se encuentran las pestañas Rows (filas) y Columns (columnas). En Measure (medida) se define la variable a ser estimada o sobre la cual se va a trabajar todas las consultas, la tabla se forma automáticamente al haber identificado y trasladado variables y medidas a la ventana de dialogo. El gráfico se forma al realizar un clic en la imagen que se encuentra en la parte superior de la consola de trabajo, en donde se permite escoger el tipo de gráfico según las características de la variable a representar. El gráfico que está a continuación muestra la generación de resultados.

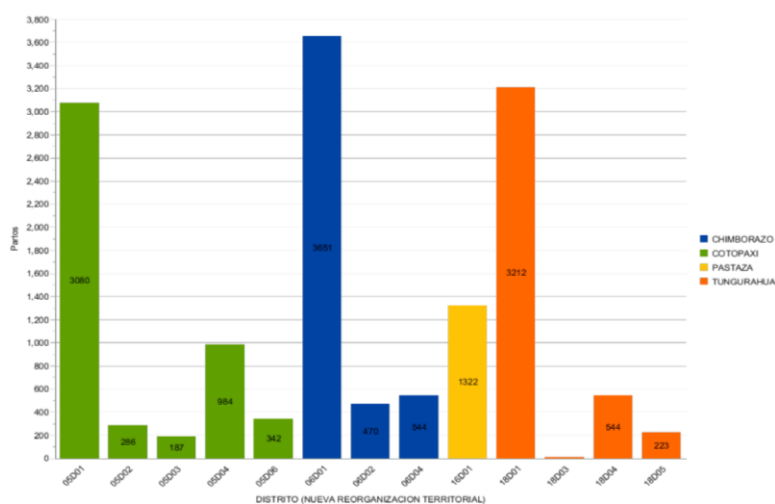
Figura 40: Reportes con Gráficos



RESULTADOS DE INDICADORES

Para dar lugar a la generación de indicadores propuestos como: nivel ocupación, dotación de camas, total de partos, total de altas, se realiza mediante la construcción de cubos de información para cada indicador, es decir primero se establece las métricas (zona, distrito y provincia), luego se determina las dimensiones que van a generar las tablas y gráficos, de esta forma se obtienen los siguientes resultados:

Figura 41: Total de partos



En el gráfico se puede observar el total de partos en las 4 provincias de la Coordinación Zonal 3 de Salud, es decir que en la Coordinación Zonal de Salud existió 14859 partos y que la mayor cantidad de partos sucedió en la provincia de Cotopaxi con 4879 partos, lo cual les permite definir si el servicio de Ginecología y Obstetricia que tiene el hospital General de Latacunga cuenta con el número adecuado de camas para atender los partos o es necesario incrementarlas.

Figura 42: Total de Altas

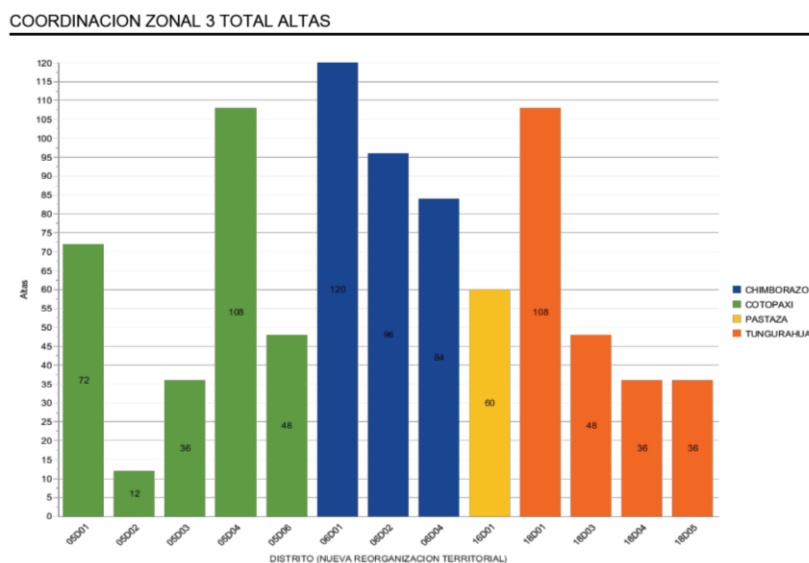


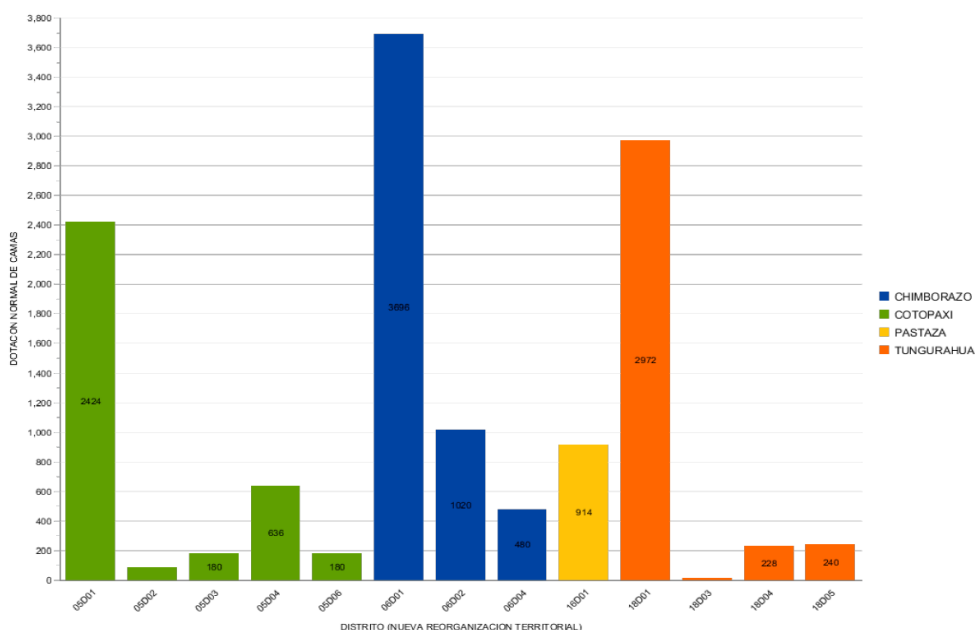
Figura 43: Total de Altas (Tabla)

COORDINACION ZONAL 3 TOTAL ALTAS

DISTRITO (NUEVA REORGANIZACION TERRITORIAL)	PROVINCIA			
	CHIMBORAZO	COTOPAXI	PASTAZA	TUNGURAHUA
	Altas	Altas	Altas	Altas
05D01	-	72	-	-
05D02	-	12	-	-
05D03	-	36	-	-
05D04	-	108	-	-
05D06	-	48	-	-
06D01	120	-	-	-
06D02	96	-	-	-
06D04	64	-	-	-
18D01	-	-	60	-
18D01	-	-	-	108
18D03	-	-	-	48
18D04	-	-	-	36
18D05	-	-	-	36
Grand Total	300	276	60	228

Analizando el total de Altas que corresponden a un mes, se observa que el total de altas es de 864 en toda la Coordinación Zonal, existiendo menos altas en la Provincia de Pastaza con 60 altas y en la Provincia de Chimborazo existe el mayor número de altas que son 300, esto les permite considerar si el número de altas realizadas es consecuente con la cantidad de insumos utilizados y con el número de médicos que tiene el hospital.

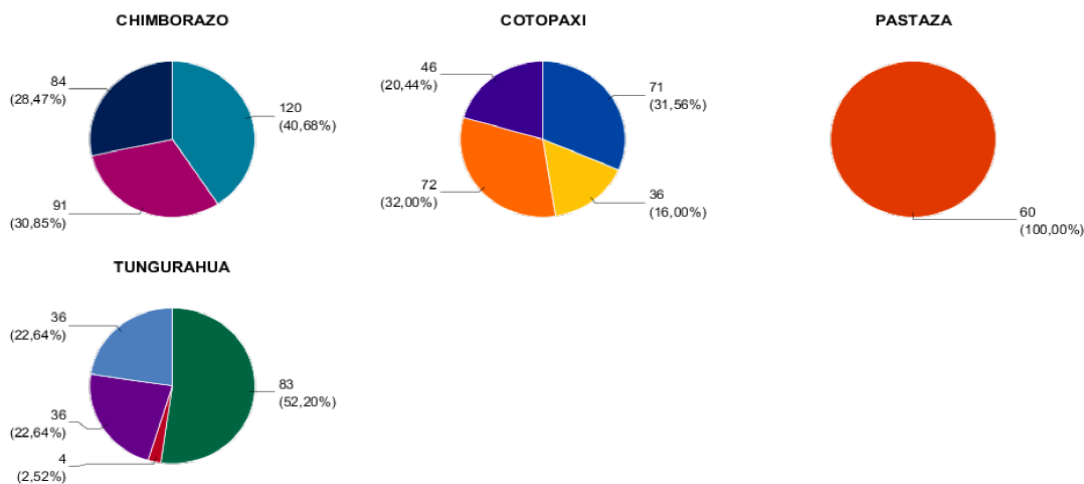
Figura 44: Dotación de Camas



En dotación de camas se puede determinar que la provincia de Tungurahua cuenta con un mayor número de camas hospitalarias a nivel de hospitales generales, su hospital general cuenta con 308 camas, en tanto que a nivel de Provincia Chimborazo es la que cuenta con un mayor número de camas

485, y la Provincia de Pastaza es la que cuenta con el menor número de camas hospitalarias por tener una sola unidad hospitalaria 60 camas hospitalarias. Conocer la dotación de camas en las unidades hospitalarias les permite identificar si el número de camas satisface la demanda de los usuarios.

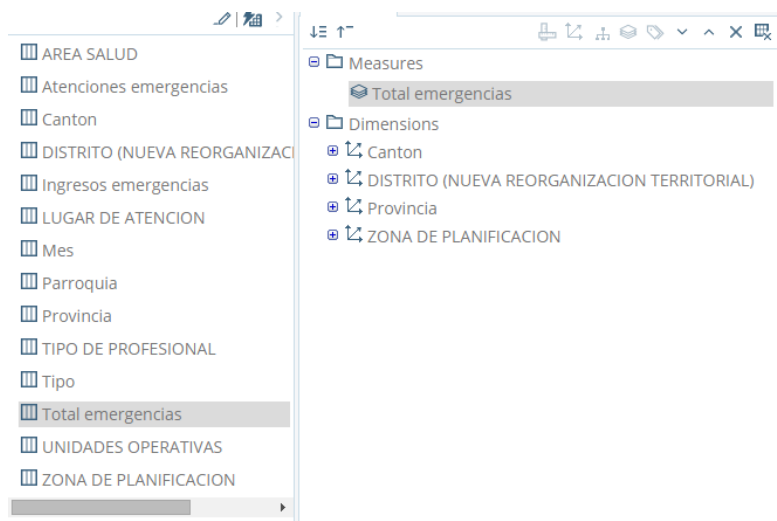
Figura 45: Nivel de Ocupación



Las cifras nos indican que el porcentaje de ocupación a nivel general es bajo, pero se puede observar que los hospitales Generales son los que tienen el más alto nivel de ocupación, en tanto que los hospitales básicos tienen un bajo nivel de ocupación. Este indicador les permite considerar, si el personal médico con el que cuenta una unidad hospitalaria es el adecuado para satisfacer la demanda de los usuarios que acuden a esta unidad hospitalaria.

Al final obtuvieron un BI para una coordinación Zonal de Salud, la información del BI corresponde a los tres niveles de atención que ha sido generado en sus distritos, con la cual se puede formar diferentes cubos para la obtención de indicadores según los requerimientos solicitados por los distintos departamentos, en la gráfica se puede observar el interfaz para obtener información.

Figura 46: Interfaz de Información



Los reportes obtenidos son archivos de tipo PDF con los resultados de los indicadores que fueron entregados a los Directores Departamentales para su análisis y toma de decisiones, además se entregó al departamento de TIC para su publicación.

5.2. Evaluación preliminar

Se hizo pruebas al BI en las fases de: consolidación de información, análisis de información y generación de reportes, originando como resultado, el perfecto funcionamiento del BI y la aceptación en un alto grado por parte de los profesionales que están a cargo de la administración de la información en Salud y quienes serán los involucrados de utilizar el BI para Salud. En el apéndice B se puede observar la documentación de las pruebas realizadas.

PLAN DE PRUEBAS AL BI

Historia de Revisión

Tabla 1: Historial

Fecha	Versión	Descripción	Autor
07/05/2015	1.0	Versión Final	Fabián Londo

Introducción

- Propósito del Plan

El propósito es documentar las pruebas de aceptación del BI aplicado a una Coordinación Zonal de Salud.

- Alcance

Verificación y determinación de las bondades del BI en la manejo de información y obtención de indicadores.

Requerimientos de Pruebas

Introducción

La generación de información en instituciones Públicas y privadas y no utilización adecuada de sistemas que puedan realizar un adecuado tratamiento, ha involucrado el desarrollo de los BI. Los mismos que exigen criterios lógicos que antes de su utilización deben ser probados por el personal que va utilizar el BI.

Filosofía de la prueba

- Generalidades

El objetivo principal de las pruebas del sistema será el de establecer un nivel de confianza que permite asegurar la aceptación del sistema por los usuarios (Estadístico Zonal, Analista Estadístico)

Se probará que la aplicación cumpla con los objetivos propuestos en la investigación, tengan las funcionalidades y características necesarias para la satisfacción del Departamento de Estadística.

- Áreas funcionales

- Funcionalidad especificada en el BI.
- Carga de archivos y procesos ETL

- Dimensiones y Hechos
- Calculo de Indicadores
- Reportes

- Categorías de resultados de prueba

1. Éxito: El resultado de la prueba es conforme al resultado esperado.
2. Aceptable: El resultado de la prueba indica que el sistema difiere de la especificación aceptada pero es aceptable, no son necesarios cambios en la aplicación, pero requiriendo un cambio en la Especificación Funcional.
3. Tolerable: El resultado de la prueba es incorrecto, la aplicación en prueba trabaja y podría ser aceptada, pero la falla deberá ser rectificada en el periodo de tiempo acordado.
4. Intolerable: El resultado de la prueba es incorrecto, y la falla debe ser corregida antes de concluir la fase de prueba.

Error: El resultado de la prueba observado es correcto, pero el resultado esperado de acuerdo

Entorno de la prueba

- Generalidades

En esta sección se da una breve descripción del entorno de prueba:

- Las pruebas se realizaron principalmente en la Dirección de Planificación, Departamento de Estadística.
- Utilizaron computadoras del departamento de Estadística.
- Para las pruebas utilizan información de la Coordinación Zonal de Salud 3

- Hardware

Tabla 2: Recursos

Recursos del BI	
Recurso	Nombre/Tipo/Número de serie
PC CORE i5, 256 Mb de memoria	ISB-24233264-1

- Software

En las computadoras de prueba se instalaron las siguientes aplicaciones de software:

- Sistema Operativo Windows 7
- JAVA2 v1.5.1 o mayor
- Data Cleaner
- Base de Datos PostgreSQL v9.2
- Pentaho V5.2

Datos de prueba

Para realizar las pruebas se tomó información correspondiente al 2013, información de bases de Internación, Emergencia, AIQ (Anestesia e Intervenciones Quirúrgicas), Laboratorio y Ginecología.

Roles y responsabilidades del equipo de pruebas

Tabla 3: Roles y Responsabilidades

Recursos humanos		
Cargo	Recursos mínimos necesarios	Responsabilidades específicas / comentarios
Administrador de pruebas	Fabián Londo	Proporcionar atención especial al funcionamiento correcto de las tareas principales del sistema. Responsabilidades: <ul style="list-style-type: none">• Proporcionar dirección técnica.• Ejecutar Pruebas• Elaborar Reportes.• Documentación de errores.

Identificación de la prueba

- Documentación de la prueba

Los resultados de la prueba son los siguientes:

Figura 47: Consola de Pentaho

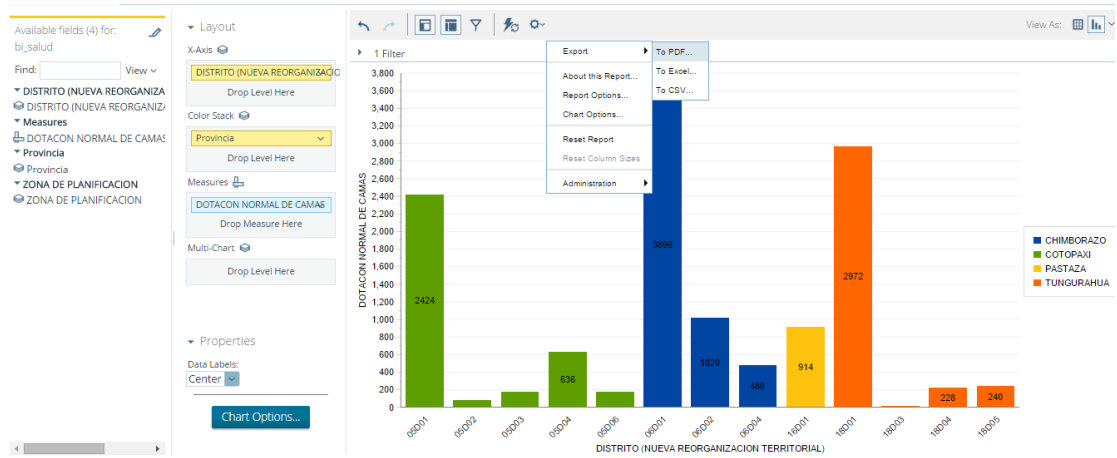


Figura 48: Resultados de la prueba (Partos)

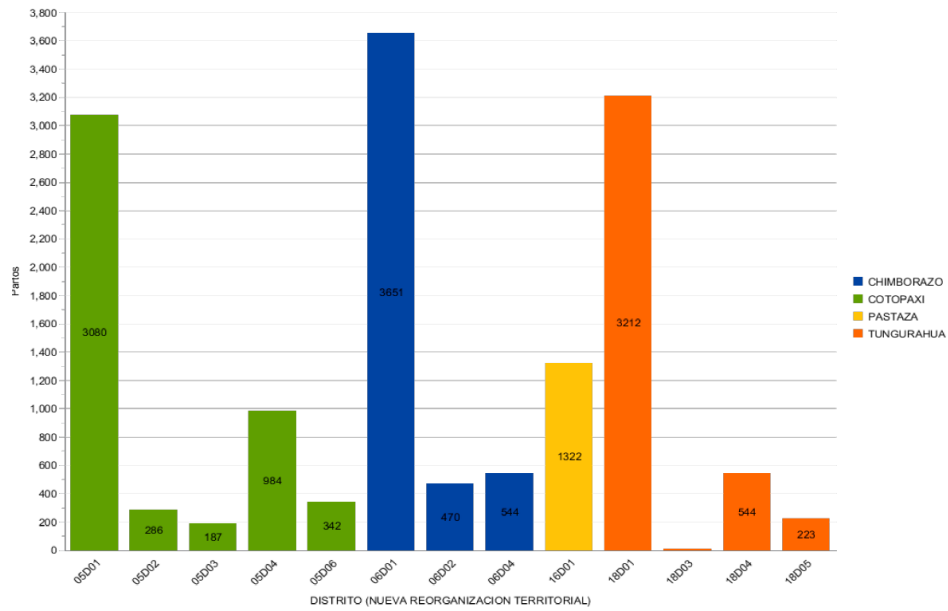
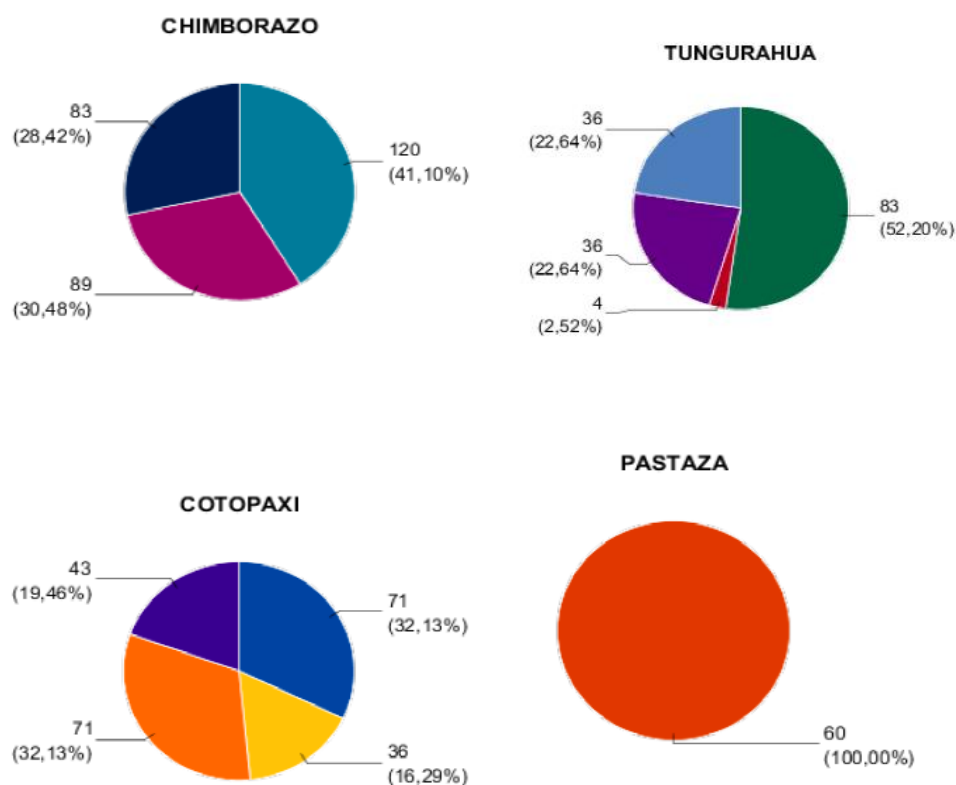


Figura 49: Resultados de la Prueba (Ocupación)



- Criterios de aceptación

El BI para salud fue probado por las tres personas integrantes del departamento de Estadística bajo los siguientes parámetros:

Tabla 4: Criterios

Aplicativo BI	Consolidación de Información	Consulta de Información	Funcionamiento de BI
Profesional 1	Excelente x Muy Bueno Bueno	Alto x Medio Bajo	Excelente Muy Bueno x Bueno
Profesional 2	Excelente x Muy Bueno Bueno	Alto Medio x Bajo	Excelente x Muy Bueno Bueno

Según los criterios de aceptación podemos decir que las pruebas realizadas al BI por los profesionales de estadística fueron de Éxito.

Requerimientos del Ambiente de Pruebas

A continuación se enumeran las características mínimas del ambiente para probar la Inteligencia de Negocios para Salud.

Hardware

- Intel CORE i5 de 600 MHz (recomendable CORE i7).
- Memoria RAM de 512 MB
- Disco Duro con capacidad libre de 30Gb.

Sistema Operativo

- Windows XP (recomendable Windows 7).

Se realizó una encuesta de satisfacción a los profesionales del Departamento de Planificación, proceso de Estadística quienes son los responsables directos de administrar la información de salud en una Coordinación, en si ellos serán los que utilicen el BI para un adecuado reporte de información e indicadores.

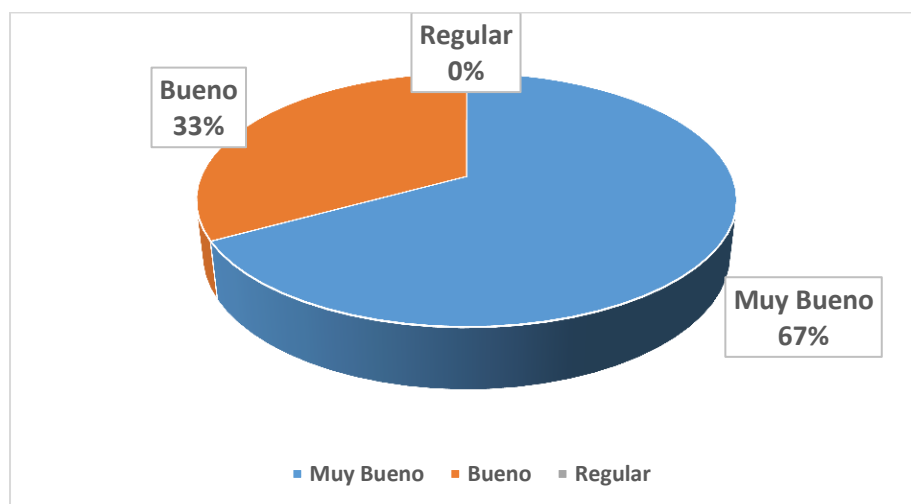
RESULTADOS DE LA ENCUESTA

1. ¿En qué escala considera de gran importancia el desarrollo de un Business Intelligence en software libre que es basado en indicadores de gestión, para una coordinación de salud?

Tabla 5: Importancia de un BI

Alternativa	Frecuencia	Porcentaje
Muy Bueno	2	67%
Bueno	1	33%
Regular	0	0
Total	3	100%

Figura 50: Importancia de un BI



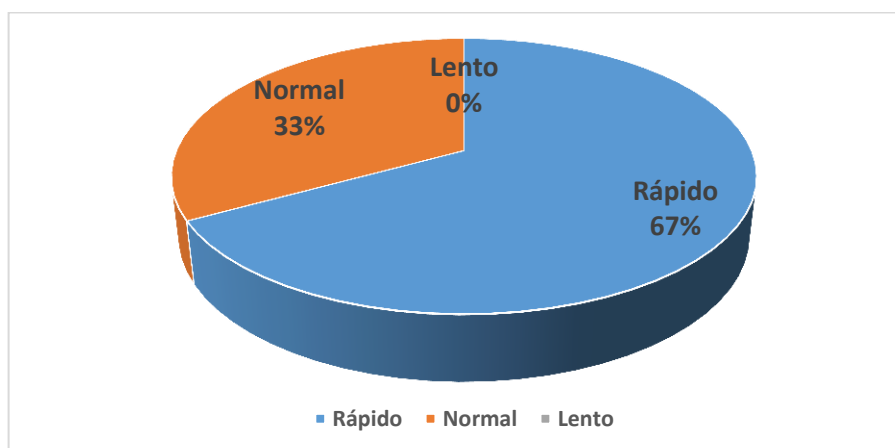
Los profesionales del proceso de estadística consideran que es Muy Bueno (67%) la implementación de un BI para desarrollar actividades de consolidación y análisis de información porque son unas de las principales funciones que realiza el departamento.

2. ¿Cómo considera el tiempo de respuesta del BI en el momento de realizar consultas?

Tabla 6: Tiempo de Respuesta a Consultas

Alternativa	Frecuencia	Porcentaje
Rápido	2	67%
Normal	1	33%
Lento	0	0
Total	3	100%

Figura 51: Tiempo de Respuesta a Consultas



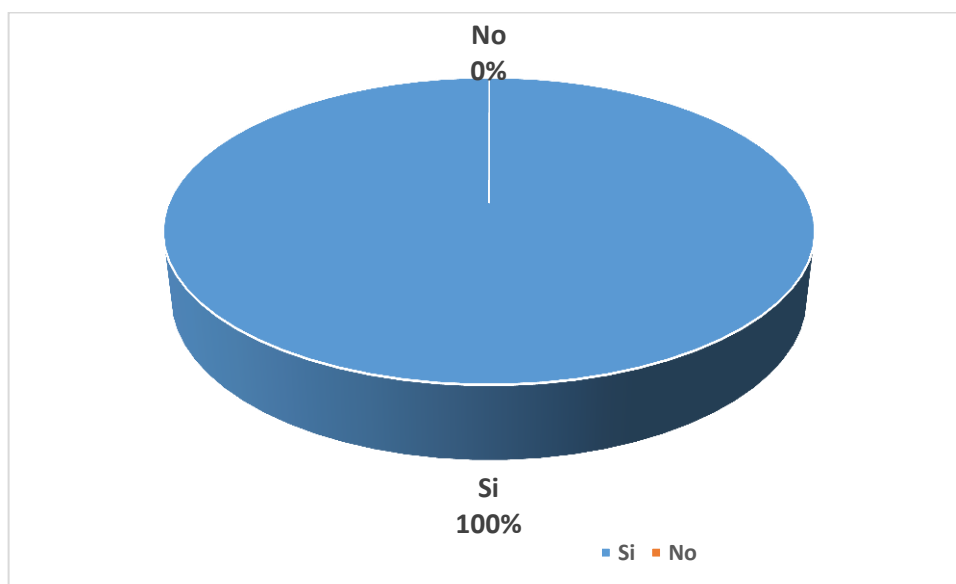
Al realizar consultas al BI los profesionales consideran que el tiempo de respuesta del aplicativo es rápido, con lo cual ellos podrán realizar de forma eficaz y a tiempo los pedidos de información realizado por los Departamentos pertenecientes a la Coordinación Zonal de Salud.

3. En el momento de hacer consultas, ¿la información obtenida es la adecuada?

Tabla 7: Respuesta a Consultas

Alternativa	Frecuencia	Porcentaje
Sí	3	100%
No	0	0%
Total	3	100%

Figura 52: Respuestas Consistentes



Elaborado por: Fabián Londo Y.

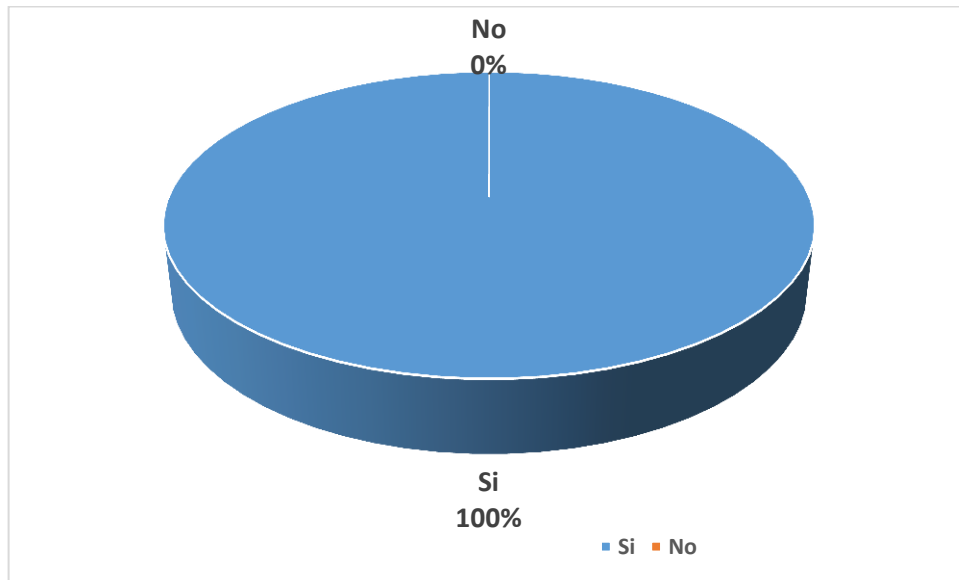
Al realizar consultas al BI por parte de los profesionales del Departamento de Estadística ellos consideran que los resultados de las consultas realizadas, son totalmente representativas y adecuadas.

4. ¿Considera fácil el uso del BI?

Tabla 8: Facilidad de Uso

Alternativa	Frecuencia	Porcentaje
Sí	3	100%
No	0	0%
Total	3	100%

Figura 53: Uso del Sistema



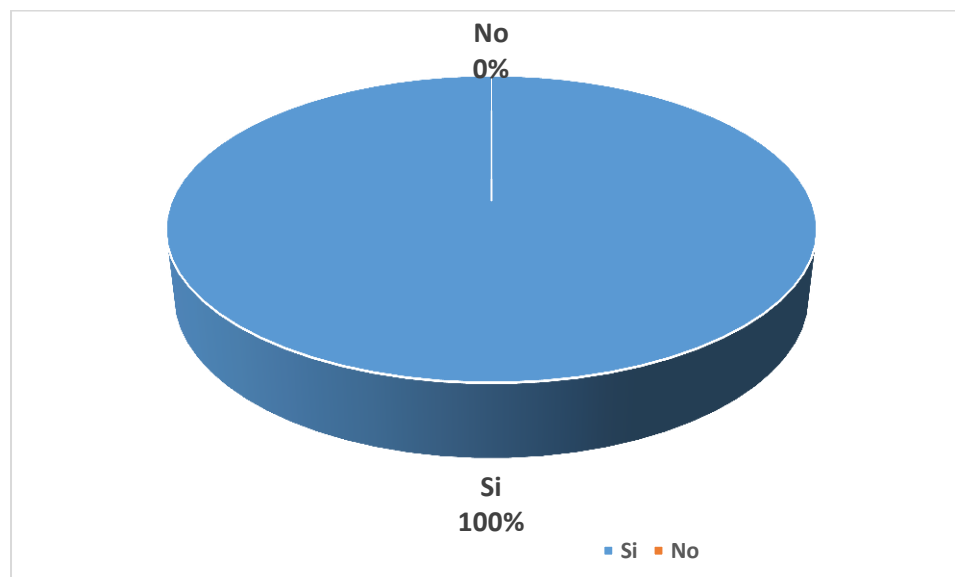
Los profesionales del proceso de Estadística luego de interactuar y utilizar el BI, determinan, que se sintieron cómodos por la versatilidad que les da el BI para la obtención de información.

5. ¿Cree que los reportes generados en el BI son entendibles?

Tabla 9: Reportes Entendibles

Alternativa	Frecuencia	Porcentaje
Sí	3	100%
No	0	0%
Total	3	100%

Figura 54: Reportes Importantes



Los profesionales de la Dirección de Planificación, proceso de Estadística, luego de interactuar, realizar consultas y obtener reportes con el BI, consideran en su totalidad, que las consultas, análisis y reportes de información generados por el BI son los requeridos y ayudan de gran manera a solventar consultas que son pedidas a nivel interno y externo a la Coordinación Zonal de Salud.

5.3. Análisis de resultados

Se verificó y validó la información utilizando un Muestreo Aleatorio Estratificado con un 90% de información confiable y un margen de error del 10%, es decir se tomó una muestra de la información recolectada en los distritos y se verificaba que sea la misma que consta en los partes físicos.

Se realizó ETL (extracción, transformación y carga) de datos provenientes de distintos tipos de archivos.

Se obtuvo una Data Warehouse en el cual reposa información de primero, segundo y tercer nivel de atención que brinda el Ministerio de Salud Pública a través de sus diferentes servicios.

Se logró definir una estructura de información para Pentaho, capaz de administrar datos de Salud de distintos años y con la perspectiva de seguir almacenando información de años anteriores para poder llegar a conformar series cronológicas de información.

Se construyó una lista inicial de indicadores contemplados según la necesidad actual de una Coordinación Zonal de Salud.

Se consiguieron reportes generados por el BI en formato PDF que contienen indicadores de salud, los mismos que serán entregados a los Directores Departamentales para la toma de decisiones.

Se diseñó un BI con Pentaho que tiene plataforma libre, lo cual permitirá realizar modificaciones o agregaciones que sean necesarias, según los requerimientos de la Coordinación Zonal de Salud.

Capítulo 6

Conclusiones y Recomendaciones

6.1. Conclusiones

Se consolidó información del primero, segundo y tercer nivel de atención en Salud, y se analizó utilizando estadística descriptiva (Medidas de Tendencia central y frecuencias) para observar la consistencia de la información, de lo cual se concluye que existe consistencia y fiabilidad de la información en un 94%.

La utilización de Software Libre para la construcción del BI permitió utilizar las diferentes librerías que dispone Pentaho para la elaboración de un BI y adecuarlas a la información que se tiene en un Coordinación Zonal de Salud, enmarcándonos dentro de las Políticas de Gobierno como es la de “Establecer como política pública para las entidades de la Administración Pública Central la utilización de Software Libre en sus sistemas y equipamientos informáticos”.

En una Coordinación Zonal de Salud se concentra información de los Distritos que la componen y en diferentes formatos y Pentaho por su flexibilidad permitió consolidar la información en una sola base.

De la base de indicadores obtenidos podemos mencionar que existe consistencia en los valores resultantes por la pertinencia que existe en cada uno de los indicadores y su consistencia en la base analizada.

Se concluye que Pentaho a través de sus diversas herramientas permitió realizar el análisis adecuado de la información, obteniendo indicadores de salud fiables que sirvió para tener un enfoque global del servicio que se está brindando.

6.2. Recomendaciones

Para un eficiente uso de herramientas de BI en las empresas sean públicas o privadas, se recomienda que en cada una de ellas se analice su situación actual en cuanto al manejo de datos e información y la manera en la que los Directivos realizan su proceso de toma de decisiones, para que de esta manera puedan elegir la mejor herramienta tecnológica para la utilización de un BI, que le permita ser eficientes y eficaces en la administración.

Se sugiere a las Coordinaciones Zonales de Salud la utilización de un Sistema BI para el almacenamiento y análisis de información estadística de salud, estructurando un plan de capacitaciones al personal de estadística sobre temas relacionados en el manejo de bases de datos y aplicaciones BI.

Se recomienda la utilización de la herramienta Pentaho para el diseño de un BI por la flexibilidad para la consolidación de varios tipos de archivos lo cual permite estructurar de manera más eficiente un BI, y que además se adapta a las políticas actuales de Gobierno de utilizar software libre.

La tecnología día tras día va avanzando, la actualización periódica de sistemas de información como de herramientas de BI son muy aconsejables en las instituciones para mantener sus funcionalidades y servicios, así como para poder ampliar su capacidad de almacenamiento y consulta, por lo cual se recomienda utilizar Pentaho.

Apéndice A

Encuesta de Satisfacción

ENCUESTA DE UTILIZACION DEL SOFTWARE PENTAHO

Objetivo: Determinar el grado de satisfacción que tienen los profesionales del proceso de Estadística al utilizar un BI (Inteligencia de Negocios).

Lugar de Aplicación: Departamento de Estadística de la Coordinación Zonal de Salud

CUESTIONARIO

Pregunta 1: En qué escala considera de gran importancia el desarrollo de un Business Intelligence en software libre que es basado en indicadores de gestión, para una Coordinación de Salud.

Muy Bueno Bueno Regular

Pregunta 2: Cómo considera el tiempo de respuesta del BI en el momento de realizar consultas.

Rápido Medio Lento

Pregunta 3: En el momento de hacer consultas al BI, la información obtenida es la adecuada.

Sí No

Pregunta 4: Considera fácil el uso del BI

Sí

No

Pregunta 5: Cree que los reportes generados en el BI son entendibles

Sí

No

Elaborado por:

Fabián Londo

Apéndice B

INSTRUCTIVO PARA INSTALAR POSTGRESQL Y PENTAHO

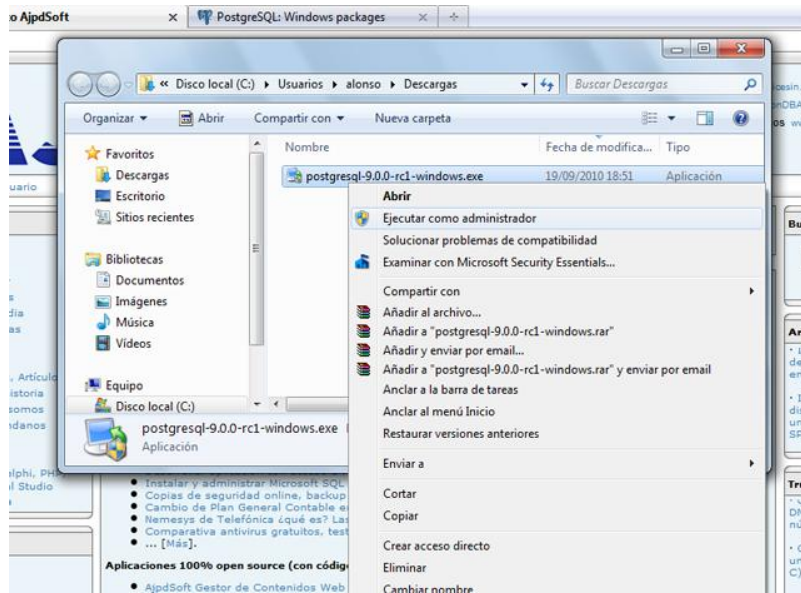
Descarga del instalador: para la descarga del instalador ingresan a la dirección web www.enterprisedb.com como se muestra en la siguiente gráfica.

Figura 55: Dirección web para descargar PosgreSQL



Una vez descargado el instalador se lo debe ejecutar haciendo doble clic en archivo descargado, en la siguiente figura evidencia el instalador descargado.

Figura 56: Instalador de PostgreSQL descargado.



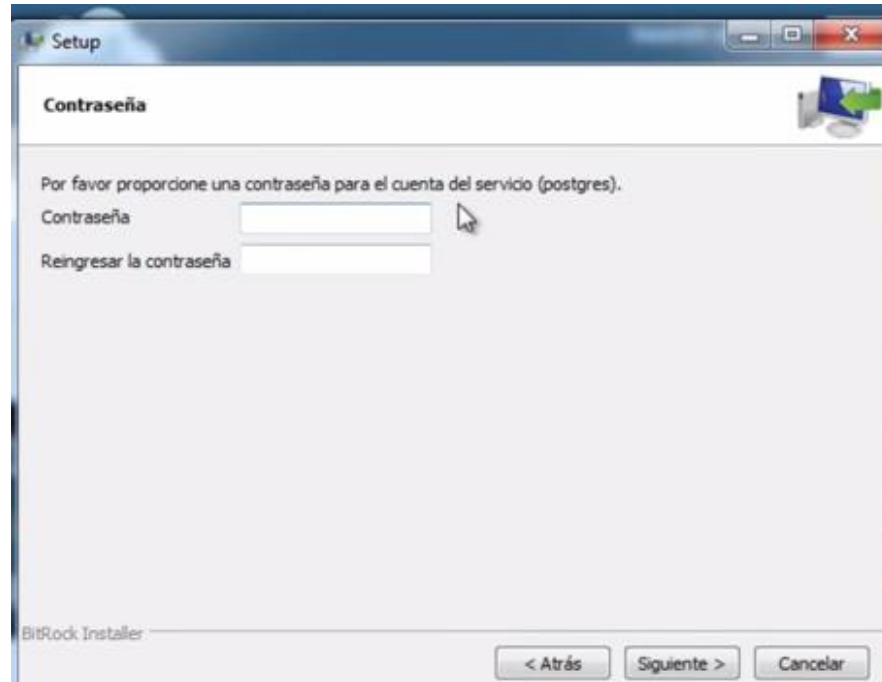
Al realizar doble clic en el archivo descargado se visualiza la siguiente ventana de instalación, para continuar se debe dar siguiente.

Figura 57: Ventana de Instalación de PostgreSQL



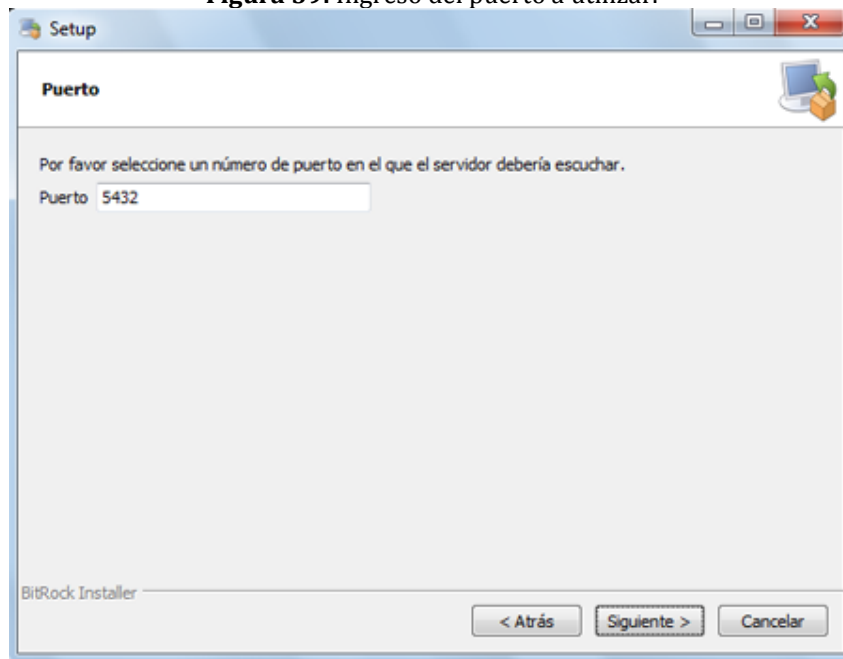
Posteriormente se despliega otra ventana en donde se debe ingresar una contraseña para el administrador de la base de datos, esta contraseña se deberá proporcionar cada vez que se desee ingresar a postgresQL para la creación de una base de datos o la conexión desde una aplicación a potsgres. En la siguiente imagen se visualiza este proceso.

Figura 58: Ventana para el ingreso de contraseña para PostgreSQL.



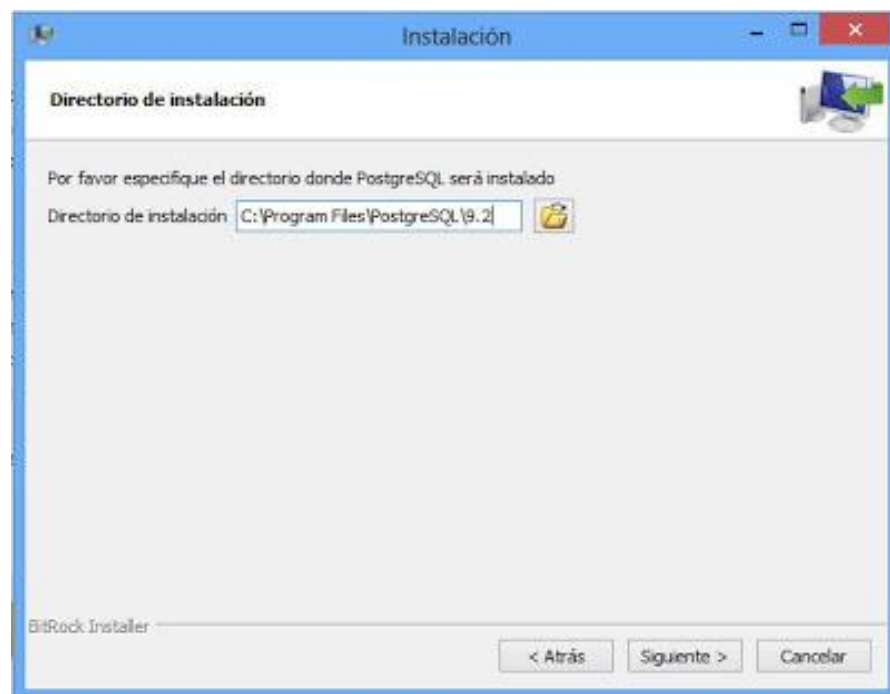
Al seguir en el proceso de instalación saldrá otra ventana en donde les pide la selección del puerto que utilizara PostgreSQL para la comunicación con otras aplicaciones, por defecto es el 5432 este puerto se puede modificar en el caso de que se desee hacerlo. La siguiente imagen visualiza este proceso.

Figura 59: Ingreso del puerto a utilizar.



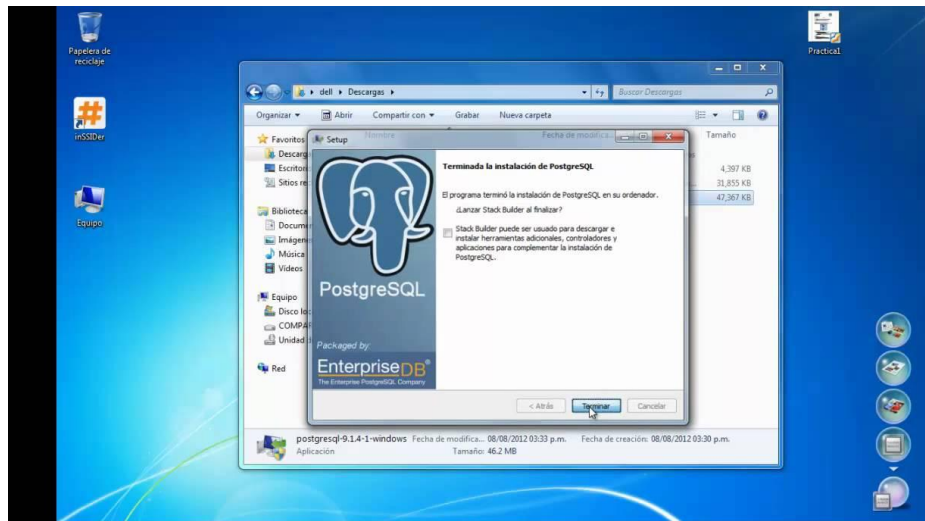
Luego sale la selección del directorio de instalación, que por defecto se guarda en el disco C en archivos de programa, si deseamos modificarlo y guardarlo en otro lugar se lo puede hacer. La siguiente imagen visualiza este proceso.

Figura 60: Directorio de Instalación.



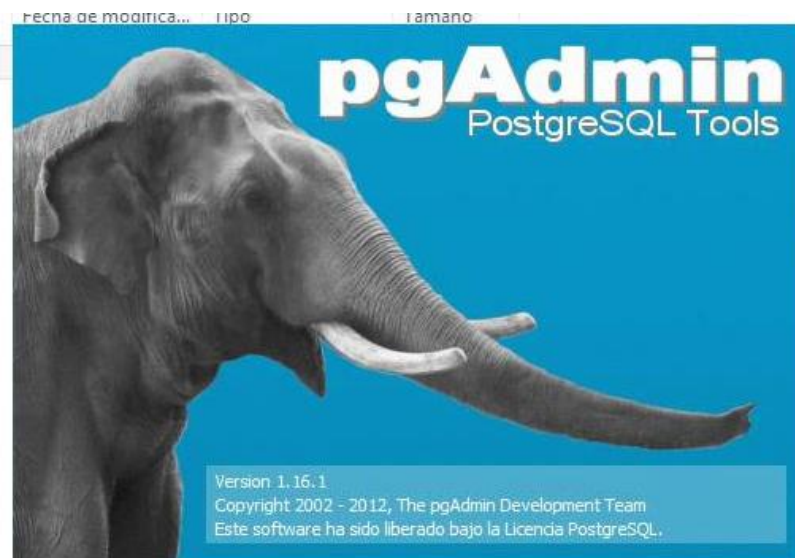
A continuación se da clic en siguiente hasta llegar a la pantalla de finalización, en la cual se da clic en terminar para finalizar la instalación. La siguiente imagen visualiza este proceso.

Figura 61: Finalización de la instalación.



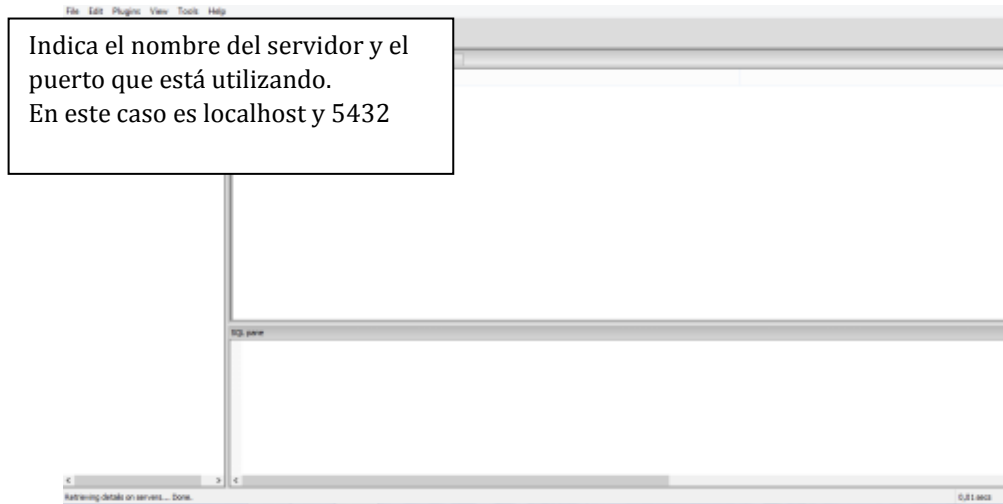
Una vez finalizada la instalación se procede a su utilización, al abrir la aplicación aparecerá la siguiente pantalla.

Figura 62: Pantalla de inicio de PostgreSQL.



A continuación les aparece la pantalla donde se va a trabajar en la creación de la base de datos, es aquí donde me va a pedir la clave para poder trabajar. La siguiente imagen visualiza el proceso.

Figura 63: Pantalla para trabajar con PostgreSQL.



Nota: El instalador incluye pgAdmin que es el que permite interactuar con Postgres para la creación de la base de datos.

La versión de Postgres instalada es la 9.2 porque es una versión más estable en relación a la 9.4.

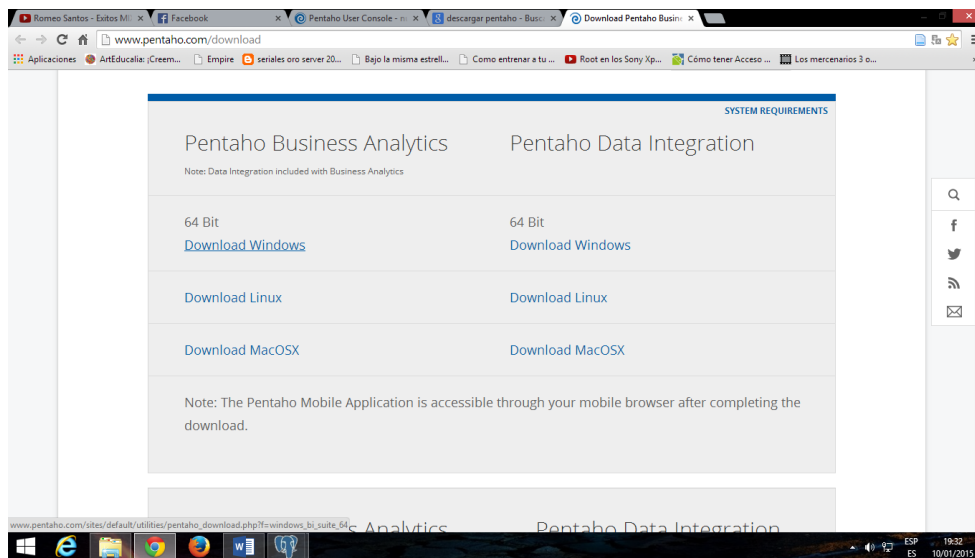
Apéndice C

INSTRUCTIVO PARA INSTALAR PENTAHO

El Link de descarga de Pentaho es el siguiente: <http://www.pentaho.com/download>

En esta pantalla se selecciona el sistema operativo y el tipo de Pentaho que se requiere instalar, en nuestro caso es el Business Analytics. En la siguiente gráfica se visualiza este proceso.

Figura 64: Pantalla de descarga de Pentaho.

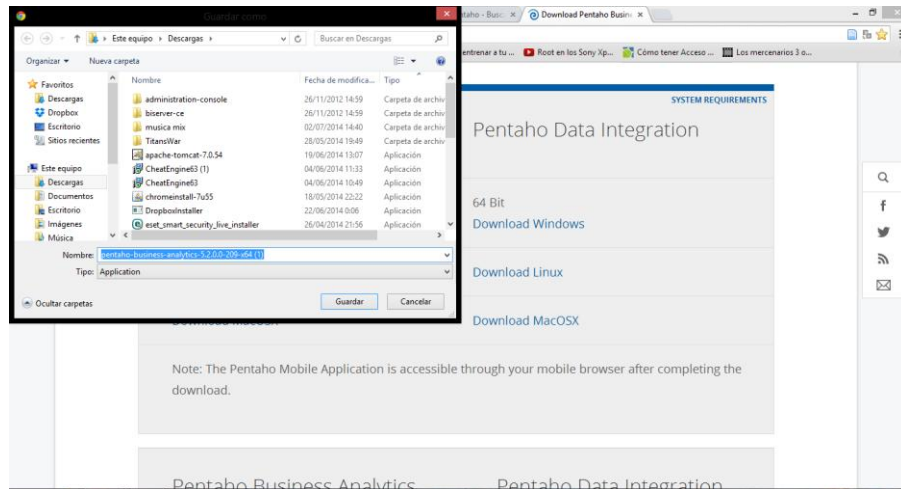


Al hacer clic en descargar les pide la ruta donde se desea guardar el instalador.

Nota: Esta versión que se va a instalar es un archivo .exe el cual ya trae las principales carpetas que se necesitan para la creación de informes, no contiene el driver para conexión a la base de datos ese debe ser descargado del Link del driver: <http://jdbc.postgresql.org/download.html>.

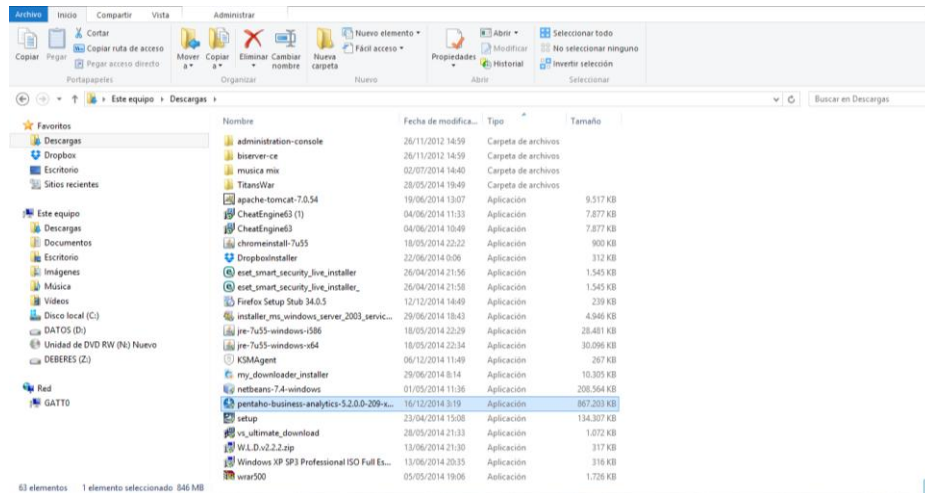
En la siguiente imagen se visualiza la ruta donde se desea instalar Pentaho.

Figura 65: Ruta de la instalación de Pentaho.



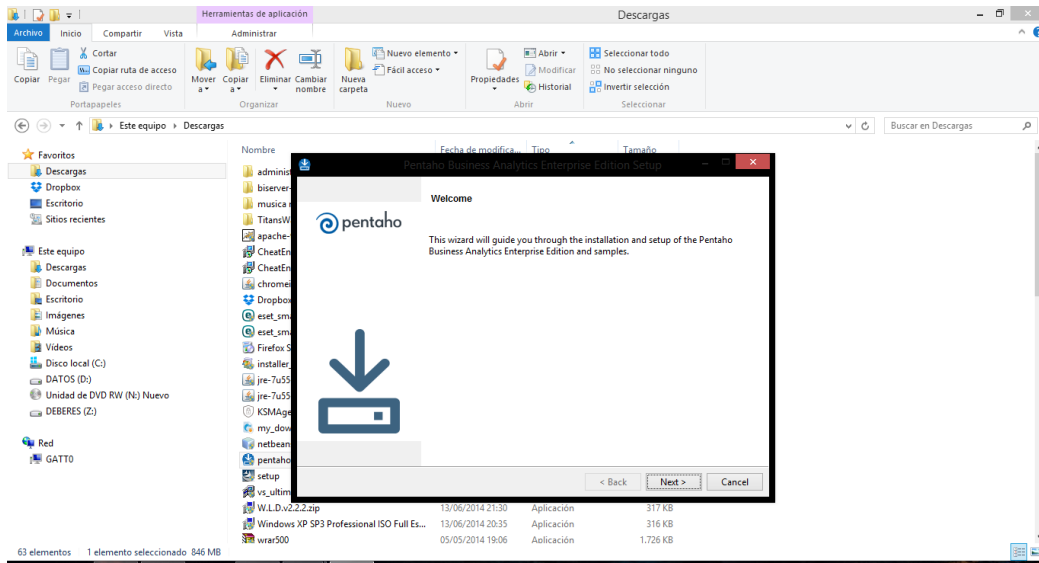
Una vez descargado el instalador se procede a la instalación, dando doble clic sobre el mismo. La grafica nos visualiza el instalador donde hacemos clic.

Figura 66: Descarga de Pentaho.



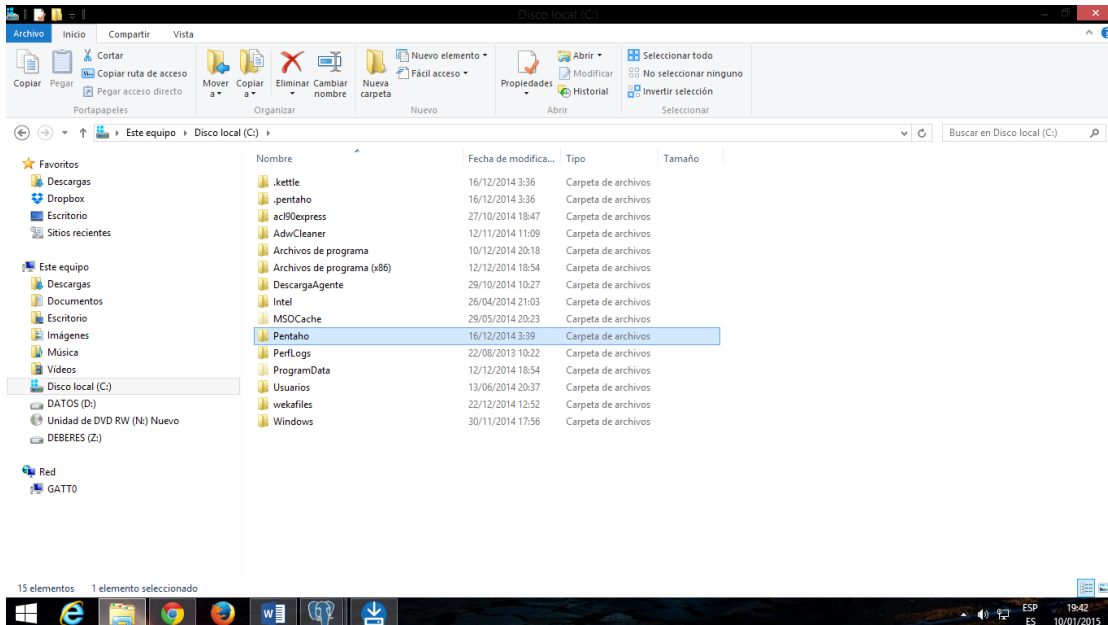
Posteriormente nos sale otra ventana, en la cual solo damos un clic en Next (siguiente). La gráfica nos visualiza este proceso.

Figura 67: Proceso de instalación.



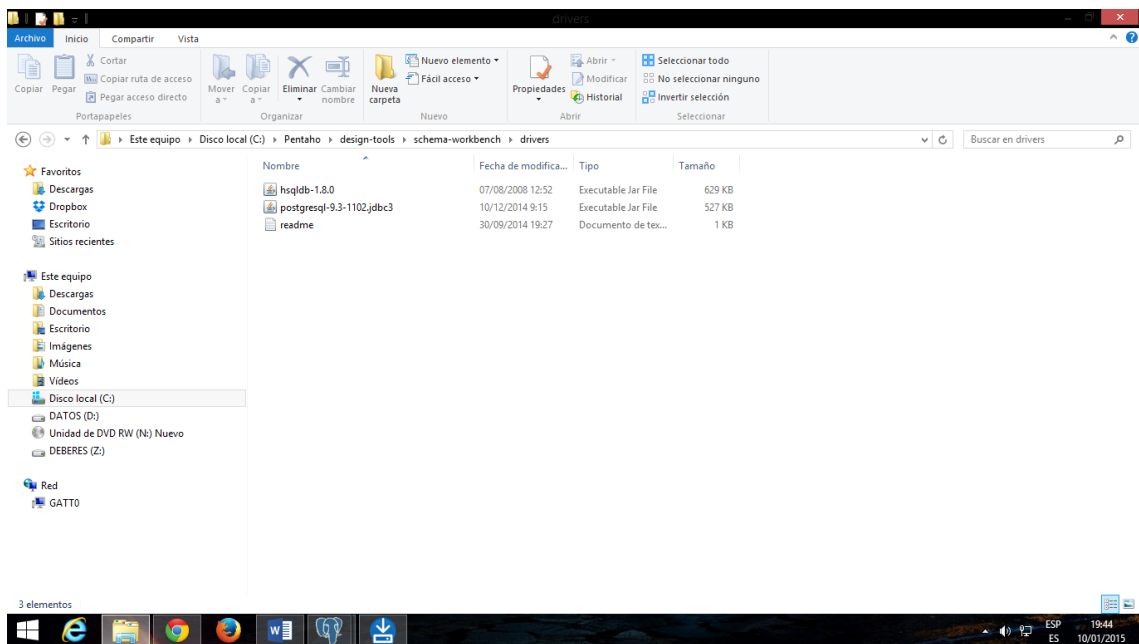
Una vez instalado, en la unidad de disco C vamos a encontrar una carpeta de nombre Pentaho. En la siguiente imagen se visualiza la carpeta donde fue instalado Pentaho.

Figura 68: Carpeta de Pentaho.



Dentro de la carpeta Pentaho se ubican en la carpeta design-tools, se abre esta carpeta y encontramos otra con el nombre Data Integration, abrimos la carpeta y encontramos una carpeta driver en ella copiamos el driver descargado anteriormente. En la siguiente imagen se visualiza la ruta donde va a ser pegado el driver descargado.

Figura 69: Carpeta para pegar Driver descargado.



Nota: Si se utiliza una base de datos diferente a postgres, el driver (jdbc) al descargar es otro, para lo cual se debe descargar el driver dependiendo de la base de datos, el driver a instalar debe ser la misma versión que la del administrador de la base de datos en este caso de la de PostgreSQL.

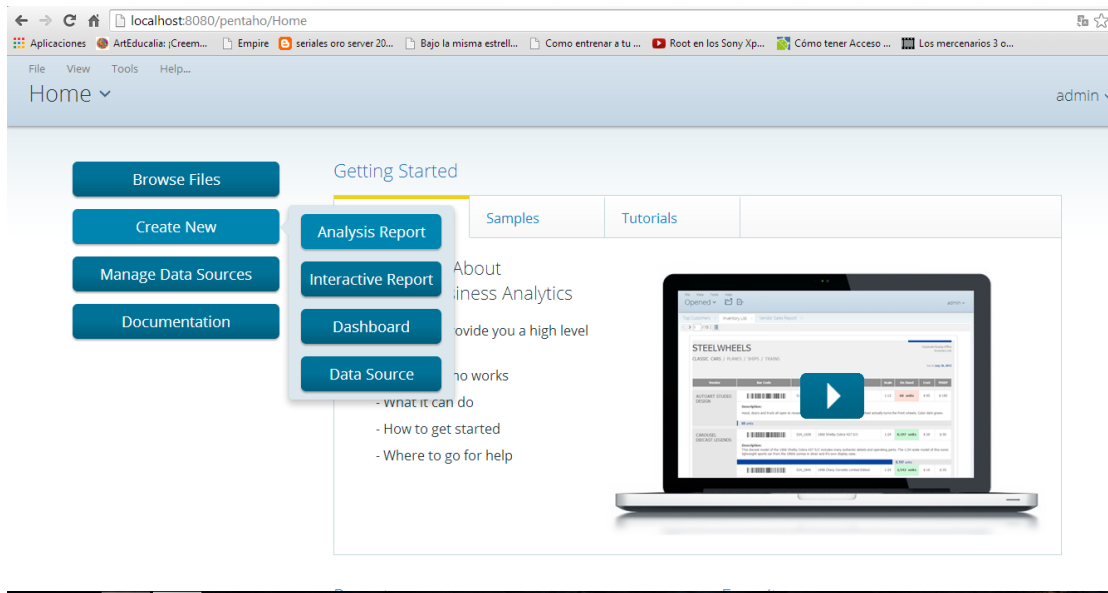
Apéndice D

Ingreso a la Consola de Pentaho

Para ingresar a la consola de usuario de Pentaho se ingresa al localhost, la dirección por defecto es: `http://localhost:8080/`.

El usuario es Admin y la contraseña password, en el siguiente gráfico se puede observar el interfaz de la consola de Pentaho.

Figura 70: Consola de Pentaho.



REFERENCIAS

- [1] UNIVERSIDAD CENTRAL DEL ECUADOR, *Estudios de Empleabilidad y necesidades de profesionales para el Sistema Nacional de Salud*, Quito, 2014.
- [2] MOYA LIGIA, *Introducción a la estadística de la salud: curso básico para estudiantes de ciencias de la salud*, Costa Rica, 2000.
- [3] HEREDIA ALVARADO JOSE ANTONIO, *Sistemas de Indicadores para la mejora y el control Integrado de la Calidad de los Procesos*, 2001.
- [4] GUTTMAN EDITH, ZORRO CARLOS, *Diseño de un Sistema de Indicadores Socio Ambientales para el Distrito Capital de Bogotá*, Julio, 2004
- [5] ESTEVEZ GARCIA JESUS FRANCISCO, *Sistema de Indicadores para el Diagnostico y Seguimiento de la Educación Superior*, México, 2007.
- [6] MINISTERIO DE SALUD PUBLICA, *Sistema de Indicadores Hospitalarios de Salud*, 2013
- [7] UNIVERSIDAD CENTRAL DEL ECUADOR, SUBIA GARCIA JORGE, *Implementación de una solución Business Intelligent para el Análisis de Gestión del Ministerio de Desarrollo Urbano y Vivienda*, Marzo, 2013
- [8] DIAZ CURTO JOSEP, *Introducción al Business Intelligent*, Mayo, 2010.
- [9] STALLMAN RICHARD, *Software Libre para una Sociedad Libre*, Noviembre, 2004
- [10] UNESCO, *Guía práctica sobre software libre: su selección y aplicación local en América Latina y el Caribe*, 2002.
- [11] ESPOCH, ILBAY ELSI, *Propuesta Metodológica para aplicar Business Intelligent, caso práctico "COHERVIC S.A."*, Octubre, 2009.
- [12] BOUMAN ROLAN, *Pentaho Solutions*, Agosto, 2009.
- [13] MOLINER LÓPEZ FRANCISCO, *Grupos A y B de Informática Bloqueo Especifico*, Agosto, 2005.
- [14] PIÑEIRO GOMEZ JOSE MANUEL, *Definición y Manipulación de Datos*, 2014.
- [15] VIVANCO MANUEL, *Muestreo Estadístico Diseño y Aplicaciones*, Diciembre, 2005.
- [16] LEMUS JUAN DAVID, *Aplicación de un Business Intelligent en Instituciones Públicas*

de Salud de Colombia, 2005.

- [17] GONZALES GERMAN, *Estudios de factibilidad de una Inteligencia de Salud en Hospitales*, Costa Rica, 2013.
- [18] GARCIA LUCINA, *Solución Matemático Computacional para la Contextualización de una estrategia Senológica en Cuba*, 2012.
- [19] RICO WILLMER, *Estudio realizado en Europa por Information Builders Ibéric mostró el costo que tiene la falta de sistemas de toma de decisiones en las organizaciones*, 2008.

Resumen Final

Desarrollo de un Business Intelligence en software libre, basado en indicadores de gestión, para una coordinación de salud.

Fabián Patricio Londo Yachambáy

95 páginas

Proyecto Dirigido por: Enrique Xavier Garcés Freire Ing. Master

La nueva estructuración de Zonas, Distritos y Circuitos que lleva adelante el estado ecuatoriano ha generado una centralización de información hacia las 9 zonas que tiene el país, las cuales se están fortaleciendo y consolidando en sus diferentes ámbitos. Sin embargo es notoria la falta de aplicativos que ayuden a analizar y procesar información para tomar decisiones oportunas y adecuadas en el momento indicado. El desarrollo de un (BI; Inteligencia de Negocios) usando software libre como estipula la política de gobierno, permite realizar el procesamiento y análisis de la información de manera rápida, fiable y oportuna mediante la utilización de técnicas estadísticas, para ello se cuenta con una base de indicadores que permitirán tomar decisiones en la parte administrativa y gerencial, el sistema de BI(Inteligencia de Negocios) está dirigido a los Directores de procesos de las Coordinaciones Zonales de Salud quienes son los encargados de tomar decisiones para que la Salud Pública sea de calidad y llegue a todos los ecuatorianos. El Producto final del trabajo son reportes realizados en Pentaho que permiten visualizar estadísticas de salud, extraídos a través de una concentración, procesamiento y análisis de información de todas las unidades operativas de primero, segundo y tercer nivel que tiene cada Coordinación Zonal de Salud.