



UNIVERSIDAD TÉCNICA DEL NORTE

INSTITUTO DE POSTGRADO



MAESTRÍA EN INGENIERÍA DE SOFTWARE

“Análisis de datos para mejorar la toma de decisiones en la distribución de agua potable del cantón Ibarra, utilizando Business Intelligence.”

Trabajo de Grado previo a la obtención del Título de Magíster en Ingeniería de Software

DIRECTOR:

Ing. Jorge Caraguay Procel. Mgs

AUTOR:

Ing. Andrés Fernando Cárdenas Pepinós

IBARRA - ECUADOR

2018

APROBACIÓN DEL TUTOR

En mi calidad de tutor de grado: “ANÁLISIS DE DATOS PARA MEJORAR LA TOMA DE DECISIONES EN LA DISTRIBUCIÓN DE AGUA POTABLE DEL CANTÓN IBARRA, UTILIZANDO BUSINESS INTELLIGENCE”, presentado por el **Ing. Andrés Fernando Cárdenas Pepinós**, para optar por el grado de Magíster en Ingeniería de Software, ha sido guiado y revisado periódicamente y cumple normas establecidas en el reglamento de estudiantes de la Universidad Técnica del Norte, por lo que doy fe que dicho trabajo reúne los requisitos y méritos suficientes para ser sometido a presentación pública y evaluación por parte del Jurado Examinador que se designe.

En la ciudad de Ibarra, 10 de mayo de 2018



Ing. Jorge Adrián Caraguay Procel MSc.

CI. 1102451687

TUTOR DE TRABAJO DE GRADO

APROBACIÓN DEL ASESOR

“ANÁLISIS DE DATOS PARA MEJORAR LA TOMA DE DECISIONES EN LA DISTRIBUCIÓN DE AGUA POTABLE DEL CANTÓN IBARRA, UTILIZANDO BUSINESS INTELLIGENCE”

Por: Ing. Andrés Fernando Cárdenas Pepinós.

Trabajo de grado de Maestría aprobado en nombre de la Universidad Técnica del Norte, por el asesor, a los diez días del mes de mayo de 2018.



.....
Ing. Daysi Elizabeth Imbaquingo Esparza, Msc.

C.I. 1002873048

ASESOR

AUTORIZACIÓN DE USO Y PUBLICACIÓN A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE

1.- IDENTIFICACIÓN DE LA OBRA

La Universidad Técnica del Norte dentro del proyecto Repositorio Digital Institucional, determinó la necesidad de disponer de textos completos en formato digital con la finalidad de apoyar los procesos de investigación, docencia y extensión de la Universidad.

Por medio del presente documento dejo sentada mi voluntad de participar en este proyecto, para lo cual pongo a disposición la siguiente información:

DATOS DE CONTACTO			
Cédula de ciudadanía	1003191044		
Apellidos y Nombres	Cárdenas Pepinos Andrés Fernando		
Dirección	AV. 13 DE ABRIL Y CALLE EL ORO		
Email	andresfcardenas1989@gmail.com		
Teléfono fijo	062546870	Celular	0987297537

DATO DE LA OBRA		
Título:	ANÁLISIS DE DATOS PARA MEJORAR LA TOMA DE DECISIONES EN LA DISTRIBUCIÓN DE AGUA POTABLE DEL CANTÓN IBARRA, UTILIZANDO BUSINESS INTELLIGENCE.	
Autor:	Cárdenas Pepinós Andrés Fernando	
Fecha:	10/05/2018	
SOLO PARA TRABAJOS DE GRADO		
PROGRAMA	Pregrado	Posgrado
		X
Título por el que opta	Magister en Ingeniería de Software	
Asesor / Director	Ing. Jorge Caraguay Procel Msc.	

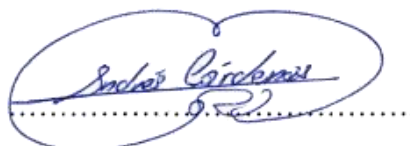
AUTORIZACIÓN DE USO A FAVOR DE LA UNIVERSIDAD

Yo, Andrés Fernando Cárdenas Pepinós con cédula de identidad Nro. 1003191044, en calidad de autor y titular de los derechos patrimoniales de la obra o trabajo de grado descrito anteriormente, hago entrega del ejemplar respectivo en formato digital y autorizo a la Universidad Técnica del Norte, la publicación de la obra en el Repositorio Digital Institucional y uso del archivo digital en la Biblioteca de la Universidad con fines académicos, para ampliar la disponibilidad del material y como apoyo a la educación, investigación y extensión; en concordancia con la Ley de Educación Superior Artículo 144.

CONSTANCIAS

El autor manifiesta que la obra objeto de la presenta autorización es original y se la desarrolló, sin violar derechos de autor de terceros, por lo tanto, la obra es original y que es el titular de los derechos patrimoniales, por lo que asume la responsabilidad sobre el contenido de la misma y saldrá en defensa de la universidad en caso de reclamación por parte de terceros.

Ibarra, a los diez días del mes de mayo de 2018

A handwritten signature in blue ink, enclosed in a blue oval. The signature reads "Andrés Cárdenas" and is followed by a dotted line.

Andrés Fernando Cárdenas Pepinós

C.I. 1003191044

AUTOR

**CESIÓN DE DERECHOS DE AUTOR
A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE**

Yo, Andrés Fernando Cárdenas Pepinós, con cédula de identidad Nro. 1003191044, manifiesto mi voluntad de ceder a la Universidad Técnica del Norte los derechos patrimoniales consagrados en la Ley de Propiedad Intelectual del Ecuador, artículos 4, 5 y 6, en calidad de autor de la obra o trabajo de grado denominado: **“ANÁLISIS DE DATOS PARA MEJORAR LA TOMA DE DECISIONES EN LA DISTRIBUCIÓN DE AGUA POTABLE DEL CANTÓN IBARRA, UTILIZANDO BUSINESS INTELLIGENCE”**, que ha sido desarrollado para optar por el título de: Magister en Ingeniería de Software en la Universidad Técnica del Norte, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En mi condición de autor me reservo los derechos morales de la obra antes citada.

En concordancia suscribo este documento en el momento que hago entrega del trabajo final en formato impreso y digital a la Biblioteca de la Universidad Técnica del Norte.



Andrés Fernando Cárdenas Pepinós

C.I. 1003191044

Autor

DEDICATORIA

A mi esposa Mayra,

Por su comprensión y apoyo para culminar con éxito esta nuevo objetivo en mi vida académica y por ser la compañera ideal para cumplir nuestras metas y objetivos trazados en nuestras vidas.

A mis padres Guillermo e Isaura,

Por siempre confiar en mí y brindarme su apoyo incondicional en las metas que me proponga.

A mis hermanos,

Por la ayuda emocional para poder culminar este objetivo.

Att: Andrés F Cárdenas.

AGRADECIMIENTO

A la gloriosa Universidad Técnica del Norte, quienes me brindaron la oportunidad de obtener mi Maestría en Ingeniería de Software, así ayudándome a crecer tanto personal como profesionalmente con todos los conocimientos impartidos.

A mi director de tesis Ing. Jorge Caraguay Msc, por todos los consejos impartidos, para cada día avanzar y culminar con éxito el presente trabajo de grado.

A los compañeros del proyecto Ibarra verde, por la dedicación y entrega presta para que el trabajo sea el mejor y poder alcanzar nuestros objetivos.

A la Empresa de Agua Potable y Alcantarillado de la ciudad de Ibarra, por la ayuda incondicional en cuanto a la obtención de la información.

Att: Andrés F Cárdenas.

INDICE

APROBACIÓN DEL TUTOR	II
APROBACIÓN DEL ASESOR	III
AUTORIZACIÓN DE USO Y PUBLICACIÓN A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE	IV
AUTORIZACIÓN DE USO A FAVOR DE LA UNIVERSIDAD	V
CONSTANCIAS	V
CESIÓN DE DERECHOS DE AUTOR A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE.....	VI
DEDICATORIA.....	VII
AGRADECIMIENTO	VIII
INDICE.....	IX
ÍNDICE DE ILUSTRACIONES	XIV
ÍNDICE DE TABLAS.....	XVI
RESUMEN	17
ABSTRACT	18
CAPITULO I.....	19
1 Introducción	19
1.1. Antecedentes	20
1.2. Planteamiento del Problema.....	21
1.3. Formulación del problema.....	22
1.4. Justificación de la investigación.....	22
1.5. Objetivos de la investigación	23
1.5.1. Objetivo General	23
1.5.2. Objetivos Específicos.....	23
1.6. Hipótesis o preguntas de investigación	24
1.6.1. Preguntas directrices	24
CAPITULO II.....	25
2 Introducción	25
2.1 Estado del Arte	25
2.2 Marco Legal	26

2.2.1	Matriz de Tensiones y Problemas	27
2.2.2	Constitución del Ecuador 2008	27
2.2.3	El Plan Nacional para el Buen Vivir	29
2.2.4	Reglamento de Régimen Académico	31
2.3	Marco Referencial	32
2.3.1	Variables e Indicadores	32
2.3.1.1	Independiente: Análisis de Consumo de Agua	32
2.3.2	Inteligencia de Negocios	33
2.3.2.1	Aplicación de Inteligencia de Negocios para análisis de consumo de agua potable	34
2.3.2.2	Balance ScoreCard	35
2.3.2.3	Evolución de las Plataformas BI	36
2.3.2.4	Desarrollo de bases de datos y aplicaciones empresariales (SAP, JD Edwards, Siebel, PeopleSoft)	37
2.3.2.5	Componentes y Arquitectura del BI	38
2.3.3	Metodologías, estrategias y análisis	39
2.3.3.1	Metodología Kimbal	39
2.3.4	Proceso de Data Analysis	40
2.3.4.1	Fuentes de Datos	41
2.3.5	Proceso de Extracción, Transformación y Carga ETL	42
2.3.6	Almacenes de Datos (Data warehouse DWH)	42
2.3.7	Modelo Multidimensional	44
2.3.8	On-Line-Analytical-Processing (OLAP)	45
2.3.9	Aplicación de BI en empresas privadas en Ecuador	46
2.3.10	Aplicación de BI en el sector público en el Ecuador	47
2.3.11	Base de Datos y Herramientas de desarrollo	48
2.3.11.1	Base de Datos Oracle 12c	48
2.3.11.2	Toad for Oracle 12.11	48

2.3.11.3	Oracle Business Intelligence.....	49
2.3.11.4	ArcGis.....	49
CAPITULO III		51
3	Introducción	51
3.1	Métodos y Materiales	51
3.1.1	Método Inductivo.....	51
3.1.2	Método Analítico – Sintético	51
3.1.3	Estrategias Técnicas.....	52
3.1.4	Caso y unidad de análisis	52
3.1.5	Recolección de Datos.....	53
3.2	Modalidad de la Investigación	53
3.2.1	Investigación de Campo.....	53
3.2.2	Investigación Documental.....	54
3.2.3	Investigación Bibliográfica.....	54
3.3	Tipo o alcance de la investigación	55
3.3.1	Investigación Exploratorio.....	55
3.3.2	Muestra o población.....	55
3.4	Instituciones participantes en este proyecto	55
3.4.1	Empresa Municipal de Agua Potable y Alcantarillado Ibarra	56
3.4.1.1	Datos generales.....	56
3.4.1.2	Misión.....	56
3.4.1.3	Visión.....	56
3.4.1.4	Integrantes del directorio	56
3.4.2	Gobierno Autónomo Descentralizado Ibarra	57
3.4.2.1	Datos generales	57
3.4.2.2	Misión.....	58
3.4.2.3	Visión.....	58
3.4.2.4	Autoridades.....	58

3.4.3 Instituto Nacional De Estadística Y Censos Ibarra.....	59
3.4.3.1 Datos generales.....	59
3.4.3.2 Misión.....	59
3.4.3.3 Visión.....	60
3.4.3.4 Organigrama.....	60
3.5 El sistema.....	60
3.5.1 Problema del sistema actual.....	61
3.5.2 Delimitación del alcance a desarrollar.....	61
3.5.3 Involucrados.....	61
3.5.4 Equipo técnico responsable.....	61
3.5.5 Factibilidad Técnica.....	62
3.5.6 Factibilidad Operativa.....	62
3.6 Desarrollo de la propuesta.....	63
3.6.1 Selección de metodología para el desarrollo de la propuesta.....	63
3.6.2 Desarrollo de la Herramienta BI.....	63
3.6.3 Arquitectura de Análisis de Datos con aplicación técnica de Business Intelligence.....	63
3.6.4 Análisis de datos.....	64
3.6.5 Procesamiento de datos.....	64
3.7 Implementación del Business Intelligence.....	65
3.7.1 Planeación del proyecto.....	65
3.7.2 Capa de datos.....	66
3.7.2.1 Datos EMAPA.....	66
3.7.2.2 Datos INEC.....	70
3.7.2.3 Datos Gobierno Autónomo Descentralizado San Miguel de Ibarra (Catastro Urbano).....	74
3.7.3 Definición de los requerimientos del negocio.....	75
3.7.3.1 Requerimientos EMAPA.....	76

3.7.3.2	Requerimientos Proyecto Ibarra Verde	77
3.7.3.3	Perspectiva de la herramienta	77
3.7.3.4	Descripción de los requerimientos.....	78
3.7.4	Capa ETL	82
3.7.5	Modelado Dimensional	84
3.7.6	Modelo físico	93
3.7.7	Implementación de la solución BI	95
3.7.8	Especificación de aplicaciones BI.....	96
3.7.9	Mantenimiento y Crecimiento del DW	99
CAPITULO IV		101
4	Introducción	101
4.1	Presentación de Resultados	101
4.1.1	Aplicación de Business Intelligence en el Análisis de Consumo y Distribución de Agua Potable.....	102
4.1.2	Definición de la muestra de estudio.....	103
4.1.2.1	Densidad Poblacional	106
4.1.2.1	Tamaño de vivienda.....	111
4.1.3	Resultados Generales de la Investigación.....	114
4.1.4	Desarrollo de Hipótesis.....	114
4.1.4.1	¿Cómo se realiza actualmente el análisis de consumo de agua potable? 115	
4.1.4.2	¿Qué tipo de análisis requiere en cuestión del consumo de agua potable? 115	
4.1.4.3	¿Cuáles son los parámetros o estándares para tomar decisiones? . 116	
4.2	Conclusiones	117
4.3	Recomendaciones.....	119
5	Referencias.....	121

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Categorías Fundamentales.....	32
Ilustración 2: Business Intelligence.....	33
Ilustración 3: Evolución Business Intelligence	37
Ilustración 4: Arquitectura Business Intelligence.....	38
Ilustración 5: Ciclo de Vida Kimball.....	39
Ilustración 6: Procesamiento de Información.....	41
Ilustración 7: Fuentes de Datos	41
Ilustración 8: Esquema ETL.....	42
Ilustración 9: Data warehouse	43
Ilustración 10: Multidimensional.....	45
Ilustración 11: Organigrama INEC.....	60
Ilustración 12: Arquitectura propuesta	64
Ilustración 13: Datos EMAPA.....	66
Ilustración 14: Campos de la Tabla Suministros	67
Ilustración 15: Campos de la Tabla Clientes	67
Ilustración 16: Campos de la Tabla Consumos	68
Ilustración 17: Campos da la Tabla Tarifas.....	68
Ilustración 18: Campos de la Tabla Producción.....	68
Ilustración 19: Campos de la Tabla Redes Comerciales	69
Ilustración 20: Diagrama Entidad Relación EMAPA.....	69
Ilustración 21: Datos INEC	70
Ilustración 22: Campos tabla Inec_Cantones	70
Ilustración 23: Campos Tabla Inec_Censos	71
Ilustración 24: Campos Tabla Inec_Habitantes_Sectores	71
Ilustración 25: Campos Tabla Inec_Parroquias.....	71
Ilustración 26: Campos Tabla Inec_Provincias	72
Ilustración 27: Campos Tabla Inec_Sectores	72
Ilustración 28: Campos Tabla Inec_Total	73
Ilustración 29: Campos Tabla Inec_Zonas	73
Ilustración 30: Diagrama Entidad Relación INEC	74
Ilustración 31: Datos Catastro Urbano	75
Ilustración 32: Esquema conceptual consumo agua potable	83

Ilustración 33: Esquema conceptual generación agua potable	83
Ilustración 34: Dimensión Tiempo	85
Ilustración 35: Dimensión Tarifa.....	86
Ilustración 36: Dimensión Redes.....	87
Ilustración 37: Dimensión Geo INEC	88
Ilustración 38: Dimensión Clientes	89
Ilustración 39: Dimensión Redes Producción	90
Ilustración 40: Tabla de hechos Consumo.....	91
Ilustración 41: Tabla de hechos Generación	92
Ilustración 42: Modelo relacional.....	93
Ilustración 43: Herramienta de Administración Oracle BI.....	94
Ilustración 44: Modelo estrella Consumo.....	95
Ilustración 45: Modelo estrella Generación.....	96
Ilustración 46: Análisis de pérdidas.....	97
Ilustración 47: Análisis por rango de consumo	98
Ilustración 48: Análisis de Densidad Poblacional	99
Ilustración 49: Datos Análisis Densidad Poblacional.....	107
Ilustración 50: Grafico Análisis Densidad Poblacional Zonas.....	107
Ilustración 51: Análisis Densidad Poblacional Sector.....	108
Ilustración 52: Densidad poblacional habitante por hectárea (hab/ha) con el consumo de agua residencial M3/mes por zona censal.	109
Ilustración 53: Zonificación de la Densidad Poblacional (hab/ha) con el consumo promedio de agua residencial m3/mes por zona censal amanzanada, parroquia urbana de la ciudad de Ibarra.	110
Ilustración 54: Tamaño de vivienda (M2) y el consumo residencial M3/mes	111
Ilustración 55: Distribución espacial tamaño de vivienda (M2) y consumo residencial M3/mes.....	112

ÍNDICE DE TABLAS

Tabla 1: Matriz de Tensiones y Problemas de la Zona 1. (2013-2017).....	27
Tabla 2: Constitución de la República del Ecuador	28
Tabla 3: Plan Nacional del Buen Vivir.....	29
Tabla 4: Reglamento de Régimen Académico	31
Tabla 5: Datos informativos EMAPA	56
Tabla 6: Miembros directorio EMAPA.....	57
Tabla 7: Datos Municipio Ibarra	57
Tabla 8: Autoridades Gobierno Autónomo Descentralizado Municipal San Miguel de Ibarra	58
Tabla 9: Datos INEC	59
Tabla 10: Beneficiarios.....	61
Tabla 11: Equipo técnico responsable	62
Tabla 12: Requerimientos EMAPA.....	76
Tabla 13: Requerimientos Proyecto Ibarra Verde	77
Tabla 14: Análisis de Perdidas	78
Tabla 15: Análisis de Consumo por Tarifa.....	79
Tabla 16: Análisis por rango de consumo a nivel de suministro.....	79
Tabla 17: Análisis de historial de consumo.....	80
Tabla 18: Análisis de consumo por planes.	80
Tabla 19: Unión de datos.....	81
Tabla 20: Análisis de consumo por sector censal.....	81
Tabla 21: Análisis de consumo por persona a nivel de zona censal.....	81
Tabla 22: Obtener la base de suministros considerando muestra de estudio	82
Tabla 23: Formula de proyección poblacional	102
Tabla 24: Sectores INEC	103
Tabla 25: Nivel de significación del coeficiente de correlación R relación hab/ha vs M3/mes	109
Tabla 26: Significancia de los coeficientes de correlación del Libro Aplicación de estimadores estadísticos y diseño experimentales en investigaciones forestales.	113

RESUMEN

En las instituciones públicas del país, el análisis de información es parte de las estrategias básicas para mejorar sus procesos con el objetivo de mejorar la entrega de sus productos y servicios. En este aspecto la data generada día a día por los sistemas transaccionales se ha convertido en un activo importante para cada una de las instituciones, por lo tanto, el desarrollo e implementación de sistemas para la toma de decisiones son más frecuentes. La presente investigación, pretende adaptar el concepto de integridad de datos de varias instituciones y realizar una base de datos única que ayude a la toma de decisiones en varios ejes, con el objetivo de mejorar la distribución de agua potable tomando en cuenta factores de crecimiento poblacional, área de catastros, niveles de educación etc. Para llevar a cabo esta investigación se la realizo con un enfoque cualitativo descriptivo, en la cual se basa en revisiones bibliográficas de aplicación de análisis de datos, para la integración de datos se realizó entrevistas con los usuarios finales de la herramienta (stakeholder), con la finalidad de presentar una herramienta que se adapte a los requerimientos de cada uno de ellos y mejorar la toma de decisiones. La implementación de este tipo de herramientas permitió conocer los patrones de crecimiento poblacional, así como las necesidades en cuanto a la distribución de agua potable.

Palabras Claves: Análisis de Datos, Sistema Integrado, Inteligencia de Negocios, Almacenes de Datos, Toma de Decisiones.

ABSTRACT

In the public institutions of the country, the analysis of information is part of the basic strategies to improve processes with the aim of improving the delivery of their products and services. In this aspect, transactional data have become an important element for each of the institutions, for the development and implementation of systems for making more frequent decisions. The present investigation intends to adapt the concept of the integrity of the data and to realize a unique database that helps to make decisions in several axes, with the objective of improving the distribution of drinking water taking into account factors of population growth, cadastral area, education levels, etc. In order to carry out this research, the analysis is carried out with a qualitative descriptive approach, based on the bibliographical references of the data analysis application. For the integration of the data, the interviews with the final users of the tool were carried out. (stakeholder), in order to present a tool that adapts to the requirements of each of them and improves decision making. The implementation of this type of tools will allow to know the patterns of population growth, as well as the needs regarding the distribution of drinking water.

Keywords: Data Analysis, Integrated System, Business Intelligence, Data Warehouses, Decision Making.

CAPITULO I

PROBLEMA DE LA INVESTIGACION

1 Introducción

Esta investigación es parte del macro proyecto “Ibarra Verde”, proyecto que tiene el objetivo principal de crear una herramienta de análisis, basado en un mapeo integral de la zona urbana de la ciudad de Ibarra, en los Andes Norte del Ecuador, que sirva como modelo para una planificación urbana técnicamente alta, y que, a la vez, respete el sentido cultural, las normas sociales y el estado ambiental de la región. (Frolich, y otros, 2017)

En colaboración con empresa de agua potable (EMAPA-I—Empresa Municipal de Agua Potable y Alcantarillado—Ibarra), se obtuvo datos del uso de agua potable, de cada usuario doméstico igual que industrial, sobre los últimos tres años, debidamente digitalizados. Como parte del análisis de datos, estamos incorporando la ubicación geográfica, con coordenadas precisas de cada usuario para poder analizar la distribución de uso por sectores geográficos y tendencias históricas del uso de agua y energía eléctrica en un mapeo integrado e interactivo. (Frolich, y otros, 2017)

Se analizó las tendencias temporales y la distribución geográfica del uso de este recurso, los datos se unificaron en una base de Oracle para aplicar las herramientas de Oracle BI (Business Intelligence) una poderosa herramienta basada en tecnología Web con interfaz gráfica amigable para los usuarios y que permitió reunir todos los datos y transformarlos en información que sirve para toma de decisiones. (Frolich, y otros, 2017)

EMAPA-I, es una empresa que a través de una administración eficiente, suministra servicios de agua potable y saneamiento con calidad, para mejorar las condiciones de vida de los habitantes del cantón, enmarcados en valores, principios y normativas vigentes. (emapa, 2017)

1.1. Antecedentes

El avance en el área de análisis de datos y el desarrollo de tecnologías de la información y comunicaciones (TIC'S), han dado lugar al desarrollo de la era del conocimiento en todas las áreas y ámbitos posibles, esta está siendo aplicada en varias instituciones públicas como privadas. El desarrollo de aplicaciones relacionadas con la generación de conocimiento a partir del análisis datos generados en las diversas instituciones cada día va en aumento para estar a la par de las instituciones pioneras en la utilización de estas herramientas.

La disponibilidad del agua potable es característica fundamental de una zona urbana bien desarrollada. A la vez, su ausencia tiende a definir zonas marginales y barrios subdesarrollados. En el siglo XXI, se está aplicando herramientas de análisis digital, a través de sistemas de computación de alto poder, para realizar una planificación integral que incorpore múltiples elementos que afecte el suministro y la demanda de este recurso, así formando parte de lo que se denomina una urbe inteligente. (Frolich, y otros, 2017)

El reto de crear herramientas digitales de alta calidad para analizar el uso de recursos en la zona urbana es una respuesta lógica al crecimiento de la ciudad, tal fenómeno que define la evolución del ser humano, como especie global, a tal punto que se ha denominado una nueva especie Homo evolutis , en nuestros tiempos. Durante el siglo XX, la mitad de la población humana ha migrado a las áreas urbanas, principalmente en el hemisferio norte. La Organización de las Naciones Unidas (ONU) ha determinado que en el año 2013 se llegó al punto de inflexión en que la mayoría de la población mundial vive en ciudades (UN-DESA, 2014). Las ventajas del ámbito urbano son numerosas, no solamente por el deseo del vibrante distrito central, sea cultural, bancario o industrial, sino incluso por las economías de escala, facilidad de provisionar necesidades vitales, la concentración de diversidad humana, la posibilidad de fomentar innovación y fácil acceso a nuevas tecnologías. A la vez, la migración al ámbito urbano presenta el riesgo de una disminución de la calidad de vida, hasta niveles de miseria entre el segmento pobre de la población. Este proceso de degradación en las áreas urbanas ocurre cuando se pierde el acoplamiento entre el ser humano y la realidad biogeográfica de la zona. (Frolich, y otros, 2017)

En el caso de la ciudad de Ibarra-Ecuador, encontramos una situación ideal para la aplicación de una herramienta BI, ya que en la actualidad se tiene mucha información generada por el consumo de agua potable que nos ayudaría a guiar el desarrollo de las zonas del cantón Ibarra. En las últimas dos décadas, el nivel técnico de planificación de la empresa municipal de agua potable ha aumentado al punto de tener disponibles cantidades de datos digitales sobre múltiples aspectos de la vida urbana, sobre todo del uso de agua potable. Pero no se encuentra todavía una centralización de análisis para guiar los técnicos y público en general, la inevitable expansión de la zona urbana sobre el transcurso de este siglo. Entonces existe una oportunidad grande de formar una herramienta interactiva. (Frolich, y otros, 2017)

En el caso del uso de agua, el análisis pretende incorporar una vista hacia la totalidad del uso y su relación con el diseño de la ciudad y la situación de la población de cada zona. Hay una tendencia, de procedencia del hemisferio Norte, pero con reconocimiento de los tiempos ancestrales, a incorporar áreas de jardín, de bosque, de pajonal, y de agricultura dentro de la ciudad, todo relacionado, por supuesto, con la distribución del agua y el sistema cultural de los residentes. Es importante, entonces, poder comparar, interactivamente, la situación demográfica, económica, social y ambiental de cada zona para incorporar dentro de las sugerencias de planificación en referencia a la distribución y suministro del recurso agua.

1.2. Planteamiento del Problema

Se ha realizado un previo análisis general de todas las instituciones que intervienen en esta investigación, descubriendo que cada una maneja su información aisladamente y por lo tanto no pueden realizar análisis de información conjunta para toma de decisiones.

Al tener varios sistemas y bases de datos transaccionales en cada una de las instituciones, que generan datos importantes con el objetivo de cumplir y ejecutar procesos en cada una de estas.

El no tener una infraestructura lógica y física donde se puedan consolidar y analizar la información generada por las diversas entidades a través de una herramienta de Business Intelligence (BI), se dificulta la presentación de información depurada, consolidada y sobre todo de fácil visualización, lo que provoca realizar trabajos extensos para realizar análisis propios de la empresa y más aún cuando los análisis son dependientes de la información proveniente de otras instituciones.

Al ser el agua potable un recurso natural no renovable, necesita ser tratado de una manera especial en cuanto a su distribución, consumo y desaprovechamiento. Por lo cual si no se implementan acciones que ayuden a optimizar el consumo y distribución, sobre todo a disminuir el desaprovechamiento, tendremos problemas de abastecimiento en un futuro, afectando el estilo de vida de la población del cantón Ibarra que va en aumento día a día.

Al tener gran y valiosa información generada por las diversas instituciones se podrían generar análisis los cuales permitan unir diversos índices que permitan mejorar los niveles de consumo y sobre todo de desperdicio de agua potable.

1.3. Formulación del problema

¿La aplicación de Business Intelligence en el análisis del consumo de agua potable, contribuirá con la toma de decisiones enfocadas a mejorar la calidad de vida de los habitantes del cantón Ibarra, en cuanto al mejoramiento de la distribución y consumo de agua potable?

1.4. Justificación de la investigación

El aumento de la población en los cascos urbanos de las ciudades es cada vez mayor, por ende el consumo de agua potable va en aumento día a día, al no tener una herramienta que permita conocer cuáles son los sectores que van aumentando la demanda de este recurso natural sobre el tiempo, limita el poder tomar decisiones para emprender campañas y acciones que permita reducir y optimizar el consumo y a la vez ayude a mejorar la distribución de agua potable.

Las herramientas de toma de decisiones en la actualidad son muy importantes en las grandes empresas, ya que estas ayudan a consolidar información, en la actualidad EMAPA posee una gran cantidad de información la cual puede ser analizada con el fin de mejorar los índices de consumos, distribución y pérdidas de agua potable.

Con la ayuda de las herramientas informáticas facilitan realizar cruces de información de varias entidades, en este caso de EMAPA, INEC y Catastral, con lo cual obtenemos no solo análisis a nivel de consumos y distribución de agua potable, sino también unirlo con información como niveles de educación, personas económicamente activas, población total, extensión del área, área de construcción, etc.

1.5. Objetivos de la investigación

1.5.1. Objetivo General

Desarrollar una herramienta Business Intelligence para análisis del consumo de agua del cantón en Ibarra, analizando los datos existentes para ayudar en la toma de decisiones.

1.5.2. Objetivos Específicos

- Investigar técnicas de Business Intelligence que beneficien al análisis de información relacionada al consumo de agua potable.
- Analizar los datos de consumo de agua potable en el cantón Ibarra para obtener información válida y procesada.
- Implantar una herramienta de Business Intelligence que se adapte a los requerimientos de la Empresa Municipal de Agua Potable y Alcantarillado de Ibarra y Proyecto Ibarra Verde.
- Presentar los resultados del análisis de la información en forma interactiva para tomar decisiones fácilmente.

1.6. Hipótesis o preguntas de investigación

¿El análisis de información del consumo de agua potable mediante herramientas Business Intelligence mejorara el abastecimiento y contribuirá el mejoramiento de vida de los habitantes del cantón Ibarra?

1.6.1. Preguntas directrices

- ¿Cómo se realiza actualmente el análisis de consumo del agua potable?
- ¿Qué tipo de análisis requiere en cuestión del consumo de agua potable?
- ¿Cuáles son los parámetros o estándares para tomar decisiones?

CAPITULO II

MARCO TEORICO

2 Introducción

En este capítulo analizamos soluciones con implementación de BI, realizando una fundamentación teórica a los principales componentes utilizados en estas investigaciones, se analizaron las diferentes metodologías utilizadas para la creación de almacenes de datos, desde el momento de la extracción, transformación y carga de la información (ETL), principios fundamentales del desarrollo de soluciones BI.

2.1 Estado del Arte

Se realizó visita a varias bibliotecas virtuales y repositorios de universidades donde se encontró varias investigaciones relacionadas, las cuales citamos a continuación.

En el país se realizó la investigación del diseño y aplicación de un data Warehouse para empresas de distribución y comercialización de agua potable, en la cual hace énfasis en que la aplicación de BI es recomendado ya que nos ayuda en la transferencia de conocimiento a partir de los datos generados por los sistemas transaccionales mediante tableros de mando o control, en los cuales se obtienen información valiosa en cuanto al estado de la empresa y poder identificar factores que nos ayude a mejorar la toma de decisiones, para poder aprovechar al máximo los beneficios de la aplicación de BI se debe conocer el modelo de negocio donde se aplica, como resultados se pudo identificar los factores por los cuales los consumos son elevados e ciertas zonas y también se pudo reducir el índice de desaprovechamiento de agua. (Lascano Sumbana, 2016)

El uso de herramientas BI en empresas públicas y privadas en el país se caracteriza por el tamaño y enfoque del negocio de cada una de estas, ya que en la actualidad solo empresas privadas consideradas grandes han podido consolidar el uso de herramientas BI, en el caso de entidades públicas el factor determinante para la utilización de estas herramientas es el ámbito en el cual se desempeña. En ambos casos el factor costo beneficio ha sido determinante en la

aplicación ya que para una correcta aplicación se necesita recursos tanto económicos para la adquisición o arrendamiento de infraestructura, así como personal debidamente capacitado para el desarrollo y uso de estos sistemas. Entre las empresas privadas que han optado por la utilización de sistemas BI se encuentran las dedicadas a la banca y expendio de productos y en el sector privado las entidades pioneras son las consideradas de sectores estratégicos como son eléctricas, petroleras, etc.

La aplicación de sistemas BI en cuanto a la distribución de agua potable no es nuevo en el mundo, ya que en la ciudad de México se realizó la implementación de un sistema BI para mejorar la distribución de agua potable, en la cual han llegado a ahorrar hasta un 10% es decir aproximadamente unos 11 millones de litros de agua por hora, tomando en cuenta que tiene alrededor de 8.8 millones de habitantes. La aplicación de BI se lo realizó en diversas áreas del negocio, lo que ha permitido ayudar en la racionalización, análisis de consumo, clientes morosos y detección de robo de agua potable. (Cota , 2011)

En la actualidad se está manejando la teoría de creación de ciudades inteligentes, ciudades que tienen el objetivo de proyectar un crecimiento sustentable en cuanto al aprovisionamiento de las necesidades básicas y crecimiento controlado. Entre los beneficios directos a esta investigación es la de implantación de redes de distribución inteligentes las cuales nos permiten detectar en tiempo real averías, atascos o fugas de agua potable. Son proyectos pilotos que se los están realizando en países como Alemania y España.

2.2 Marco Legal

Para el desarrollo de esta investigación nos basamos en varios campos legales, los cuales se describen a continuación.

2.2.1 Matriz de Tensiones y Problemas

Para identificar fácilmente cuales son los problemas de cada una de las zonas se tiene como base la MATRIZ DE TENSIONES Y PROBLEMAS (PNBV 2013-2017), en la cual se detalla el contexto, los ejes, las tensiones y problemas de cada una de estas, en estés caso específico nos centraremos en el análisis de la matriz perteneciente a la zona 1, en la cual va a ser aplicada la investigación.

Vamos hacer referencia a dos problemas que se encuentran enunciados dentro de esta matriz, los cuales son:

Tabla 1: Matriz de Tensiones y Problemas de la Zona 1. (2013-2017)

Contexto	Eje	Tensiones y Problemas
HÁBITAT SUSTENTABLE	AMBIENTE	Gestión integral de cuencas hidrográficas: cultura que incentive el ahorro y el uso racional del agua.
	SERVICIOS BÁSICOS	Ampliar la cobertura de servicios básicos: agua potable y saneamiento ambiental en sectores urbanos y rurales.

Fuente: Elaboración Propia

Nota: Se toma artículos que interesa en la investigación.

2.2.2 Constitución del Ecuador 2008

En la constitución vigente actualmente aprobada en Montecristi en el año 2008, hace referencia el uso de agua y en la cual privilegia el uso de esta para el consumo humano y riego, también en varios artículos detalla los derechos y obligaciones en cuanto al acceso del líquido vital, para lo cual detallamos los artículos que van en concordancia a la investigación realizada.

CONSTITUCIÓN

Art. 12.- El derecho humano al agua es fundamental e irrenunciable. El agua constituye patrimonio nacional estratégico de uso público, inalienable, imprescriptible, inembargable y esencial para la vida.

Art. 32.- La salud es un derecho que garantiza el Estado, cuya realización se vincula al ejercicio de otros derechos, entre ellos el derecho al agua, la alimentación, la educación, la cultura física, el trabajo, la seguridad social, los ambientes sanos y otros que sustentan el buen vivir.

Art. 66.- Se reconoce y garantizará a las personas:

2. El derecho a una vida digna, que asegure la salud, alimentación y nutrición, agua potable, vivienda, saneamiento ambiental, educación, trabajo, empleo, descanso y ocio, cultura física, vestido, seguridad social y otros servicios sociales necesarios.

Art. 264.- Los gobiernos municipales tendrán las siguientes competencias exclusivas sin perjuicio de otras que determine la ley:

1. Planificar el desarrollo cantonal y formular los correspondientes planes de ordenamiento territorial, de manera articulada con la planificación nacional, regional, provincial y parroquial, con el fin de regular el uso y la ocupación del suelo urbano y rural.

4. Prestar los servicios públicos de agua potable, alcantarillado, depuración de aguas residuales, manejo de desechos sólidos, actividades de saneamiento ambiental y aquellos que establezca la ley.

Art. 318.- El agua es patrimonio nacional estratégico de uso público, dominio inalienable e imprescriptible del Estado, y constituye un elemento vital para la naturaleza y para la existencia de los seres humanos. Se prohíbe toda forma de privatización del agua.

Art. 375.- El Estado, en todos sus niveles de gobierno, garantizará el derecho al hábitat y a la vivienda digna, para lo cual:"

"6. Garantizará la dotación ininterrumpida de los servicios públicos de agua potable y electricidad a las escuelas y hospitales públicos.

Art. 411.- El Estado garantizará la conservación, recuperación y manejo integral de los recursos hídricos, cuencas hidrográficas y caudales ecológicos

asociados al ciclo hidrológico. Se regulará toda actividad que pueda afectar la calidad y cantidad de agua, y el equilibrio de los ecosistemas, en especial en las fuentes y zonas de recarga de agua. La sustentabilidad de los ecosistemas y el consumo humano serán prioritarios en el uso y aprovechamiento del agua.

Art. 412.- La autoridad a cargo de la gestión del agua será responsable de su planificación, regulación y control. Esta autoridad cooperará y se coordinará con la que tenga a su cargo la gestión ambiental para garantizar el manejo del agua con un enfoque ecosistémico.

Art. 415.- El Estado central y los gobiernos autónomos descentralizados adoptarán políticas integrales y participativas de ordenamiento territorial urbano y de uso del suelo, que permitan regular el crecimiento urbano, el manejo de la fauna urbana e incentiven el establecimiento de zonas verdes. Los gobiernos autónomos descentralizados desarrollarán programas de uso racional del agua, y de reducción reciclaje y tratamiento adecuado de desechos sólidos y líquidos. Se incentivará y facilitará el transporte terrestre no motorizado, en especial mediante el establecimiento de ciclo vías.

Fuente: Elaboración Propia

Nota: Se toma artículos que interesa en la investigación.

2.2.3 El Plan Nacional para el Buen Vivir

Como política de estado se ha generado el Plan del Buen Vivir, en el cual se plasman orientaciones para planificar el desarrollo nacional, erradicar la pobreza, promover el desarrollo sustentable y la redistribución equitativa de los recursos y la riqueza para alcanzar un buen vivir, dentro de este plan se han seleccionado varios objetivos que están acorde a la investigación.

Tabla 3: Plan Nacional del Buen Vivir

<i>Plan Nacional</i>		
<i>OBJETIVO</i>	<i>POLITICA</i>	<i>LITERAL</i>
4.	4.3	4.3 a
Fortalecer las capacidades y potencialidades	Promover espacios formales y educación	Democratizar el acceso al conocimiento, no fortaleciendo los acervos de datos, la información científica y los saberes diversos en todos sus formatos, desde

de ciudadanía la permanente para espacios físicos y virtuales de libre el intercambio de acceso, reproducción y circulación en conocimientos y red, que favorezcan el aprendizaje y el saberes para la intercambio de conocimientos. sociedad aprendiente.

4.4

4.4.i

Mejorar la calidad de la educación en todos sus niveles y modalidades, para la generación de conocimiento y la formación integral de personas creativas, solidarias, responsables, críticas, participativas y productivas, bajo los principios de igualdad, equidad social y territorialidad

Asegurar en los programas educativos la inclusión de contenidos y actividades didácticas e informativas que motiven el interés por las ciencias, las tecnologías y la investigación, para la construcción de una sociedad socialista del conocimiento.

4.4.n

Diseñar e implementar herramientas e instrumentos que permitan el desarrollo cognitivo-holístico de la población estudiantil.

4.9	4.9.k.
Impulsar la formación en áreas de conocimiento no tradicionales que aportan a la construcción del Buen Vivir	Generar programas con contenidos educativos que fortalezcan el proceso de aprendizaje escolar, la alfabetización digital y el aprendizaje para la vida, a través de las TIC.

Fuente: Elaboración Propia

Nota: Se toma artículos que interesa en la investigación.

2.2.4 Reglamento de Régimen Académico

En el reglamento de régimen académico se enuncia que las instituciones de educación superior deben contribuir en la investigación científica, social y tecnológica para solucionar problemas del país, así como formar a los estudiantes para adaptarse al entorno de desarrollo científico, tecnológico y económico, con el fin de constituirse en agente efectivo del cambio y desarrollo social para aprovechar las oportunidades que le ofrece la sociedad en donde va a desenvolver su vida y ejercicio profesional, por lo cual para la elaboración de este trabajo se citan los artículos que avalan la realización de esta investigación:

Tabla 4: Reglamento de Régimen Académico

Reglamento de Régimen Académico
Artículo 3. Literal d. articular la formación académica y profesional, la investigación científica, tecnológica y social, y la vinculación con la colectividad en un marco de calidad, innovación y pertinencia.
Literal h. Impulsar el conocimiento de carácter multi, inter y trans disciplinarios en la formación de grado y postgrado, la investigación y la vinculación con la colectividad.
Literal j. Desarrollar la educación superior bajo la perspectiva del bien público social, aportando a la democratización del conocimiento para la garantía de derechos y la reducción de inequidades.

Fuente: Elaboración Propia

Nota: Se toma artículos que interesa en la investigación

2.3 Marco Referencial

Para la interpretación del trabajo realizado en cuanto al análisis de variables independientes y dependientes se realizó el diagrama que especifica los temas a desarrollar como se muestra en la Ilustración 1.

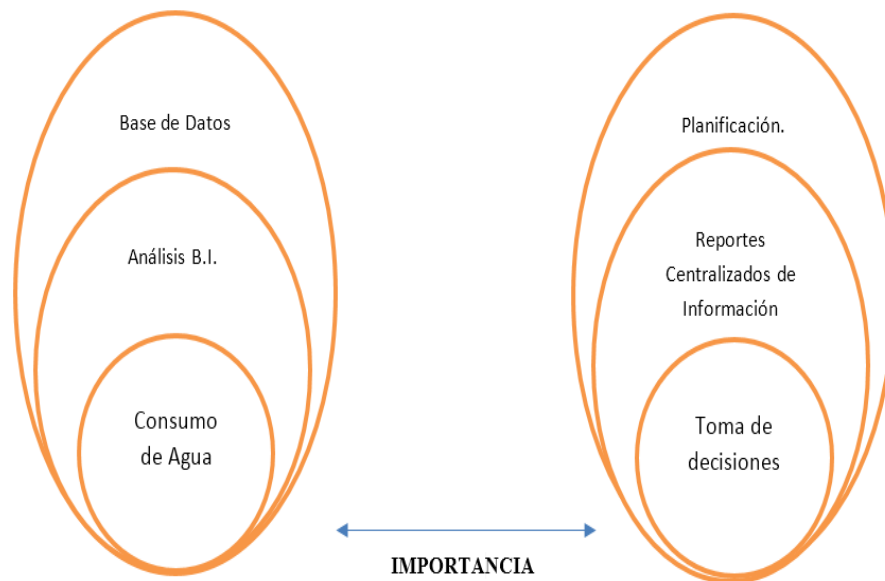


Ilustración 1: Categorías Fundamentales

Fuente: Elaboración Propia

Los temas a desarrollar en el Marco Teórico del Informe Final serán:

- Base de Datos
- Análisis Business Intelligence
- Consumo de Agua
- Reportes Centralizados de Información
- Sistema Integrado para la Toma de Decisiones

2.3.1 Variables e Indicadores

2.3.1.1 Independiente: Análisis de Consumo de Agua.

Indicadores:

- Consumo en m3
- Tarifas de consumo
- Cruce de información con variables de otras instituciones
- Estándares de consumo por tipos de tarifas

Dependiente: Toma de decisiones

Indicadores:

- Procesos Centralizados
- Reporte interactivos
- Sistema BI

2.3.2 Inteligencia de Negocios

Business Intelligence es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios.



Ilustración 2: Business Intelligence

Fuente: Internet

Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, podemos definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la compañía) en información estructurada, para su explotación directa (reporting, análisis OLTP / OLAP, alertas...) o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio.

La inteligencia de negocio actúa como un factor estratégico para una empresa u organización, generando una potencial ventaja competitiva, que no es otra que proporcionar información privilegiada para responder a los problemas de negocio: entrada a nuevos mercados, promociones u ofertas de productos, eliminación de islas de información, control financiero, optimización de costes, planificación de producción, análisis de perfiles de clientes, rentabilidad de un producto concreto, etc.

Los principales productos de Business Intelligence que existen hoy en día son:

- Cuadros de Mando Integrales (CMI)
- Sistemas de Soporte a la Decisión (DSS)
- Sistemas de Información Ejecutiva (EIS)

Por otro lado, los principales componentes de orígenes de datos en el Business Intelligence que existen en la actualidad son:

- Datamart
- Data warehouse

Los sistemas y componentes del BI se diferencian de los sistemas operacionales en que están optimizados para preguntar y divulgar sobre datos. Esto significa típicamente que, en un data warehouse, los datos están desnormalizados para apoyar consultas de alto rendimiento, mientras que en los sistemas operacionales suelen encontrarse normalizados para apoyar operaciones continuas de inserción, modificación y borrado de datos. En este sentido, los procesos ETL (extracción, transformación y carga), que nutren los sistemas BI, tienen que traducir de uno o varios sistemas operacionales normalizados e independientes a un único sistema des normalizado, cuyos datos estén completamente integrados.

2.3.2.1 Aplicación de Inteligencia de Negocios para análisis de consumo de agua potable

La aplicación de políticas en cuanto al consumo y distribución del agua potable es imprescindible para aprovechar al máximo el recurso y evitar el

desabastecimiento. La importancia del agua potable en la actualidad se da debido al desabastecimiento en algunas épocas del año por factores ambientales y por el crecimiento de las urbes los que significa un aumento de necesidad de agua en diversos sectores.

Con la utilización de la inteligencia de negocios se analiza la información para identificar consumos anómalos que no cumplan con los patrones establecidos como litros diarios por persona, consumos máximos mensuales por tarifas, márgenes de pérdidas, etc.

Para el análisis de información se realizó utilizando técnicas de recopilación de datos de diversas instituciones que nos faciliten comprender las tendencias de consumo.

Con la información recopilada se procedió a aplicar la técnica OLAP, propia de la inteligencia de negocio, que permite realizar la construcción de los cubos multidimensionales, de los cuales obtenemos los elementos de análisis o facts como son máximos consumo, porcentajes de pérdidas, crecimiento por zonas, y sectores.

Para complementar la investigación se aplicó la técnica de Data Mining, propia de la inteligencia de negocios, la cual nos permitió identificar patrones y correlaciones en cada uno de los análisis realizados.

2.3.2.2 Balance ScoreCard

Es una herramienta de gestión que traduce la estrategia de la empresa en un conjunto coherente de indicadores

La visión y la estrategia de negocios dictan el camino hacia el que deben encaminarse los esfuerzos individuales y colectivos de una empresa. La definición de estrategias por naturaleza es complicada, pero la implementación de la misma representa el mayor obstáculo en la mayoría de las ocasiones. El reto corresponde en identificar exactamente lo que debe monitorearse, para comunicar en todos los niveles de la empresa, si se están alcanzando las estrategias a través de acciones muy puntuales.

Balanced Scorecard es la principal herramienta metodológica que traduce la estrategia en un conjunto de medidas de la actuación, las cuales proporcionan la estructura necesaria para un sistema de gestión y medición.

El acceso a los principales almacenes de información brinda la posibilidad de presentar los resultados de desempeño y entender por qué están dándose esos resultados.

El Balanced Scorecard induce una serie de resultados que favorecen la administración de la compañía, pero para lograrlo es necesario implementar la metodología y la aplicación para monitorear, y analizar los indicadores obtenidos del análisis. Entre otros podemos considerar las siguientes ventajas:

- Alineación de los empleados hacia la visión de la empresa.
- Comunicación hacia todo el personal de los objetivos y su cumplimiento.
- Redefinición de la estrategia en base a resultados.
- Traducción de la visión y estrategias en acción.
- Favorece en el presente la creación de valor futuro.
- Integración de información de diversas áreas de negocio.
- Capacidad de análisis.
- Mejoría en los indicadores financieros.
- Desarrollo laboral de los promotores del proyecto.

2.3.2.3 Evolución de las Plataformas BI

El término Business Intelligence (Inteligencia de negocios), se lo viene manejando desde hace tiempo atrás como se lo detalla en la Ilustración 3, con el objetivo de que los departamentos de sistemas les provean de información en reportes estándares, personalizados y facilitar la toma de decisiones; el objetivo principal es la transformación de gran cantidad de datos a información. (GUEVARA VEGA, 2015)

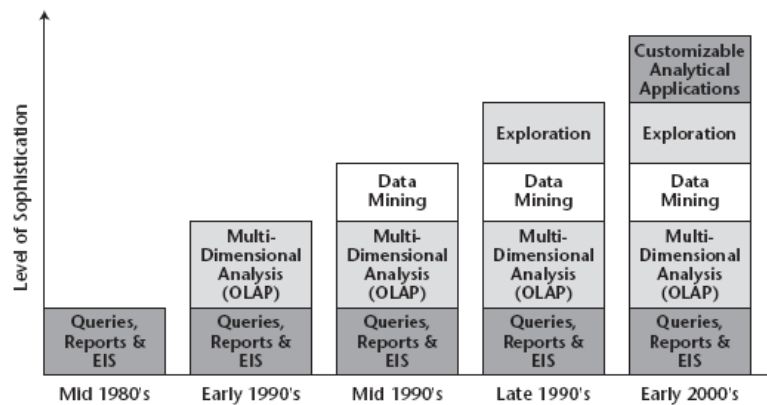


Ilustración 3: Evolución Business Intelligence

Fuente: <http://carlosproal.com/dw/dw01.html>

En 1969, se da el inicio de las bases de datos, para lo cual se definió varias reglas de evaluación para lo que en ese entonces se llamarían bases de datos relacionales, estas reglas permitían evaluar a los administradores de sistemas que manejaban relaciones, inicialmente los modelos relacionales no eran usados debido a que tenían problemas en cuanto al rendimiento y competitividad de base de datos jerárquicas. (DATAPRIX, s.f.). En la actualidad el uso de base de datos relacional se lo realiza en la mayor parte de sistemas transaccionales existentes en las diversas empresas, las base de datos relacionales se han constituido la primera fuente de datos para los sistemas de análisis de datos BI.

2.3.2.4 Desarrollo de bases de datos y aplicaciones empresariales (SAP, JD Edwards, Siebel, PeopleSoft).

En los años 80's, luego del desarrollo empieza el crecimiento del uso de base de datos, por lo cual se desarrolla el lenguaje de consultas o declarativos que se lo denomina Structured Query Language (SQL), el cual permite efectuar consultas con el fin de recuperar y realizar cambios de forma sencilla. Mejorando así las bases de datos anteriores que se diseñó para el procesamiento de transacciones. (Universidad Poliecnica de Valencia, s.f.).

En los años 90, surgen las primeras aplicaciones enfocadas en el uso de BI y DW, entre las principales IBM OLAP Server, Cognos, SAS y Oracle que hasta la actualidad han seguido en el mercado con las mejoras correspondientes. Todas las herramientas que iniciaron en estos años fueron denominadas aplicaciones BI 1.0,

ya que su almacenamiento era limitado y el funcionamiento mantenía ciertos inconvenientes. (GUEVARA VEGA, 2015).

Ya en los años 2000, con la mejora de la tecnología aparecen nuevos términos, en los cuales BI ya se considera una herramienta transversal en la mayoría de organizaciones, ya que con la información generada por los ERP's y CRM's se alimentaban con el objetivo de generar conocimiento a partir de los datos. En esta etapa los empleados, consultores, clientes, proveedores y público en general podían acceder a la información para poder tomar decisiones de forma correcta basada en información exacta y actualizada. (GUEVARA VEGA, 2015).

2.3.2.5 Componentes y Arquitectura del BI

Los componentes para implementar una solución BI nace de los sistemas que generan información sean estos bases de datos, ERP'S, CRM's, ficheros de texto, etc. Sobre los cuales se pueda realizar una transformación estructural para optimizar su proceso analítico. (Sinnexus, 2016), como se detalla en la Ilustración 4.

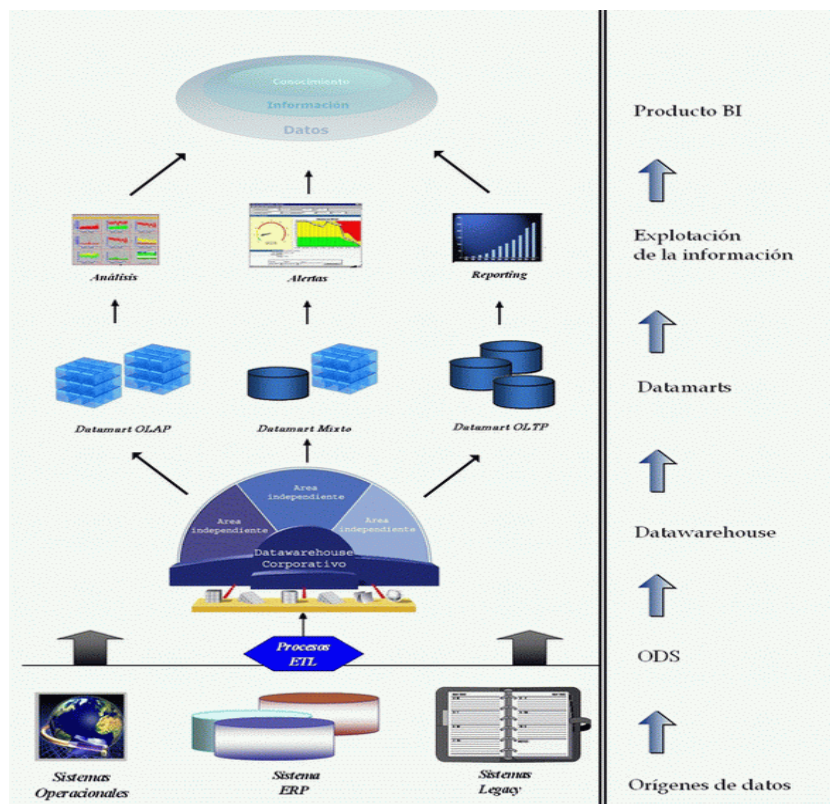


Ilustración 4: Arquitectura Business Intelligence

Fuente: (Sinnexus, 2016)

2.3.3 Metodologías, estrategias y análisis

2.3.3.1 Metodología Kimbal

Para el desarrollo de este proyecto se hace referencia a la metodología propuesta por Kimbal, la cual es empleada para la construcción de almacenes de datos DW orientados a un determinado ámbito de análisis. (WEB 2.0 - MediaWiki, 2014).

Como se aprecia en la Ilustración 5 propuesta por Brito M, se representa el ciclo de vida dimensional del negocio.

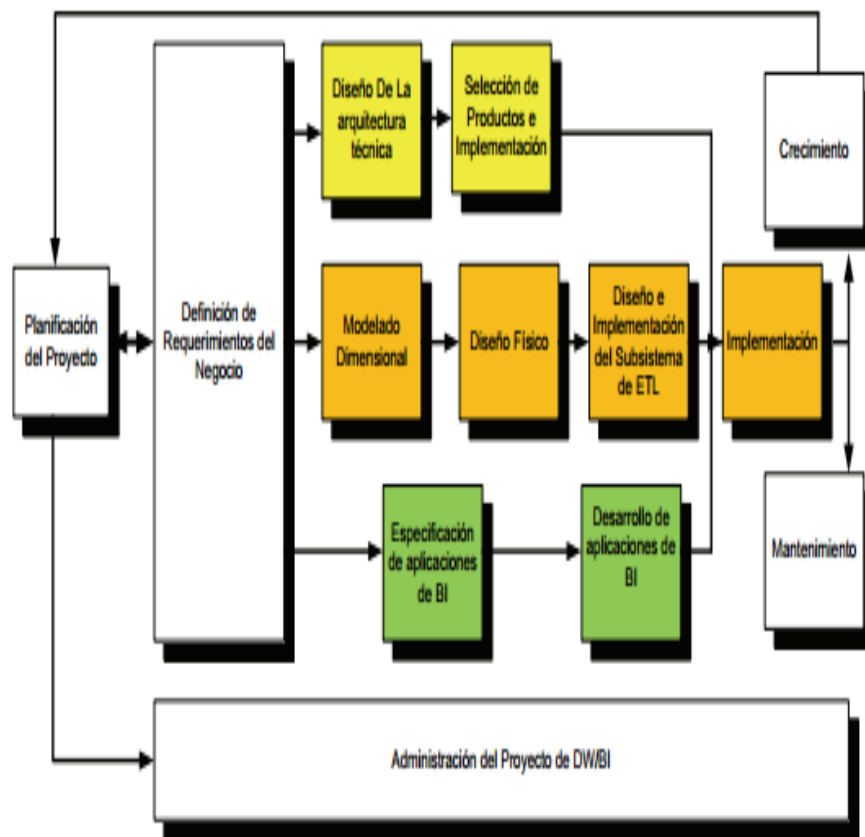


Ilustración 5: Ciclo de Vida Kimball.

Fuente: <http://ucv2014inteligenciadenegocios.blogspot.com/2014/09/post-3-metodologia-de-kimball.html>

La metodología se basa en el Ciclo de Vida Dimensional del negocio, que se encuentra basado en cuatro principios básicos.

- **Concentrarse en el negocio.**- Se basa en identificar los requerimientos del negocio y su valor asociado, para desarrolla las

relaciones sólidas con el negocio y enfocar los análisis de mejor manera. (Brito, 2014).

- **Construir una infraestructura de información adecuada.-** Se enfoca en el diseño de la base de información única, integrada y de fácil uso, con un rendimiento alto para reflejar los requerimientos de implementación de una forma adecuada. (Azuaje, 2014)
- **Realizar entregas en incrementos significativos.-** Se lo aplica en la creación de los almacenes de datos DW, en incrementos entregables cada cierto tiempo, por recomendación podría ser en plazos de 6 a 12. Los incrementos deben ser identificados para determinar el orden de aplicación, se recomienda la utilización de metodologías ágiles de construcción de software. (Azuaje, 2014)
- **Ofrecer la solución completa.-** Se lo efectúa al proporcionar los elementos necesarios para realizar las entregas de valor necesario a los usuarios finales. Se debe tener en este punto los almacenes de datos sólidos, bien diseñados y con calidad, se entregará las herramientas de consulta ad hoc, aplicaciones de informes, análisis avanzado, capacitación, soporte y documentación. (Azuaje, 2014)

2.3.4 Proceso de Data Analysis

El análisis de datos es la ciencia que permite inspeccionar, purificar y modelar los datos, con la finalidad de que las compañías, industrias y/o organizaciones tomen las mejores decisiones empresariales.

El análisis de datos se lo puede enfocar a todas las áreas en las cuales se posea datos y se requiera tener información válida para toma de decisiones.

Se describe el procesamiento de datos y transformación a información, como se muestra en la Ilustración 6.



Ilustración 6: Procesamiento de Información

Fuente: http://contenidosdigitales.ulp.edu.ar/exe/Sistema_de_informacion_integral/image001.png

2.3.4.1 Fuentes de Datos

Las fuentes de datos son los sistemas transaccionales instalados en las compañías, industrias y/o organizaciones, en los cuales se registran todos los datos estructurados según las necesidades de cada una de estas, los datos pueden ser obtenidos de sistemas como ERP, CRM, SCM, SAP, etc., los cuales pueden ser entregados realizando conexiones directas a las fuentes de datos de cada uno de estos o por medio de archivos planos, los cuales serán analizados, como se observa en la Ilustración 7.

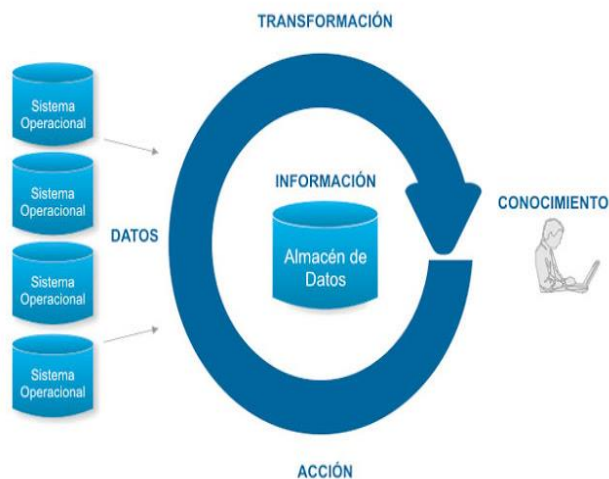


Ilustración 7: Fuentes de Datos

Fuente: <http://www.cdrconsultores.com.ve/inteligencia-de-negocios.html>

2.3.5 Proceso de Extracción, Transformación y Carga ETL

El termino ETL se simplifica del concepto de cada una de estas, “Extraer” los datos físicamente de las distintas fuentes de información. “Transformar” los datos para estructurarlos en los distintos modelos de análisis. “Cargar” (Load en inglés) los datos en una base de datos analítica optimizada, como se observa en la Ilustración 8.” (Gómez Morales, 2013)

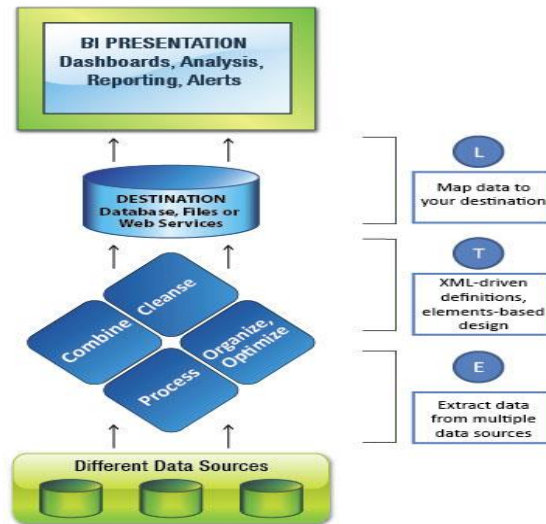


Ilustración 8: Esquema ETL

Fuente: (DATAPRIX, s.f.)

2.3.6 Almacenes de Datos (Data warehouse DWH)

Son grandes almacenes de datos que se crean y estructuran con la finalidad de analizar la información requerida, se lo realiza a diferentes niveles de detalle llamado granularidad, de cada una de las áreas o procesos que posee una organización. Es la base de la analítica ya que es conocida como base de datos estratégica o multidimensional. (Castro Roza, 2013).

Un Data warehouse es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta. La creación de un data warehouse representa en la mayoría de las ocasiones el primer paso, desde el punto de vista técnico, para implantar una solución completa y fiable de Business Intelligence. (Sinnexus, 2016).

Los DW se crearon con la finalidad de cambiar el enigma de que los datos sean validados y utilizados por las áreas técnicas con altos conocimiento en estructuras y estos sean utilizados por personal operativo con la finalidad de comprender de manera fácil los resultados y tomar decisiones, como se muestra en la Ilustración 9.(Codd, Codd y Salley, 1993).” (Abril Frade & Pérez Castillo, 2017).

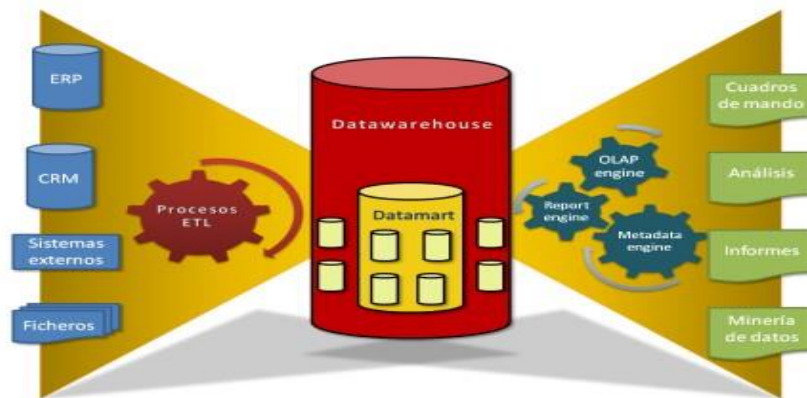


Ilustración 9: Data warehouse

Fuente: http://mundodb.es/wp-content/uploads/2013/04/Data_warehouse.jpg

Data warehouse debe cumplir con cuatro características principales que son:

- **Integrado.-** Los datos almacenados en el data warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios. (Sinnexus, 2016).
- **Temático.-** Los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del data warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar. (Sinnexus, 2016).

- **Histórico.-** El tiempo es parte implícita de la información contenida en un data warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el data warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, data warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones. (Sinnexus, 2016).

- **No Volátil.-** El almacén de información de un data warehouse existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del data warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía. (Sinnexus, 2016).

2.3.7 Modelo Multidimensional

Los modelos multidimensionales son específicos para la utilización en los almacenes de datos, las bases de datos multidimensionales (BDMB) son un tipo de base de datos optimizada para Data warehouse que se utilizan principalmente para crear aplicaciones OLAP, una tecnología asociada al acceso y análisis de datos en línea. (PowerData, 2015).

El modelo multidimensional maneja dos conceptos principales:

- **Hechos.-** Comprenden los criterios de análisis, poseen atributos llamados de hechos o de síntesis, y son de tipo cuantitativo. Sus valores también llamadas medidas se obtienen por la aplicación de una función estadística o matemática que resume un conjunto de valores en un único valor. Por ejemplo: ventas en dólares, cantidad de unidades en inventario, cantidad de unidades de producto vendidas, horas trabajadas, promedio de piezas producidas, consumo de agua potable de un sector. (Rodriguez Almeida & Da Silva Camargo, 2015)

- **Dimensiones.-** Son los criterios básicos y principales para los análisis, algunos ejemplos son: tiempo, producto, cliente, departamento, entre otras. Las dimensiones se utilizan para seleccionar y agrupar los datos en un nivel de detalle deseado. Los componentes de una dimensión se denominan niveles y se organizan en jerarquías, verbigracia, la dimensión tiempo puede tener niveles día, mes y año. (Rodríguez Almeida & Da Silva Camargo, 2015)

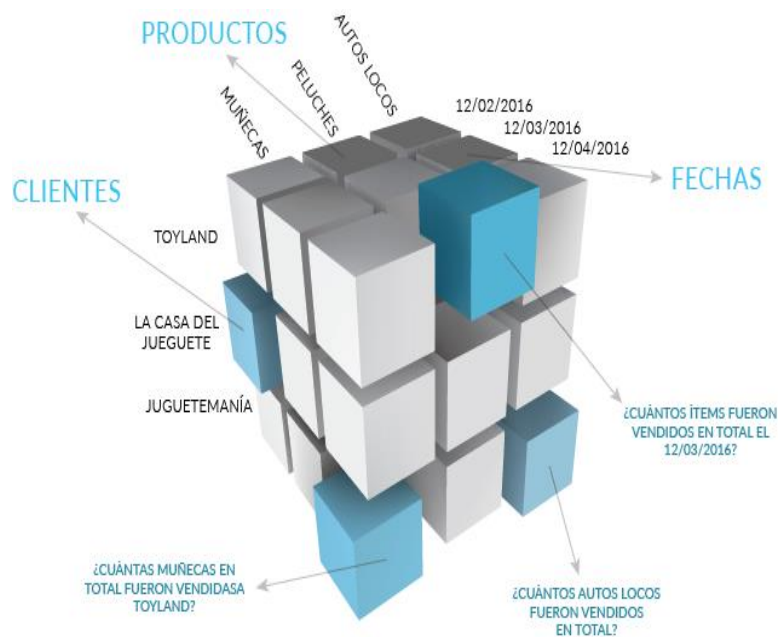


Ilustración 10: Multidimensional

Fuente: <http://i.imgur.com/Pz1wxud.png>

2.3.8 On-Line-Analytical-Processing (OLAP)

El procesamiento de la información se lo realizó de manera organizada, utilizando el procesamiento analítico en línea que permite obtener acceso a datos organizados y agregados de orígenes de datos empresariales, proporcionando una alternativa orientada al análisis de datos rápido y de navegación flexible. (GAMBOA CRUZADO, 2015)

OLAP ayudó a realizar operación de básicas como son agrupación y agregación con gran facilidad.

OLAP se clasifica por sus diferentes categorías que son:

- **ROLAP.-** Los datos tiene acceso directo desde el DW, la implementación almacena los datos en un motor relacional de tres niveles. (GAMBOA CRUZADO, 2015)
- **MOLAP.-** Permite pre calcular los datos para luego almacenarlos en una base de Datos multidimensional, el principio es optimizar los tiempos de respuesta, el resumen de la información es usualmente calculado por adelantado. (GAMBOA CRUZADO, 2015)

2.3.9 Aplicación de BI en empresas privadas en Ecuador

En cuanto la aplicación de BI en empresa privadas en el país se aplicó la técnica de muestro se seleccionaron 67 empresas privadas, los resultados obtenidos reflejan que las empresas privadas de gran tamaño tienen alta utilización de Business Intelligence para su toma de decisiones, ya que por la cantidad de información que se encuentran procesando mejora los tiempos de respuesta para obtener un informe que reflejen los datos deseados. Por otro lado el entorno que las herramientas de BI muestran es muy fácil su aprendizaje, uso y es amigable con el usuario. (Cueva Andrade, Jerez Cevallos, Díaz Zuñiga, & Ron Egas, 2014).

El principal motivo por el cual no se ha implantado Business Intelligence en empresas privadas de gran tamaño es el costo ya que para la implantación de herramientas Business Intelligence se requiere de grandes recursos económicos que no todas las empresas están dispuestas a pagar, a pesar que el valor invertido sería recuperado a mediano plazo ya que los beneficios que el uso de esta herramienta proporciona es de gran ayuda para mejorar ciertos procesos de algunas áreas que pueden encontrarse con problemas. Pero también las empresas que proveen este servicio deberían pensar la forma de tratar de atraer al mercado de empresas pequeñas para que ellos también participen en el uso de las herramientas de BI y de igual manera que las empresas grandes y medianas puedan mejorar sus técnicas y procedimientos. (Cueva Andrade, Jerez Cevallos, Díaz Zuñiga, & Ron Egas, 2014).

Los sectores en los cuales son más requeridos la implantación de Business Intelligence en las empresas privadas ecuatorianas son las dedicadas a la Banca y a las Industrias dedicadas a Bienes y Consumo en donde la Inteligencia de Negocio se vuelve parte fundamental en la toma de decisiones gerenciales. (Cueva Andrade, Jerez Cevallos, Díaz Zuñiga, & Ron Egas, 2014).

Como resultados de la aplicación del BI en el sector privado, cuando se implementa Business Intelligence en una empresa la toma de decisiones es más efectiva, ya que con la ejecución de informes precisos se puede llegar a obtener acciones eficientes que mejoren los procesos en ciertas áreas de negocio. (Cueva Andrade, Jerez Cevallos, Díaz Zuñiga, & Ron Egas, 2014).

2.3.10 Aplicación de BI en el sector público en el Ecuador

La aplicación de herramientas BI en el sector privado se ve reflejado en la investigación realizada por Cueva Cecilia, donde aplicando la técnica de muestreo se escogieron 37 empresas públicas para el análisis respectivo, los resultados obtenidos reflejan que la utilización de Business Intelligence en empresas públicas de gran tamaño en el sector público es Alta, al igual que en el sector privado los informes a nivel gerencial son herramientas muy útiles al momento de la toma de decisiones, por este motivo su uso es muy requerido en todo tipo de empresa de mediano y gran tamaño. (Cueva Andrade, Jerez Cevallos, Díaz Zuñiga, & Ron Egas, 2014)

En la investigación también se identifican factores por los cuales no se aplica el BI en las empresas públicas del Ecuador, existen 3 principales causas por las cuales las empresas del sector público no utilizan Business Intelligence, estas son: Costo, Falta de Personal y Falta de Conocimiento sobre BI. En el caso de las empresas públicas el costo es una variable que influye mucho al momento de contratar servicios externos como es el caso de herramientas BI ya que ellos cuentan con una partida presupuestaria 8 asignada por el gobierno anualmente, no contar con un capital propio y al ser una herramienta con un costo medianamente alto algunos de ellos no pueden acceder a este tipo de servicio que brindaría mejoras en sectores de sus empresas que pueden tener dificultad de mejorar y en las áreas que se encuentran bien mantenerlas de esa forma o seguir creciendo, pero

se debe tomar en cuenta que por el artículo de uso de software libre implantado a las empresas públicas, éstas no pueden tener un acceso como una empresa privada. (Cueva Andrade, Jerez Cevallos, Díaz Zuñiga, & Ron Egas, 2014).

Se identificó los sectores públicos donde la utilización de herramientas BI, en el sector Público constatamos que la utilización de Business Intelligence lo lidera el sector Energético, ya que debido a ser una parte importante dentro de la economía ecuatoriana las decisiones que se tomen en torno a este tema se transforma en cruciales para el país, por otro lado los demás sectores se encuentran parejos en la utilización de Business Intelligence. Estos resultados se encuentran tan parejos debido a la muestra obtenida, ya que puede que en algunos sectores el porcentaje real sea más alto y en otros más bajo, pero para alcanzar un mercado más amplio en los sectores empresariales se debe tomar en cuenta todas las áreas de negocio. (Cueva Andrade, Jerez Cevallos, Díaz Zuñiga, & Ron Egas, 2014).

Como resultados de la aplicación del BI en el sector público, “al igual que en el sector privado, las empresas publicas piensan que la toma de decisiones utilizando herramientas Business Intelligence mejoró notablemente y se pudo optimizar algunas áreas que no se las tomaba en cuenta y que ahora brindan un mejor servicio y calidad a sus consumidores. (Cueva Andrade, Jerez Cevallos, Díaz Zuñiga, & Ron Egas, 2014).

2.3.11 Base de Datos y Herramientas de desarrollo

2.3.11.1 Base de Datos Oracle 12c

En este desarrollo se está utilizando la versión de Oracle 12 c para almacenar los datos de las diferentes entidades que proveen la información, Oracle 12c ha cambiado radicalmente con respecto a las versiones anteriores, ha tenido muchas mejoras en el optimizador basado en costos (CBO), en la portabilidad de sus bases de datos, en la flexibilidad al realizar los cambios tanto físicos como lógicos y en su arquitectura haciéndola adecuada para hacer frente al nuevo mundo de “Cloud Computing”. (Oracle, 2014).

2.3.11.2 Toad for Oracle 12.11

Trabajamos con el Toad for Oracle 12.11 para la verificación de la información y sobre todo para la administración de la base de datos como tal, Toad

Software es un conjunto de herramientas de gestión de bases de datos de Quest que los desarrolladores de bases de datos, administradores de bases de datos y analistas de datos utilizan para gestionar bases de datos relacionales y no relacionales utilizando SQL. (Quest, 2014).

2.3.11.3 Oracle Business Intelligence

Para el análisis de información se está trabajando con la solución de Oracle Business Intelligence 12c, Oracle BI Enterprise Edition (a veces denominado Oracle Business Intelligence) proporciona una gran variedad de capacidades de inteligencia empresarial para realizar recopilación, presentación y entrega de datos de manera oportuna y fácil de entender. (Oracle, 2016).

2.3.11.4 ArcGis

Al tener que realizar la unión de datos georeferenciados, utilizamos la herramienta de ArcGis para realizar este trabajo, que comprende en unir la información de as diversas fuentes mediante capas y realizar la unión respectiva, ArcGIS es un completo sistema que permite recopilar, organizar, administrar, analizar, compartir y distribuir información geográfica. Como la plataforma líder mundial para crear y utilizar sistemas de información geográfica (SIG), ArcGIS es utilizada por personas de todo el mundo para poner el conocimiento geográfico al servicio de los sectores del gobierno, la empresa, la ciencia, la educación y los medios. ArcGIS permite publicar la información geográfica para que esté accesible para cualquier usuario. El sistema está disponible en cualquier lugar a través de navegadores Web, dispositivos móviles como smartphones y equipos de escritorio. (Esri, s.f.).

CAPITULO III

METODOLOGÍA

3 Introducción

En este capítulo, detallamos la información relevante para el cumplimiento de cada uno de los objetivos que se plantearon para realizar esta investigación, en primera instancia vamos a describir la metodología y herramienta de investigación utilizadas para centrarnos así en el método científico a utilizar en la investigación. En segunda instancia y con la ayuda de la información del marco referencial se buscara dar solución al problema formulado. Como tercer y último paso en este capítulo, detallamos paso a paso la implementación de la solución, aplicando la metodología, herramientas y técnicas seleccionadas previamente.

3.1 Métodos y Materiales

3.1.1 Método Inductivo

Este método lo utilizamos para obtener conclusiones que parten de hechos particulares aceptados como válidos, planteando un razonamiento que fluye de lo particular o individual hasta lo general. El método se inicia con un estudio individual de los hechos y se formulan conclusiones universales que se postulan como leyes, principios o fundamentos de una teoría (Abreu, 2014)

Su aplicación la realizamos en la etapa de obtener los resultados del análisis y las conclusiones de la investigación, ayudó para recopilar la información obtenida de los análisis de forma sistémica y lograr obtener resultados de fácil comprensión.

3.1.2 Método Analítico – Sintético

El análisis permitió descomponer en partes algo complejo (unión de datos de diversas entidades públicas), en desintegrar un hecho o una idea en sus partes, para mostrarlas, describirlas, numerarlas y para explicar las causas de los hechos o fenómenos que constituyen el todo. Posteriormente se puede aplicar el método comparativo, con el objetivo de establecer las principales relaciones que existen entre las variables de estudio. (Abreu, 2014)

Permitió analizar los datos obtenidos en el diagnóstico para llegar a determinar las estrategias, recursos, materiales y medios que intervienen en el proceso de implementación y desarrollo de las herramientas de gestión.

En esta investigación al tener la información ya recopilada, el método inductivo ayudó a determinar la mejor manera de relacionar los datos entre sí para poder analizarlos conjuntamente, tomando en cuenta que la investigación se basa en estudio de áreas geográficas delimitadas, la mejor opción fue relacionar los datos mediante coordenadas geográficas UTM Datum Pasad56 zona 17s.

3.1.3 Estrategias Técnicas

Se utilizaron las siguientes técnicas:

- **Entrevista:** Se realizaron varias entrevistas de trabajo con el personal del departamento de Agua no Contabilizada y Departamento de tecnologías de la Información y Comunicaciones de EMAPA, así como con el personal del proyecto Ibarra Verde, con el objetivo de definir los requerimientos de la implementación del BI.

- **Observación:** Se la realizaron visitas a las instalaciones de la Empresa Municipal de Agua Potable y Alcantarillado de Ibarra y poder verificar los procesos y las actividades que desarrollan, principalmente en el departamento de Agua no Contabilizada y Departamento de tecnologías de la Información y Comunicaciones de EMAPA.

3.1.4 Caso y unidad de análisis

El caso de estudio propuesto se lo realizó para el proyecto Ibarra Verde, el cual recopila los datos de la Empresa Municipal de Agua Potable y Alcantarillado del cantón Ibarra (EMAPA-I) a nivel de consumo desde el 2014 información del Instituto Nacional de Estadísticas y Censos (INEC) del censo realizado en el año 2010 en el cual consta toda la información a nivel de sectores censales e información proporcionada por el departamento de Avalúos y Catastros del Municipio de Ibarra de cada uno de los predios, con toda esta información se realiza una unión tomando en cuenta el posicionamiento geográfico para realizar

un análisis de consumo de agua por diversas variables como son, tarifas, sectores, educación del sector, dimensión del sector, etc.

3.1.5 Recolección de Datos

Para la recolección de datos, todas las instituciones públicas participantes proporcionan la información en archivos planos directamente al representante del proyecto Ibarra Verde, debido a que por seguridad no se tiene un acceso directamente a las bases de datos y tampoco ejecutar scripts que permita extraer la información, la información proporcionada será cargada a una base de datos transaccional Oracle, la cual fue diseñada para cumplir con los requerimientos previos al análisis.

El convenio de provisión de datos se lo realizó entre Ibarra Verde y cada una de las entidades participantes (Ver Anexos), en el cual detalle el alcance de provisión de datos de cada una de las entidades participantes.

3.2 Modalidad de la Investigación

La investigación que realizamos fue cuantitativa predominante, debido a que poseemos la información que vamos a analizarla y procesarla.

Para llevar a efecto la presente indagación se utilizamos algunos tipos de investigación como es: de Campo, Documental, Exploratorio, Descriptiva, bibliográfica que se detalla a continuación.

3.2.1 Investigación de Campo

Se realizó observando directa en el campo de estudio para verificar la realidad de la Empresa Municipal de Agua Potable y Alcantarillado de Ibarra, así como con el equipo de trabajo del proyecto Ibarra Verde, trabajando conjuntamente en con la información proporcionada, y en tiempo en que ocurren los fenómenos objetos de estudio; para realizar un diagnóstico y posteriormente una solución al problema observado.

3.2.2 Investigación Documental

Se realizó investigando todo lo relacionado al entorno a EMAPA en cuanto a la recolección, producción y distribución de agua potable en toda su área de concesión, se enfocó principalmente en el plan estratégico vigente, en el cual detalla como objetivo mejorar la distribución para satisfacer las necesidades de una población que va en crecimiento día a día, en cuanto al aporte documental del Proyecto Ibarra Verde, se revisó documentación en cuanto al crecimiento poblacional sostenible dentro de las urbes, lo que ayuda a entender que la planificación en cuanto al crecimiento poblacional ayuda a prevenir que las nuevas áreas pobladas carezcan de servicios básicos.

El desarrollo de la investigación se realizó tomando en cuenta recomendaciones de la OMS en cuanto al consumo mínimo y máximo de agua potable en Latinoamérica, así mismo en el tema de distribución de agua potable se tomó en cuenta los índices de desperdicio y pérdidas de agua potable con el fin reducir lo antes mencionado.

3.2.3 Investigación Bibliográfica

Es indispensable y se la realizó investigando en textos de información y documentos existentes, para determinar cuál es el conocimiento real sobre el tema a investigar y poder generar criterios en base a dichas fuentes, desarrollando así la fundamentación científica, filosófica y legal; y sobre todo tener una idea clara de la importancia del desarrollo del portafolio electrónico, basado principalmente en documentación que contenga información sobre la metodología de Kimball, ya que esta es la base de desarrollo de software, así como en bases legales como la Constitución, Plan Nacional de Buen Vivir, Reglamento de Régimen Académico y recomendación de la Organización Mundial de Salud (OMS) en cuanto al uso de agua potable.

3.3 Tipo o alcance de la investigación

3.3.1 Investigación Exploratorio

La aplicamos con fin el localizar oportunidades potenciales para la institución, así como comprobar hipótesis o alternativas que permitan establecer las líneas de actuación más convenientes.

Su aplicación se la realizo indagando en la misión de cada una de las entidades interesadas en el desarrollo de esta solución, con la intención de definir los objetivos y estrategias conjuntamente con los interesados en la investigación, ayudó a definir cuáles son los problemas principales de cada una de las entidades y definir los pasos a seguir para realizar la investigación propuesta.

3.3.2 Muestra o población

En esta investigación no trabajamos utilizando una muestra calculada en base al universo de datos proporcionado por EMAPA, realizamos varios filtros que ayudaron a enfocarnos en la totalidad de un grupo de suministros que cumplían todas las características necesarias para ser considerados dentro de esta investigación. La descripción de todos los filtros o características específicas que deben cumplir lo suministros la detallamos en el punto 4.1.1 Definición de la Muestra de Estudio.

3.4 Instituciones participantes en este proyecto

El proyecto se caracteriza por ser multi-institucional, en la cual cada una de las instituciones aportan con información valiosa para realizar el análisis. En este caso vamos a trabajar con 3 entidades públicas que proporcionan información valiosa como son EMAPA, Municipio de Ibarra e INEC.

3.4.1 Empresa Municipal de Agua Potable y Alcantarillado Ibarra

3.4.1.1 Datos generales

Tabla 5: Datos informativos EMAPA

NOMBRE	Empresa Municipal de Agua Potable y Alcantarillado de Ibarra
SIGLAS	EMAPA-I
TIPO	Empresa Municipal
FECHA DE CREACIÓN	12 de Agosto de 1969
TELEFONO	062 951 670
CASILLA POSTAL	100105
PORTAL WEB	http://www.emapaibarra.gob.ec/index.php
PROVINCIA	Imbabura
CANTON	Ibarra
CIUDAD	Ibarra
CALLE PRINCIPAL	Sucre 7-77
CALLE SECUNDARIA	Pedro Moncayo

Fuente: Elaboración Propia

3.4.1.2 Misión

“Somos una empresa que a través de una administración eficiente, suministra servicios de agua potable y saneamiento con calidad, para mejorar las condiciones de vida de los habitantes del cantón, enmarcados en valores, principios y normativas vigentes” (emapa, 2017).

3.4.1.3 Visión

“Seremos una empresa reconocida por la ciudadanía al dotar servicios de agua potable, saneamiento y tratamiento de aguas residuales oportunas y de calidad a través del desarrollo eficiente, autosustentable e integral de nuestro personal, en apego al marco legal aplicable y comprometido con la mejora continua.” (emapa, 2017)

3.4.1.4 Integrantes del directorio

“El Directorio está conformado por 5 miembros que son: a) Por el Alcalde o su delegado que deberá ser un funcionario o funcionaria del Gobierno Autónomo

Descentralizado, por el tiempo de duración para el cual fue elegido, quién lo presidirá; b) Por el Concejal que cumpla la función de Presidente de Obras Públicas, por un período de dos años; c) Por dos concejales delegados por el Concejo Municipal con sus respectivos suplentes, por un período de dos años; d) Por el representante de juntas parroquiales del Cantón Ibarra, por un período de dos años; y, e) El Gerente General de la empresa, actuará como Secretario con voz, pero sin voto, por un período de cuatro años.” (emapa, 2017)

Tabla 6: Miembros directorio EMAPA.

DIRECTORIO
PRESIDENTE
Ing. Álvaro Castillo
SECRETARIO
Ing. Arturo Fuentes
MIEMBROS
Arq. Ramiro Páez Valencia
Tnlg. Carmen Luisa Zehnder
Srta. Carina Rivadeneira
Sr. Hugo Aguirre

Fuente: Elaboración Propia

3.4.2 Gobierno Autónomo Descentralizado Ibarra

3.4.2.1 Datos generales

Tabla 7: Datos Municipio Ibarra

NOMBRE	Gobierno Autónomo Descentralizado Municipal San Miguel de Ibarra
AREA	O Avalúos y Catastros
DEPARTAMENTO	
TIPO	Gobierno Autónomo Descentralizado Municipal
TELEFONO	06 3700 200
CASILLA POSTAL	100105
PORTAL WEB	https://www.ibarra.gob.ec/web/
PROVINCIA	Imbabura
CANTON	Ibarra

CIUDAD	Ibarra
CALLE PRINCIPAL	Calle García Moreno 6-31
CALLE SECUNDARIA	Calle Simón Bolívar.

Fuente. Elaboración Propia

3.4.2.2 Misión

“Somos un gobierno municipal que a través de una administración eficiente, fomenta el desarrollo integral del cantón, brindando servicios de calidad enmarcados en valores, principios y normativas, para mejorar las condiciones de vida de sus habitantes.” (Ibarra, 2017).

3.4.2.3 Visión

“Ser un gobierno incluyente, reconocido por la ciudadanía por brindar servicios públicos de calidad, cumpliendo con los principios de gobernabilidad, para alcanzar un desarrollo ordenado, económico, social, turístico, productivo y seguro. Posicionando al cantón Ibarra en el año 2019 como referente nacional e internacional.” (Ibarra, 2017).

3.4.2.4 Autoridades

“En el contexto de una Administración Municipal innovadora, promotora del desarrollo local y cuyos esfuerzos se orientan fundamentalmente a la atención de las demandas ciudadanas, se ha elaborado el Plan Estratégico Participativo Institucional del Gobierno Autónomo Descentralizado Municipal San Miguel de Ibarra 2014 – 2019, teniendo presente que la calidad de vida de una población, se mide por la calidad de los servicios que recibe.” (Ibarra, 2017).

Tabla 8: Autoridades Gobierno Autónomo Descentralizado Municipal San Miguel de Ibarra

AUTORIDADES
ALCALDE
Ing. Alvaro Castillo
CONCEJALES
Ing. Andrea Elizabeth Scacco
Ing. Carina Elizabeth Rivadeneira Ruiz
Ing. Diana Elena Harrington Herrera

Srta. Carla Gabriela Proaño Chiriboga

Ing. Esther Hilda Herrera Tapia

Ing. Juan Manuel Mantilla Echeverría

Ing. Carlos Lenin Arias Brito

Arq. Jaime Ramiro Páez Valencia

Ing. Laureano Alencastro

Fuente: Elaboración Propia

3.4.3 Instituto Nacional De Estadística Y Censos Ibarra

3.4.3.1 Datos generales

Tabla 9: Datos INEC

NOMBRE	Instituto Nacional De Estadística Y Censos Ibarra
SIGLAS	INEC
TIPO	Gubernamental
TELEFONO	(593-2) 2232303
PORTAL WEB	http://www.ecuadorencifras.gob.ec/institucional/home
PROVINCIA	Pichincha
CANTON	Quito
CIUDAD	Quito
CALLE PRINCIPAL	Juan Larrea
CALLE SECUNDARIA	José Riofrío

Fuente: Elaboración Propia

3.4.3.2 Misión

“Coordinar, normar y evaluar la producción de la información estadística oficial proveniente del Sistema Estadístico Nacional, mediante la planificación, ejecución y análisis de operaciones estadísticas oportunas y confiables, así como de la generación de estudios especializados que contribuyan a la toma de decisiones públicas y privadas y a la planificación nacional” (Instituto Nacional de Estadística y Censos, 2017).

3.4.3.3 Visión

“Ser un referente a nivel nacional e internacional por la calidad, oportunidad e innovación en la producción de información estadística, en la articulación de las entidades del Sistema Estadístico Nacional y en la generación de conocimiento a través de estudios especializados.” (Instituto Nacional de Estadística y Censos, 2017)

3.4.3.4 Organigrama



Ilustración 11: Organigrama INEC

Fuente: <http://www.ecuadorencifras.gob.ec/la-institucion/>

3.5 El sistema

Todas las instituciones poseen sistemas ERP, por lo cual la información se va generando y actualizando según las necesidades de cada una. Las arquitecturas y base de datos son diferentes en cada institución dependiendo mucho de sus políticas, en unos casos utilizan herramientas de paga y en otras software libre.

Los sistemas de cada institución generan información según la misión de cada una de estas, cumpliendo los objetivos para los cuales fueron implementados.

3.5.1 Problema del sistema actual

En la actualidad, no existe un sistema en el cual se pueda realizar un análisis de la información uniendo varias instituciones, cada institución maneja sus propios sistemas de información los cuales generan información muy valiosa para realizar análisis.

Al existir mucha información generada por las instituciones parte de este proyecto, se tiene una gran oportunidad para obtener los datos y analizarlos según nuestras necesidades, en este caso enfocaremos a realizar análisis de consumos y distribución de agua potable, tomando en cuenta variables como número de habitantes, extractos sociales, áreas de sectores y demás.

3.5.2 Delimitación del alcance a desarrollar

La implementación de una aplicación Business Intelligence se limita a analizar los datos provistos por las instituciones antes mencionadas. El desarrollo de este análisis va enfocado para utilización en el Proyecto Ibarra Verde, con el objetivo de crear un análisis socio-ambiental interactivo para el del uso de agua y así guiar al desarrollo de las zonas estudiadas.

3.5.3 Involucrados

Al ser un proyecto multi institucional, los involucrados desean tener acceso a la herramienta desarrollada, por lo cual se ha identificado a los dos beneficiarios directos de esta investigación.

Tabla 10: Beneficiarios

BENEFICIARIO DIRECTO	Proyecto Ibarra Verde
BENEFICIARIO DIRECTO	Empresa Municipal de Agua Potable y Alcantarillado Ibarra.

Fuente: Elaboración Propia

3.5.4 Equipo técnico responsable

El equipo técnico es el responsable de desarrollar y dar mantenimiento a la aplicación hasta que esta se encuentre entregada y puesta en producción. Luego

de la entrega las instituciones beneficiadas serán las encargadas de asignar al personal de mantener la aplicación actualizada y en correcto funcionamiento.

Tabla 11: Equipo técnico responsable

Participante	Ocupación	Oficina	Experiencia
Ing. Andrés Fernando Cardenas Pepinós	Analista de TIC'S	Dirección de Tecnologías y Comunicaciones Empresa Eléctrica Regional Norte.	de 4 años en desarrollo de aplicaciones informáticas.

Fuente: Propia

3.5.5 Factibilidad Técnica

Las instituciones antes mencionadas prestan todo el apoyo respetivo en cuanto al acceso de datos y guías en temas específicos por parte del personal de cada institución, al ser Ibarra Verde un proyecto nuevo no posee infraestructura propia en la cual se pueda implementar la solución propuestas, realizando gestiones personales con el apoyo del equipo de trabajo de proyecto se consiguió se proporcione un servidor por un lapso limitado de tiempo, al finalizar este periodo con un tiempo factible de anterioridad se notificara para la extracción y migración de la máquina virtual a donde los dueños del proyecto crean conveniente.

El licenciamiento de las herramientas para BI se consiguió gracias a la gestión del Dr. Larry Frolich representante del proyecto Ibarra Verde.

3.5.6 Factibilidad Operativa

El proyecto Ibarra Verde y las instituciones participantes en este proyecto, dan todas las facilidades para proporcionar los datos que van a ser utilizados en la implementación de esta herramienta.

La Universidad Técnica del Norte, promueve e impulsa las investigaciones que tenga como objetivo mejorar el estilo de vida de las personas.

El equipo de trabajo del proyecto Ibarra Verde tiene toda la predisposición de sacar este proyecto adelante, con la ayuda de sus diferentes integrantes en las

diversas áreas, como son los ingenieros en recursos renovables, investigadores de campo y desarrolladores de herramientas informáticas.

3.6 Desarrollo de la propuesta

En esta sección trataremos y presentamos el modelo de inteligencia de negocio a presentar, seleccionaremos la metodología con la cual desarrollaremos la solución tecnológica. Se definirá la arquitectura tecnológica a desarrollar basada en la metodología seleccionada, así como temas y principios técnicos que son parte del desarrollo de software.

3.6.1 Selección de metodología para el desarrollo de la propuesta

Tomando en cuenta la investigación bibliográfica, la metodología que se adapta a nuestro desarrollo es la propuesta por Ralph Kimball, la cual posee una gran cantidad de desarrollos relevantes similares al que vamos a realizar.

3.6.2 Desarrollo de la Herramienta BI

Para el desarrollo de la propuesta, hemos seleccionado la utilización de las herramientas de la familia Oracle, tanto en base de datos como para Business Intelligence, con estos garantizaremos la confiabilidad y disponibilidad de la información ya que Oracle es una de las mejores si no la mejor herramienta para manejo de bases de datos.

Cabe detallar que la información proporcionada por las diversas empresas parte del proyecto Ibarra Verde, es entregado en archivos planos por lo cual el registro de la información en la base de datos no tiene ningún problema.

3.6.3 Arquitectura de Análisis de Datos con aplicación técnica de Business Intelligence

La arquitectura del desarrollo, hace énfasis en la unión y transformación de los datos provistos por las diversas entidades y generación de información consolidada con el fin de analizarla y tomar decisiones, el principio de unión de datos es que estén debidamente geo referenciados.

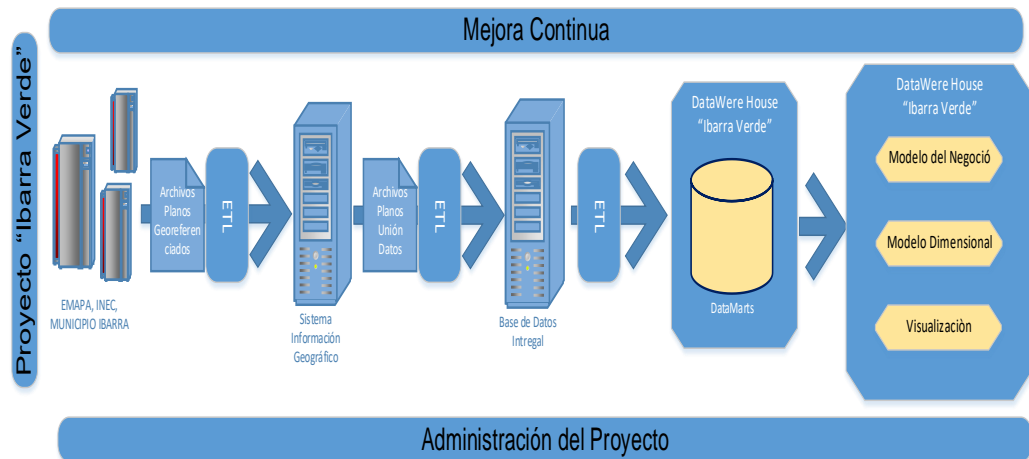


Ilustración 12: Arquitectura propuesta

Fuente: Propia

3.6.4 Análisis de datos

Como se detalló anteriormente esta investigación incluye varias instituciones, por lo cual cada una de estas ayudaron con datos en formato plano debidamente geo referenciados, con el objetivo de consolidar toda esta información y realizar un análisis general, en el cual incluya, dotación de agua potable, número de habitantes, extensión de predios , etc. En este punto se utilizó la herramienta de ArcGis, la cual permitió unir y consolidar toda la información.

3.6.5 Procesamiento de datos

Los datos provistos por las diferentes instituciones fueron tratados para lograr consolidar los datos y obtener resultados consolidados tomando en cuenta los factores de cada una de las instituciones.

El procesamiento de datos se detalla en el diagrama de arquitectura (Ilustración 12), en la cual se detalla la transformación de los datos desde archivos planos hasta obtener la información consolidada en los DataMarts.

Los procesos ETL de cada una de la entidades públicas, se las cargara dependiendo de cada necesidad, se han generado métodos ETL para facilitar la carga de información de cada una de las instituciones, estos debido a que no se posee una conexión directa a los sistemas de las diferentes empresas. En el caso

de EMAPA la carga será con mayor frecuencia ya que los datos se generan mensualmente.

3.7 Implementación del Business Intelligence

3.7.1 Planeación del proyecto

Al ser el proyecto Ibarra Verde un proyecto nuevo e innovador, no tiene una base o metodologías de desarrollo de software definido, por lo cual el la implementación de esta herramienta nosotros proponemos la metodología de desarrollo que más se adaptan a nuestras necesidades, en el caso de EMAPA tampoco poseen desarrollos propios en cuanto al área de Business Intelligence, por lo cual la aplicación de la metodología servirá como base para futuros desarrollos por parte de EMAPA en el área especificada.

El desarrollo y pacificación del desarrollo del proyecto está basado en la metodología propuesta por Kimball.

Como todo desarrollo de software y más por ser nuevo tiene riesgos los cuales deben ser mitigados y tomados en cuenta en cada una de las fases.

Para citar algunos de los riesgos más importantes en el desarrollo del proyecto consideramos los siguientes:

- Definir con exactitud los objetivos de la investigación y desarrollo de software, para mitigar el riesgos de crear expectativas difíciles de cubrir
- Al ser un proyecto multi-institucional, se debe contar con los respaldos físicos de la predisposición administrativa para garantizar el acceso y aprovisionamiento de la información.
- Toda la información proveída por cada una de las instituciones se manejan con la respectiva confidencialidad del caso.
- Se debe asegurar que la investigación y el desarrollo del software cumpla con las necesidades de cada usuario.

El resultado final de todo el desarrollo es tener una herramienta 100% funcional basado en las necesidades de cada usuario.

3.7.2 Capa de datos

Como se detalló anteriormente lo datos provistos por las diferentes entidades fueron archivos planos, los cuales fueron trasladados a una base de datos relacional, con el objetivo de luego realizar la consolidación de todos los datos en los esquemas estrella.

Con los datos ya en la base de datos relacional realizamos el proceso de construcción de los DataMart, para poder realizar los análisis respectivos.

3.7.2.1 Datos EMAPA

En el caso de EMAPA, la información que fue entregada fue la de los registros de consumo mensual de cada medidor, la información fue entregada en archivos planos los cuales para facilidad de manejo fueron trasladados a una base de datos relacional, como se muestra en la Ilustración 13.

CUENTA	CEDULA	NOMBRES	CALLES	TARIFA	REG	ANIO	CONCATENADO	NORTE	ESTE	PLANES	RUT	SECUENCI	ENER	FEBBER	MARZ	ABRI	
138899	9900013809	IVAN MANTILLA PEREZ	AV. ATAHUALPA 25-163	RES. CIUDAD	RUD	2016	15020010						10	34	42	28	38
138910	1705792458	ROSA SANCHEZ ALVARADO	AV. ATAHUALPA 25-101	RES. CIUDAD	RUD	2016	15020060	820588.1	100366531.7	15	2		80	3	7	5	4
138911	1001776673	ROSA ESTHER REALTE ZAMBRANO	AV. ATAHUALPA	RES. CIUDAD	RUD	2016	15020060	820588.6	100366448.2	15	2		80	38	39	40	40
138912	10009927186	ELVIA MAGDALENA BASTIDAS ERAZO	AV. ATAHUALPA	RES. CIUDAD	RUD	2016	15020090	820571	100366555.8	15	2		90	20	14	16	16
138913	1000866192	LUIS OSWALDO GUDINO HIGUERA	AV. ATAHUALPA 25-59	RES. CIUDAD	RUD	2016	15020110	820599.2	10036673.44	15	2		110	36	29	30	31
138914	1001985156	DOLGA MAGDALENA GUDINO HIGUERA	AV. ATAHUALPA	SEDAO. CIUDAD	RUD	2016	15020120	820592	100366991.8	15	2		120	24	17	23	22
138915	1000315094	JULIO CESAR PINEDA AGUIRRE	PASAJE PATRICIO GRANDA 1-56	SEDAO. CIUDAD	RUD	2016	15020130	820595.8	100367207.6	15	2		130	0	0	6	79
138916	1002146803	PAZ JIMENA LIEFERNA POTO	AV. ATAHUALPA	RES. CIUDAD	RUD	2016	15020140	820595.8	10036712.9	15	2		140	15	10	9	12
138917	1001597531	VENUS EDITH PORTILLA MORAN	FRAY PEDRO BEDON 1-08	RES. CIUDAD	RUD	2016	15020150	820601.7	10036750.7	15	2		150	12	10	12	10
138918	1000790673	IGACIO RODRIGO ESPINOSA ERAZO		RES. CIUDAD	RUD	2016	15020170	820602.0	10036771.9	15	2		170	27	33	25	24
138919	1210394101	JORGE EDUARDO Y ANO SAMPOLVAL GARCIA		COM. CIUDAD	RUD	2016	15020180	820607	10036795.01	15	2		180	27	29	26	40
138920	1000747848	WASHINGTON B. ESPINOSA SARZOSA	EDUARDO ALMEIDA 1-75 Y T. GOMEZ	RES. CIUDAD	RUD	2016	15020190	820610.8	10036819.2	15	2		190	63	38	13	43
138921	1001310919	ANTONIO WASHINGTON CHAVEZ BENALCAZAR	AV. ATAHUALPA	RES. CIUDAD	RUD	2016	15020200	820611.3	10036830.44	15	2		200	75	28	28	27
138922	1000827640	HERNANDO ALEX MENA BUALES	AV. ATAHUALPA 25-67	RES. CIUDAD	RUD	2016	15020220	820615.7	10036851.7	15	2		220	35	25	30	29
138923	100791010	VICTOR MERA	AV. ATAHUALPA 25-53	SEDAO. CIUDAD	RUD	2016	15020230	820617	10036866.9	15	2		230	49	41	37	41
138924	1001314619	HERLINDA ESPERANZA ROLAÑOS HERNANDEZ	AV. ATAHUALPA 23-29	RES. CIUDAD	RUD	2016	15020240	820619	10036880.4	15	2		240	10	11	11	21
138925	9900013825	RICARDO F. VERA ROSADO	AV. ATAHUALPA 23-25	COM. CIUDAD	RUD	2016	15020240	820615	10036880.4	15	2		240	10	45	6	20
138926	100125086	DAVID GONZALO VASQUEZ ERAZO	HERNAN GONZALEZ DE SAA 11-91	COM. CIUDAD	RUD	2016	15020270	820623.9	10036900.5	15	2		270	34	27	27	24
138927	170076056	TERESA MARIANA DAZA PEÑON	CUARTO ESQUINAS	SEDAO. CIUDAD	RUD	2016	15020280	820625.9	10036887.1	15	2		280	15	14	14	13
138928	9900013828	ANGEL ALMEIDA MIGUEL	HERNAN GONZALEZ DE SAA 11-73	RES. CIUDAD	RUD	2016	15020290	820592.9	10036882.4	15	2		290	24	29	27	24
138929	1000101829	LAURO OSWALDO THADO SALGADO	IMBAYA 1-28	SEDAO. CIUDAD	RUD	2016	15020310	820574.9	10036869.3	15	2		110	9	10	10	9
138930	1701927751	TERESITA DE JESUS SUAREZ HEREDIA	IMBAYA 1-62	SEDAO. CIUDAD	RUD	2016	15020330	820575.2	10036830.6	15	2		330	26	9	12	16
138931	1001295458	MARIA ANTONIETA ALARCON HIDALGO	IMBAYA	DISCAPACIDAD CIUDAD	RUD	2016	15020340	820543.4	10036785.8	15	2		340	23	20	15	14
138932	400896050	JOSE FRANCISCO YANILIN CUBARAN	IMBAYA	SEDAO. CIUDAD	RUD	2016	15020350	820543	10036772.4	15	2		350	11	29	29	11
138933	9900013833	ARTURO TORRES LUIS	IMBAYA 1-105	RES. CIUDAD	RUD	2016	15020360	820542.5	10036780.6	15	2		360	29	10	14	26

Ilustración 13: Datos EMAPA

Fuente: Propia

Una vez con la información en archivos planos, esta fue trasladada a las tablas de nuestra base de datos relacional, como se muestra en la Ilustración 14 la tabla suministro.

Name	Datatype	Not Null
ID_CUENTA	NUMBER	<input type="checkbox"/>
ID_TARIFA	VARCHAR2(10)	<input checked="" type="checkbox"/>
ID_RED	VARCHAR2(10)	<input checked="" type="checkbox"/>
IDENTIFICACION...	VARCHAR2(13)	<input checked="" type="checkbox"/>
ID_SECTOR	VARCHAR2(20)	<input checked="" type="checkbox"/>
PLAN	VARCHAR2(5)	<input checked="" type="checkbox"/>
RUTA	VARCHAR2(5)	<input checked="" type="checkbox"/>
SECUENCIA	VARCHAR2(5)	<input checked="" type="checkbox"/>
CONCATENADO	VARCHAR2(20)	<input checked="" type="checkbox"/>
DIRECCION	VARCHAR2(50...)	<input checked="" type="checkbox"/>
COOR_NORTE	NUMBER	<input checked="" type="checkbox"/>
COOR_ESTE	NUMBER	<input checked="" type="checkbox"/>

Ilustración 14: Campos de la Tabla Suministros

Fuente: Propia

Para los datos de los clientes se utilizó una estructura básica, debido a que los datos como teléfono, email y demás nos son de interés para esta investigación como se muestra en la Ilustración 15.

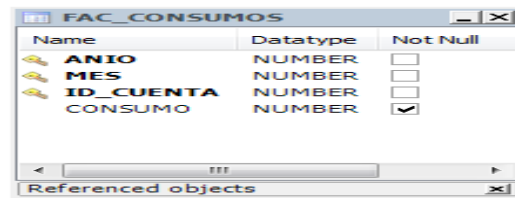
Name	Datatype	Not Null
IDENTIFICACION_CLI	VARCHAR2(13)	<input type="checkbox"/>
NOMBRES	VARCHAR2(500)	<input checked="" type="checkbox"/>

Ilustración 15: Campos de la Tabla Clientes

Fuente: Propia

La mayor cantidad de la información es la perteneciente a los consumos mensuales por cada suministros, en esta tabla encontramos los datos de fecha del

registro de consumo y la cantidad de consumo realizado por el sus ministros en metros cúbicos como se observa en la Ilustración 16.

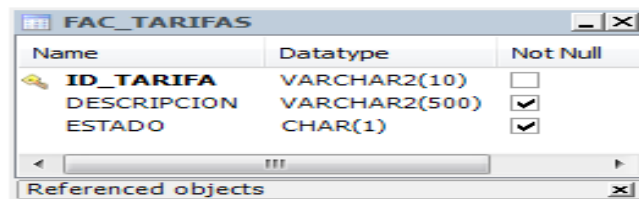


Name	Datatype	Not Null
ANIO	NUMBER	<input type="checkbox"/>
MES	NUMBER	<input type="checkbox"/>
ID_CUENTA	NUMBER	<input type="checkbox"/>
CONSUMO	NUMBER	<input checked="" type="checkbox"/>

Ilustración 16: Campos de la Tabla Consumos

Fuente: Propia

Para las tarifas se ha realizado la estructura en la cual conste el ID_TARIFA el cual es el que actualmente maneja la empresa para su identificación con su debida descripción y su estado, al realizar la investigación de decidió registrar solo las tarifas que se encontraban activas en ese momento como se muestra en la Ilustración 17.

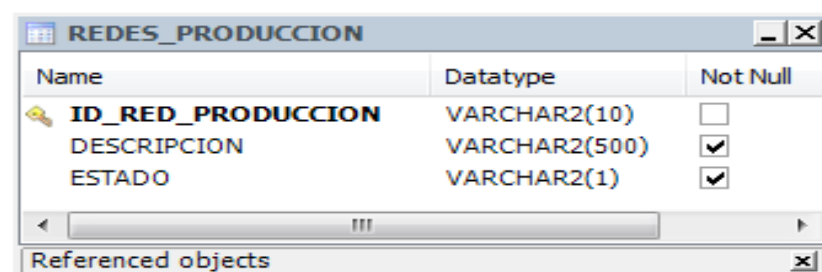


Name	Datatype	Not Null
ID_TARIFA	VARCHAR2(10)	<input type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>
ESTADO	CHAR(1)	<input checked="" type="checkbox"/>

Ilustración 17: Campos da la Tabla Tarifas

Fuente: Propia

Para el registro de las plantas de producción o también conocidos como tanques de producción, se implementó la estructura para almacenar la información de cada uno de estos tanques en la cual se detalla su identificativo propio así como su descripción como se muestra en la Ilustración 18.



Name	Datatype	Not Null
ID_RED_PRODUCION	VARCHAR2(10)	<input type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>
ESTADO	VARCHAR2(1)	<input checked="" type="checkbox"/>

Ilustración 18: Campos de la Tabla Producción

Fuente: Propia

Para la identificación de sectorial o de comercialización se detallan las redes de comercialización, la cual consta de información propia como id y descripción que sirven para la identificación de estas, cabe detallar que un tanque o planta de distribución puede abastecer a varias redes de comercialización como se muestra en la Ilustración 19.

Name	Datatype	Not Null
ID_RED_COMERCIAL	VARCHAR2(10)	<input type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>
ESTADO	CHAR(1)	<input checked="" type="checkbox"/>
ID_RED_PRODUCCION	VARCHAR2(10)	<input checked="" type="checkbox"/>

Ilustración 19: Campos de la Tabla Redes Comerciales

Fuente: Propia

Al tener cada una de las estructuras desarrolladas con sus debidas relaciones, se obtiene en modelo relacional de los datos entregados por EMAPA, diagrama que se muestra en la Ilustración 20.

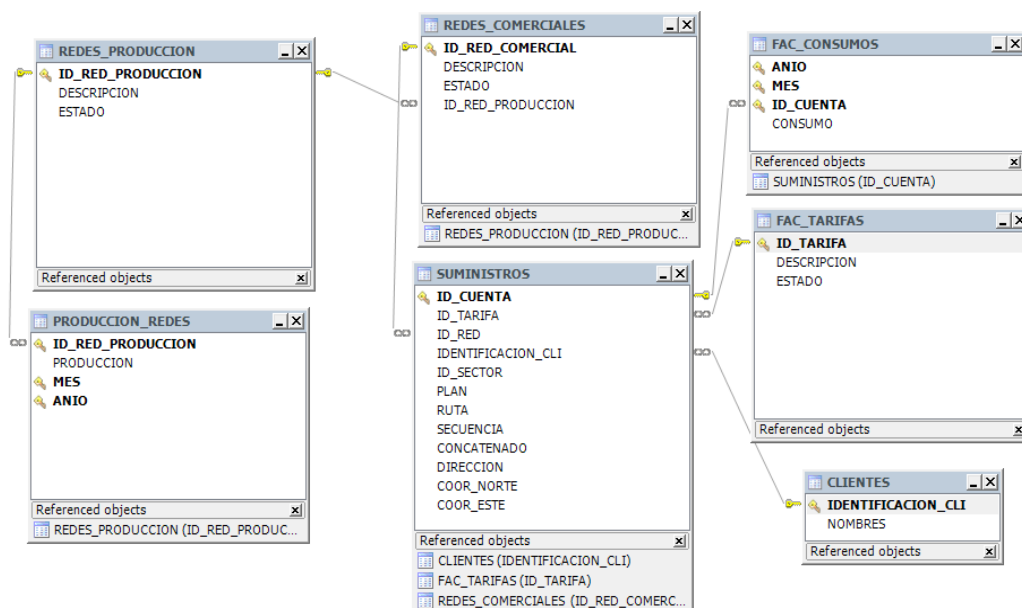


Ilustración 20: Diagrama Entidad Relación EMAPA

Fuente: Propia

3.7.2.2 Datos INEC

En el caso del INEC, la información que fue provista fue con respecto a la información censal 2010, en la cual se detalla las provincias, cantones, parroquias, zonas, sectores, población, educación y demás, igualmente para facilidad de manejo se lo traslado a una base de datos relacional, como se muestra en la Ilustración 21.

The screenshot shows an Excel spreadsheet with a grid of data. The columns are labeled with letters A through X, and the rows are numbered 1 through 40. The data appears to be a list of records for different cantons or sectors, with columns for identifiers, names, and various numerical values. The spreadsheet interface includes the standard Excel ribbon with tabs like 'ARCHIVO', 'INICIO', 'INSERTAR', etc.

Ilustración 21: Datos INEC

Fuente: Propia

Los datos al ser entregados en archivos planos se realizó el debido tratamiento para llevarlos a nuestra base de datos relacional, de los cuales de detallan las siguientes entidades.

Tenemos la estructura de los cantones, la cual consta de su id y su respectiva descripción como se muestra en la Ilustración 22.

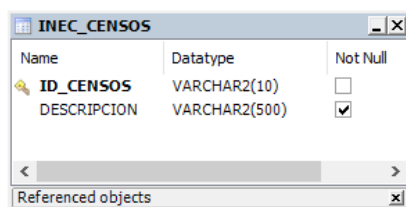
The screenshot shows a window titled 'INEC_CANTONES' displaying the structure of a database table. It lists two columns: 'ID_CANTON' and 'DESCRIPCION'. 'ID_CANTON' is of type VARCHAR2(10) and is not null. 'DESCRIPCION' is of type VARCHAR2(500) and is null. There is a 'Referenced objects' section at the bottom.

Name	Datatype	Not Null
ID_CANTON	VARCHAR2(10)	<input type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>

Ilustración 22: Campos tabla Inec_Cantones

Fuente: Propia

Al tener información solo del censo 2010, pero con la proyección de que la aplicación este vigente muchos años se creó la estructura de censos, para poder almacenar la información de censos posteriores como se muestra en la Ilustración 23.




Name	Datatype	Not Null
ID_CENSOS	VARCHAR2(10)	<input type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>

Ilustración 23: Campos Tabla Inec_Censos

Fuente: Propia

Para el almacenamiento de la información de área y número de habitantes por sector y por censo, se creó la estructura de habitantes sectores, la cual almacena la información relacionada con cada sector censal, identificativos del censo, área en metros cuadrados del área y número de habitantes como se muestra en la Ilustración 24.

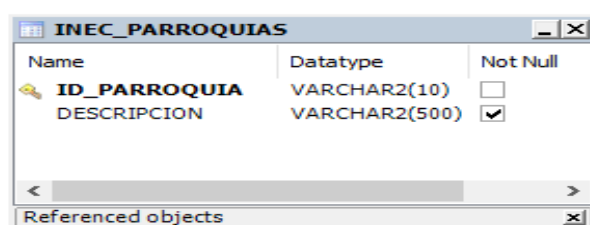


Name	Datatype	Not Null
ID_CENSOS	VARCHAR2(10)	<input type="checkbox"/>
ID_SECTOR	VARCHAR2(20)	<input type="checkbox"/>
HABITANTES	NUMBER	<input checked="" type="checkbox"/>
AREA	NUMBER	<input checked="" type="checkbox"/>

Ilustración 24: Campos Tabla Inec_Habitantes_Sectores

Fuente: Propia

Se creó la estructura para la información de las parroquias, en la cual podemos almacenar la codificación de la parroquia y su descripción como se muestra en la Ilustración 25.



Name	Datatype	Not Null
ID_PARROQUIA	VARCHAR2(10)	<input type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>

Ilustración 25: Campos Tabla Inec_Parroquias

Fuente: Propia

Se creó la estructura de provincias, para almacenar la información de las provincias, con su respectiva codificación y descripción como se muestra en la Ilustración 26.

Name	Datatype	Not Null
ID_PROVINCIA	VARCHAR2(10)	<input type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>

Ilustración 26: Campos Tabla Inec_Provincias

Fuente: Propia

Se creó la estructura de cada sector el cual se compone del id propia e información relacionada como es, provincia, cantos, zona, parroquia y descripción, cabe detallar que una zona puede tener varios sectores como se observa en la Ilustración 27.

Name	Datatype	Not Null
ID_SECTOR	VARCHAR2(20)	<input type="checkbox"/>
ID_PROVINCIA	VARCHAR2(10)	<input checked="" type="checkbox"/>
ID_CANTON	VARCHAR2(10)	<input checked="" type="checkbox"/>
ID_ZONA	VARCHAR2(13)	<input checked="" type="checkbox"/>
ID_PARROQUIA	VARCHAR2(10)	<input checked="" type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>

Ilustración 27: Campos Tabla Inec_Sectores

Fuente: Propia

Para la totalización de datos relacionados a cada sector, se creó una estructura en la cual se relaciona el sector con cada uno de los valores entregados por el INEC del último censo como se muestra en la Ilustración 28.

Name	Datatype	Not Null
ID_SECTOR	VARCHAR2(30)	<input checked="" type="checkbox"/>
TOTAL_POBLACION	NUMBER	<input checked="" type="checkbox"/>
AREA	NUMBER	<input checked="" type="checkbox"/>
DENSIDAD	NUMBER	<input checked="" type="checkbox"/>
NBI	NUMBER	<input checked="" type="checkbox"/>
DEF_CUALIT	NUMBER	<input checked="" type="checkbox"/>
DEF_CUANT	NUMBER	<input checked="" type="checkbox"/>
PEA_ASALAR	NUMBER	<input checked="" type="checkbox"/>
AGUA_POR_R	NUMBER	<input checked="" type="checkbox"/>
AGUA_POR_1	NUMBER	<input checked="" type="checkbox"/>
SERVICIO_H	NUMBER	<input checked="" type="checkbox"/>
LUZ_ELECTRICA	NUMBER	<input checked="" type="checkbox"/>
ILUMINACION	NUMBER	<input checked="" type="checkbox"/>
TOTAL_VIVIENDAS	NUMBER	<input checked="" type="checkbox"/>
PERSONAS_ECO_ACT	NUMBER	<input checked="" type="checkbox"/>
PERSONAS_MAYOR_5	NUMBER	<input checked="" type="checkbox"/>
PERSONAS_PRIMARIA	NUMBER	<input checked="" type="checkbox"/>
PERSONAS_SECUNDARIA	NUMBER	<input checked="" type="checkbox"/>
PERSONAS_SUPERIOR	NUMBER	<input checked="" type="checkbox"/>

Ilustración 28: Campos Tabla Inec_Total

Fuente: Propia

Se creó la estructura de manejo de zonas censales, la cual consta de su id propio y de descripción para poder identificar fácilmente como se muestra en la Ilustración 29.

Name	Datatype	Not Null
ID_ZONA	VARCHAR2(13)	<input type="checkbox"/>
DESCRIPCION	VARCHAR2(500)	<input checked="" type="checkbox"/>

Ilustración 29: Campos Tabla Inec_Zonas

Fuente: Propia

El modelo entidad relación final, se puede observar en la Ilustración 30, la cual nos muestra todas las tablas con sus respectivas relaciones.

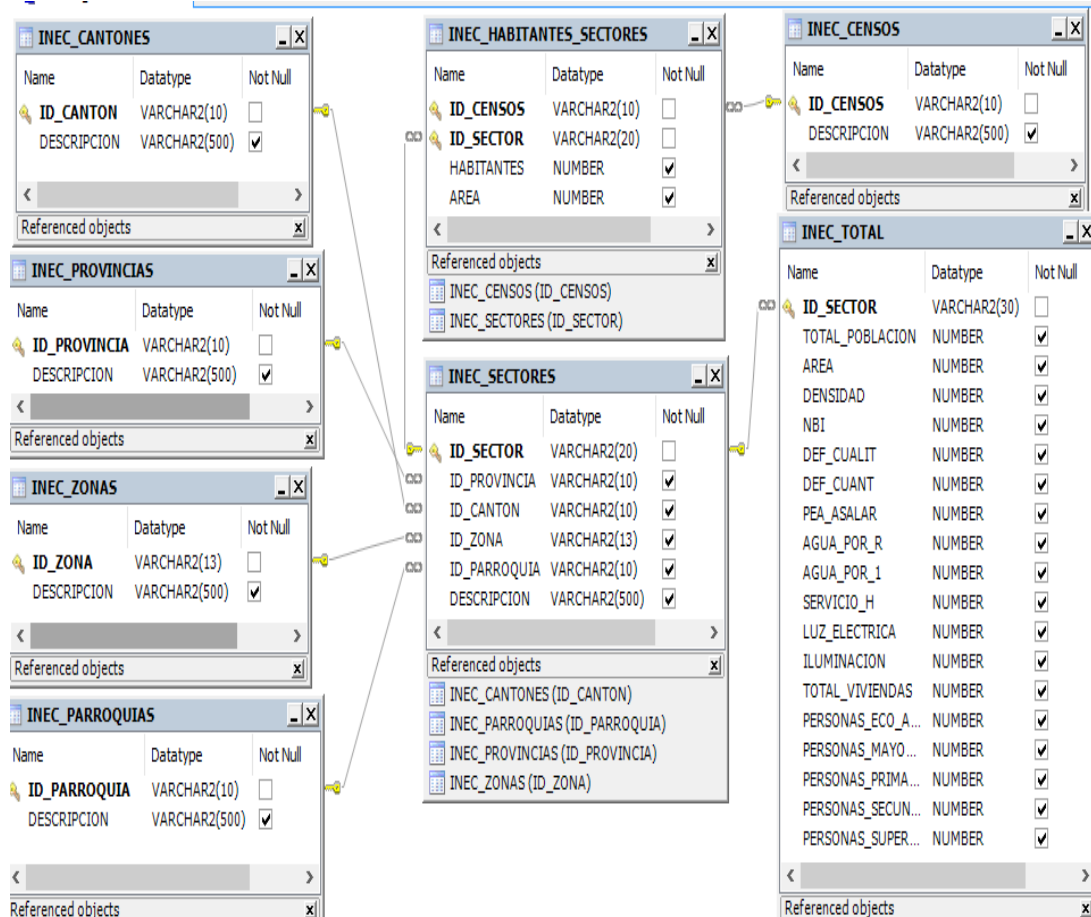


Ilustración 30: Diagrama Entidad Relación INEC

Fuente: Propia

3.7.2.3 Datos Gobierno Autónomo Descentralizado San Miguel de Ibarra (Catastro Urbano)

En el caso del Municipio de Ibarra, la información proporcionada fue la relacionada con el catastro urbano, la cual fue entregada en archivos Excel, tomando en cuenta cada predio debidamente geo-referenciado. Esta información fue unida directamente en la etapa de transformación y carga de información, como se

muestra en la Ilustración 31, la información fue registrada en la tabla INEC_TOTAL.

DPA_ZONA	nro_predio	clave_cata	tipo	area_total_m2	valor_terr	area_const	valor_edif	valor_como_corregido
100150001	002	100103031113002000	Urbano	1427,08	56583,01	43,27	7746,60	64329,61
100150001	008	100102020903008000	Urbano	150	966,74	40,00	4767,68	5734,42
100150001	003	100102020702003000	Urbano	270,940002	5500,92	39,02	9366,24	14867,17
100150001	013	100102020704013000	Urbano	333,390015	6837,23	55,60	15436,90	22274,12
100150001	019	100102020712019000	Urbano	266,1	5457,23	62,88	13689,25	19146,48
100150001	051	100103031334051000	Urbano	474,75	19775,38	28,20	6309,23	26084,61
100150001	006	100103031318006000	Urbano	1695,45	112747,43	484,59	161074,96	273822,38
100150001	002	100103031127002000	Urbano	7473,569824	400675,27	439,30	84399,81	485075,08
100150001	040	100103031331040000	Urbano	350	13461,00	203,52	77892,78	91353,78
100150001	021	100103031101021000	Urbano	932,5	64622,25	148,05	46860,73	111482,98
100150001	008	100103031401008000	Urbano	218,169998	4429,53	41,42	11219,53	15649,06
100150001	018	100103031323018000	Urbano	1467,23	102706,10	173,28	57888,38	160594,48
100150001	018	100103031404018000	Urbano	270	5371,11	62,94	17218,96	22590,07
100150001	015	100103031318015000	Urbano	569,75	22590,30	129,17	43120,41	65710,71
100150001	005	100103031009005000	Urbano	2847,13	112887,28	160,17	57575,43	170462,71
100150001	053	100103031334053000	Urbano	242,01	10182,57	88,42	18916,09	29098,66
100150001	006	100103031331010000	Urbano	702,809998	27866,07	141,80	59753,54	87619,61
100150001	006	100103031305006000	Urbano	11016	223658,95	205,42	51580,66	275239,61
100150001	027	100102020706027000	Urbano	278,660004	5657,66	46,16	17174,99	22832,65
100150001	009	100102020722009000	Urbano	303,799988	6168,08	99,70	35751,33	41919,41
100150001	001	100102020903001000	Urbano	150	976,50	109,44	40610,12	41586,62
100150001	007	100102020903007000	Urbano	150	966,74	53,40	9807,09	10773,82

Ilustración 31: Datos Catastro Urbano

Fuente: Propia

3.7.3 Definición de los requerimientos del negocio

En este análisis el objetivo principal es la identificación de las necesidades de cada uno de los interesados en el desarrollo de esta solución tecnológica, con la cual se pueda visualizar resultados y poder tomar decisiones en cada una de las áreas.

Se han tomado en cuenta los requerimientos principales de las dos entidades interesadas en este desarrollo que son EMAPA y el Proyecto Ibarra Verde, a continuación detallamos cada uno de los requerimientos por institución, los requerimientos serán valorados por la prioridad del 1 al 10:

3.7.3.1 Requerimientos EMAPA

Tabla 12: Requerimientos EMAPA

Número	Requerimiento	Descripción	Prioridad
RF1	Análisis de Perdidas	Generar un reporte de análisis, en donde se pueda visualizar la producción de cada tanque y su respectivo consumo proyectado en la sumatoria de todos los suministros que se relacionan con el tanque.	9
RF2	Análisis de Consumo por Tarifa	Generar un reporte de análisis, en donde se pueda visualizar la comparativa anual en los consumos por tarifa para identificar cuáles son las tarifas con mayor consumo.	6
RF3	Análisis por rango de consumo a nivel de suministro	Generar un reporte de análisis, en donde se pueda visualizar los consumos de los suministros por rangos de consumos.	8
RF4	Análisis de historial de consumo	Generar un reporte de análisis, en donde se pueda visualizar y analizar el consumo histórico de un suministro en específico.	7
RF5	Análisis de consumo por planes	Generar un reporte de análisis, en donde se pueda visualizar el consumo por plan de distribución.	6

Fuente: Propia

3.7.3.2 Requerimientos Proyecto Ibarra Verde

Tabla 13: Requerimientos Proyecto Ibarra Verde

Número	Requerimiento	Descripción	Prioridad
RF6	Unión de datos.	Realizar la unión de datos de EMAPA, con los datos de INEC, CATASTRAL.	10
RF7	Análisis de consumo por sector censal.	Generar un reporte de análisis, en donde se pueda analizar el consumo de cada sector censal por año, mes, tarifa y rango de consumo.	7
RF8	Análisis de consumo por persona a nivel de zona censal.	Generar un reporte de análisis, en donde se pueda analizar el consumo por persona diaria de agua potable.	8
RF9	Generar una base de suministros considerados muestra de estudio	Generar una base de suministros, que serán la base de estudio para los cruces de información de consumo de agua potable con las demás variables obtenidas de las demás instituciones.	

Fuente: Propia

3.7.3.3 Perspectiva de la herramienta

Los responsables del Proyecto Ibarra Verde serán los responsables de gestionar el aprovisionamiento toda la información para que la herramienta cumpla con su objetivo principal de implementación, cabe resaltar que la información del INEC y de catastro no se actualiza con mayor frecuencia, no así como la de EMAPA, esta información se genera diariamente.

Se requiere que el Proyecto Ibarra Verde cuente con la infraestructura necesaria para que una vez presentada y aceptada la herramienta desarrollada, se haga cargo en su totalidad de la administración y utilización de esta.

3.7.3.4 Descripción de los requerimientos

Análisis de Perdidas

Tabla 14: Análisis de Perdidas

Identificación del Requerimiento	RF1
Nombre del Requerimiento	Análisis de Perdidas
Características	Se tomaran en cuenta todos los tanques de abastecimiento que cuenta la ciudad para el abastecimiento del agua potable.
Descripción del Requerimiento	<p>La herramienta dará una visualización del porcentaje de pérdida total de uno o varios circuitos de abastecimiento. Los circuitos de abastecimientos están detallados por cada tanque de provisión de agua. Cada suministro o medidor de agua debe tener la identificación a que circuito de distribución pertenece, en el caso de no tener este identificador se lo pondrá tomando en cuenta la ubicación geográfica y sobre todo el plan, ruta y sector al que pertenece.</p> <p>Para este análisis se lo realizara tomando en cuenta cuanto fue la producción de un circuito de distribución y comparándolo con el total de consumo de todos los suministros atados al circuito.</p> <p>Este requerimiento ayudará a mejorar el abastecimiento y sobre todo a reducir pérdidas y hurtos de agua, denominada agua no contabilizada.</p>
Prioridad del Requerimiento	Alto

Fuente: Propia

Tabla 15: Análisis de Consumo por Tarifa

Identificación del Requerimiento	RF2
Nombre del Requerimiento	Análisis de Consumo por Tarifa
Características	Se tomaran en cuenta todas las tarifas existentes actualmente en EMAPA.
Descripción del Requerimiento	La herramienta dará una visualización del porcentaje de consumo por cada tarifa existente, los reportes se los podrá obtener con la dimensión de tiempo año y mes. También dará la información sobre el porcentaje de suministros existentes por cada tarifa.
Prioridad del Requerimiento	Medio

Fuente: Propia

Tabla 16: Análisis por rango de consumo a nivel de suministro.

Identificación del Requerimiento	RF3
Nombre del Requerimiento	Análisis por rango de consumo a nivel de suministro
Características	Se tomaran en cuenta todos los suministros existentes.
Descripción del Requerimiento	La herramienta dará una visualización de un listado de suministros que se filtraran tomando en cuenta varios parámetros, el principal será el rango de consumo. Este análisis permitirá identificar suministros que consuman sobre los parámetros considerados normales para cada tarifa, con el objetivo de identificar

	suministros que tengas tarifas diferentes al objetivo de consumo.
Prioridad del Requerimiento	Alto
<i>Fuente: Propia</i>	

Tabla 17: Análisis de historial de consumo.

Identificación del Requerimiento	RF4
Nombre del Requerimiento	Análisis de historial de consumo
Características	El análisis se lo realizará por suministro
Descripción del Requerimiento	La herramienta dará una visualización del historial de consumo de un suministro específico, en el cual se mostrara los valores de consumo mensual.
Prioridad del Requerimiento	Medio
<i>Fuente: Propia</i>	

Tabla 18: Análisis de consumo por planes.

Identificación del Requerimiento	RF5
Nombre del Requerimiento	Análisis de consumo por planes
Características	El análisis se lo realizará de todos los suministros
Descripción del Requerimiento	La herramienta dará una visualización de la comparativa de consumo de cada uno de los planes existentes, el análisis se lo podrá realizar con diversos filtros como año, mes y tarifas.
Prioridad del Requerimiento	Medio
<i>Fuente: Propia</i>	

Tabla 19: Unión de datos.

Identificación del Requerimiento	RF6
Nombre del Requerimiento	Unión de datos
Características	El análisis debe tomar en cuenta la diversa información existente
Descripción del Requerimiento	La herramienta dará la facilidad de obtener análisis uniendo la diversidad de información que se tiene.
Prioridad del Requerimiento	Alta

Fuente: Propia

Tabla 20: Análisis de consumo por sector censal

Identificación del Requerimiento	RF7
Nombre del Requerimiento	Análisis de consumo por sector censal
Características	El análisis debe tomar en cuenta todos los suministros y sectores censales
Descripción del Requerimiento	La herramienta dará la facilidad de obtener análisis de consumo de todas las áreas censales, en la cual se detallara las tarifas respectivas.
Prioridad del Requerimiento	Media

Fuente: Propia

Tabla 21: Análisis de consumo por persona a nivel de zona censal.

Identificación del Requerimiento	RF8
Nombre del Requerimiento	Análisis de consumo por persona a nivel de zona censal
Características	El análisis debe tomar en cuenta todos los suministros y zonas censales.
Descripción del Requerimiento	La herramienta dará la facilidad de obtener análisis de consumo de consumo por persona a nivel de zonas y por litros diarios.
Prioridad del Requerimiento	Media

Fuente: Propia

Tabla 22: Obtener la base de suministros considerando muestra de estudio

Identificación del Requerimiento	RF9
Nombre del Requerimiento	Generar una base de suministros considerados muestra de estudio
Características	Debe ser la base para realizar las comparativas y cruces de correlación con las diversas variables.
Descripción del Requerimiento	Aplicar los diversos filtros dados, con el fin de obtener una muestra, la cual será la base de estudio para poder realizar cruces de variables y poder obtener resultados de correlación.
Prioridad del Requerimiento	Media

Fuente: Propia

3.7.4 Capa ETL

Una vez identificados los requerimientos se procedió al diseño multidimensional de los cubos a utilizar, estos cubos permitirán realizar la extracción de información desde el origen de datos descrito anteriormente. La extracción se la realizará directamente de la base de datos relacional desarrollada exclusivamente para la implementación de esta herramienta, en la cual se une información de varias entidades como de detalle anteriormente, esta información permitió crear las tabas de hechos y dimensiones correspondientes.

A continuación se describe la estructura de cada uno de los cubos utilizados para el desarrollo de esta herramienta.

En la Ilustración 32, se muestra el esquema conceptual para el análisis de datos de consumo de agua potable.

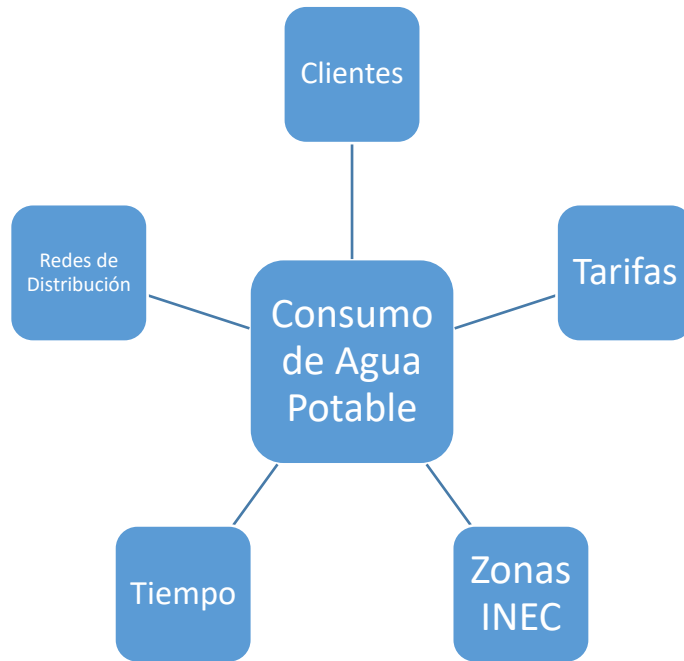


Ilustración 32: Esquema conceptual consumo agua potable

Fuente: Propia

En la Ilustración 33, se observa el esquema conceptual para el análisis de la producción de agua potable.

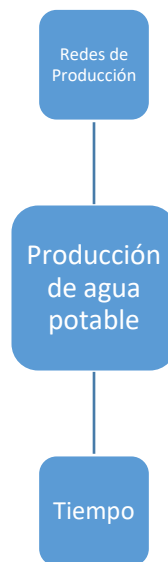


Ilustración 33: Esquema conceptual generación agua potable

Fuente: Propia

3.7.5 Modelado Dimensional

Vamos a definir los datos que permite obtener claramente qué y cómo vamos a mostrar los datos, para que cumplan con las especificaciones descritas en los requerimientos.

Nos ayuda a definir la granularidad de los datos establecidos en los requerimientos de desarrollo, con el objetivo de cumplir en lo mayor posible el detalle que permitirá entregar de mejor manera la información perteneciente a cada análisis.

Realizaremos la selección de cada elemento

- DataMart
- Dimensiones
- Hechos

DataMart.- como lo hemos citado anteriormente este trabajo va enfocado al análisis de información de varias instituciones, y de forma específica de EMAPA, mismo que contribuirá a tener un Data Warehouse consolidado de todas las instituciones participantes.

Dimensiones.- se ha definido las dimensiones que utilizaremos para los análisis previstos.

Hechos.- definido por una transacción o evento que se realiza, estas serán representadas con una colección de partes de datos.

Presentamos las dimensiones utilizadas en este desarrollo y describiendo la granularidad de cada una:

En la Ilustración 34 se observa la dimensión tiempo utilizada para esta implementación.

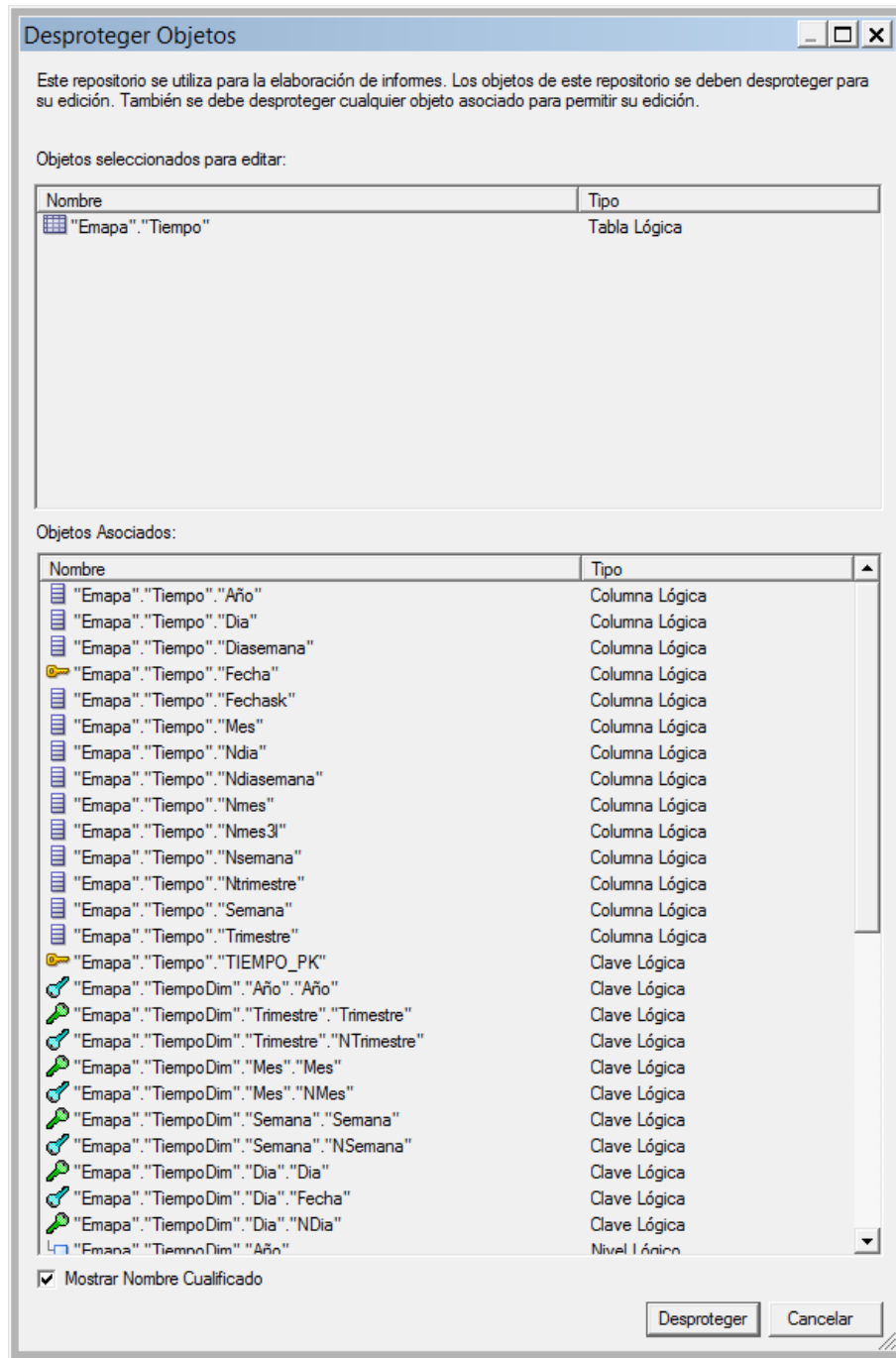


Ilustración 34: Dimensión Tiempo

Fuente: Propia

En la Ilustración 35 se muestra la dimensión tarifa, en la cual se describe cada una de las tarifas con las cuales cuenta EMAPA y podemos realizar análisis.

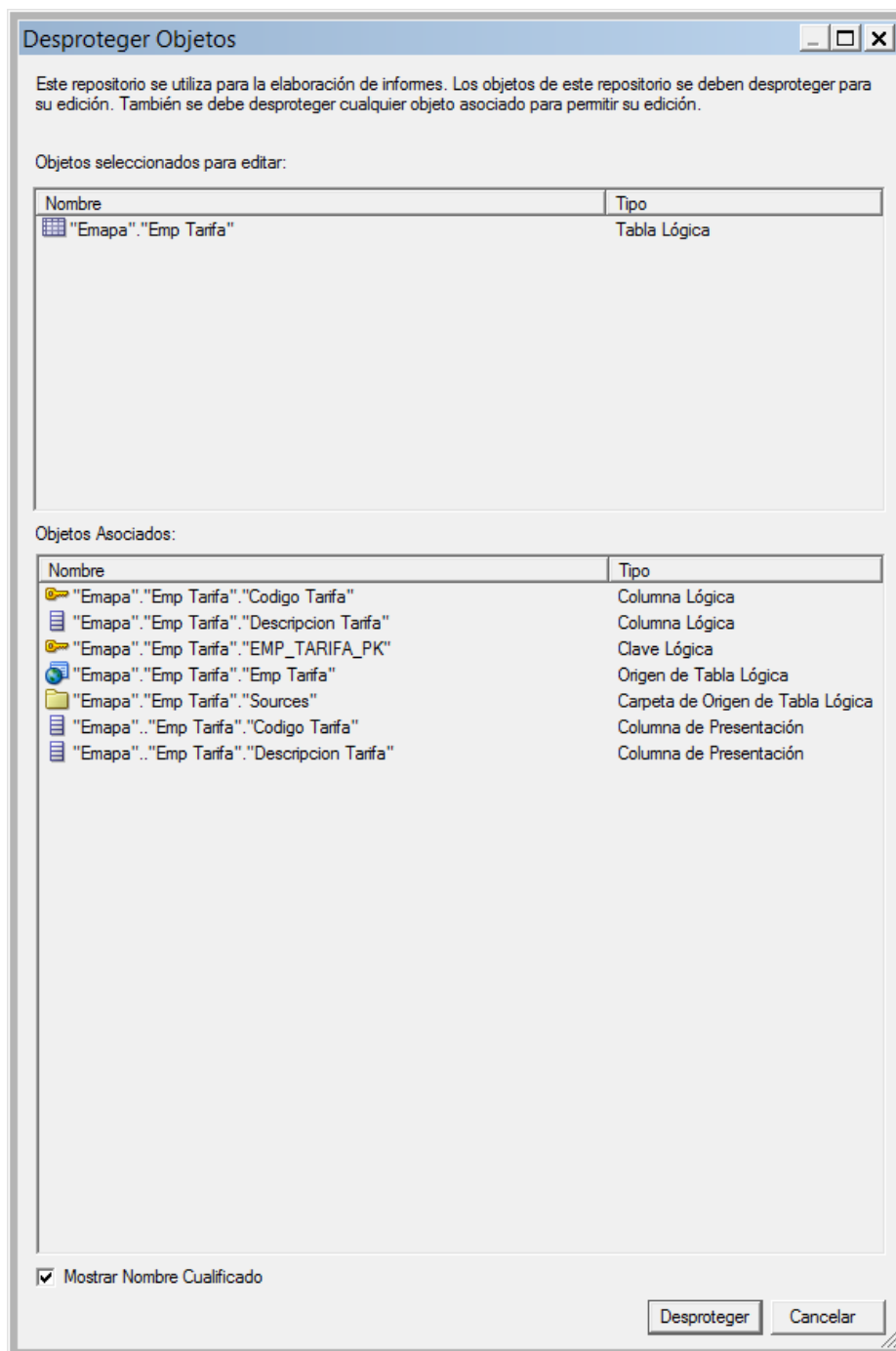


Ilustración 35: Dimensión Tarifa

Fuente: Propia

En la Ilustración 36, se observa la tabla de dimensiones Redes, la cual consta con la información de las redes de comercialización para los respectivos análisis.

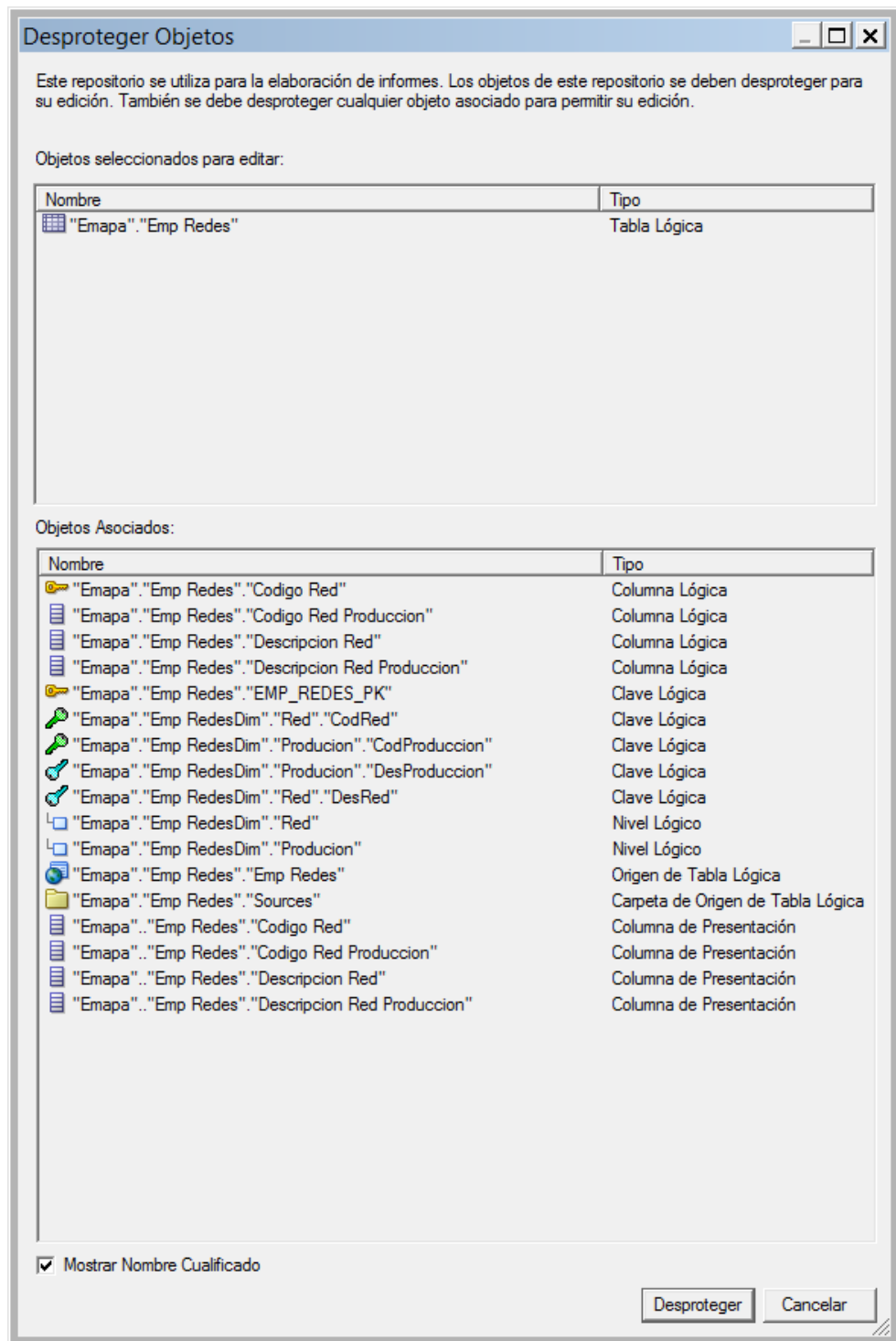


Ilustración 36: Dimensión Redes

Fuente: Propia

En la Ilustración 37, se observa la dimensión perteneciente a la información INEC y de catastro, cual nos facilitara información para los análisis correspondientes.

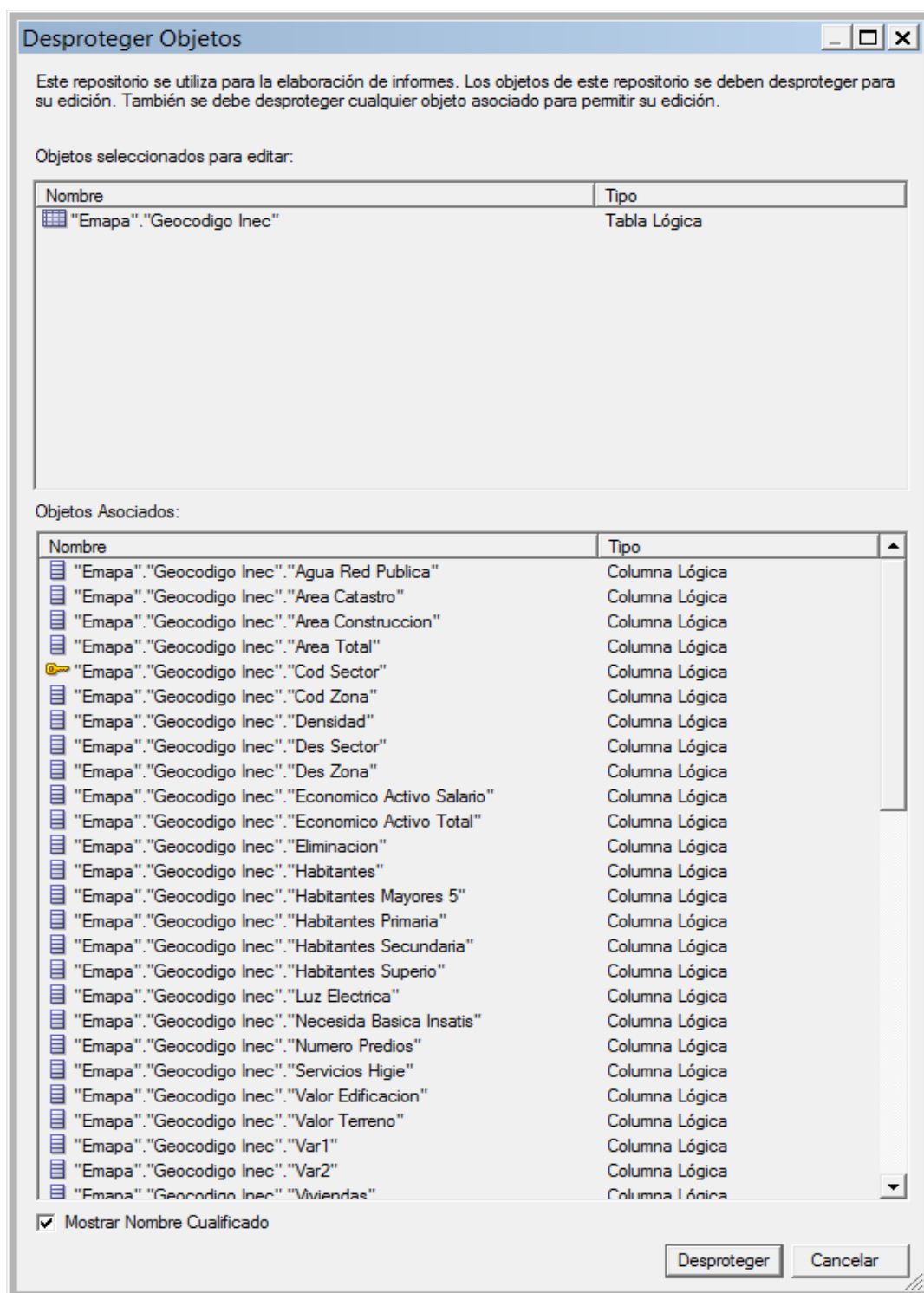


Ilustración 37: Dimensión Geo INEC

Fuente: Propia

En la Ilustración 38, se observa la dimensión perteneciente a los clientes, la cual contiene la información necesaria para los diferentes análisis.

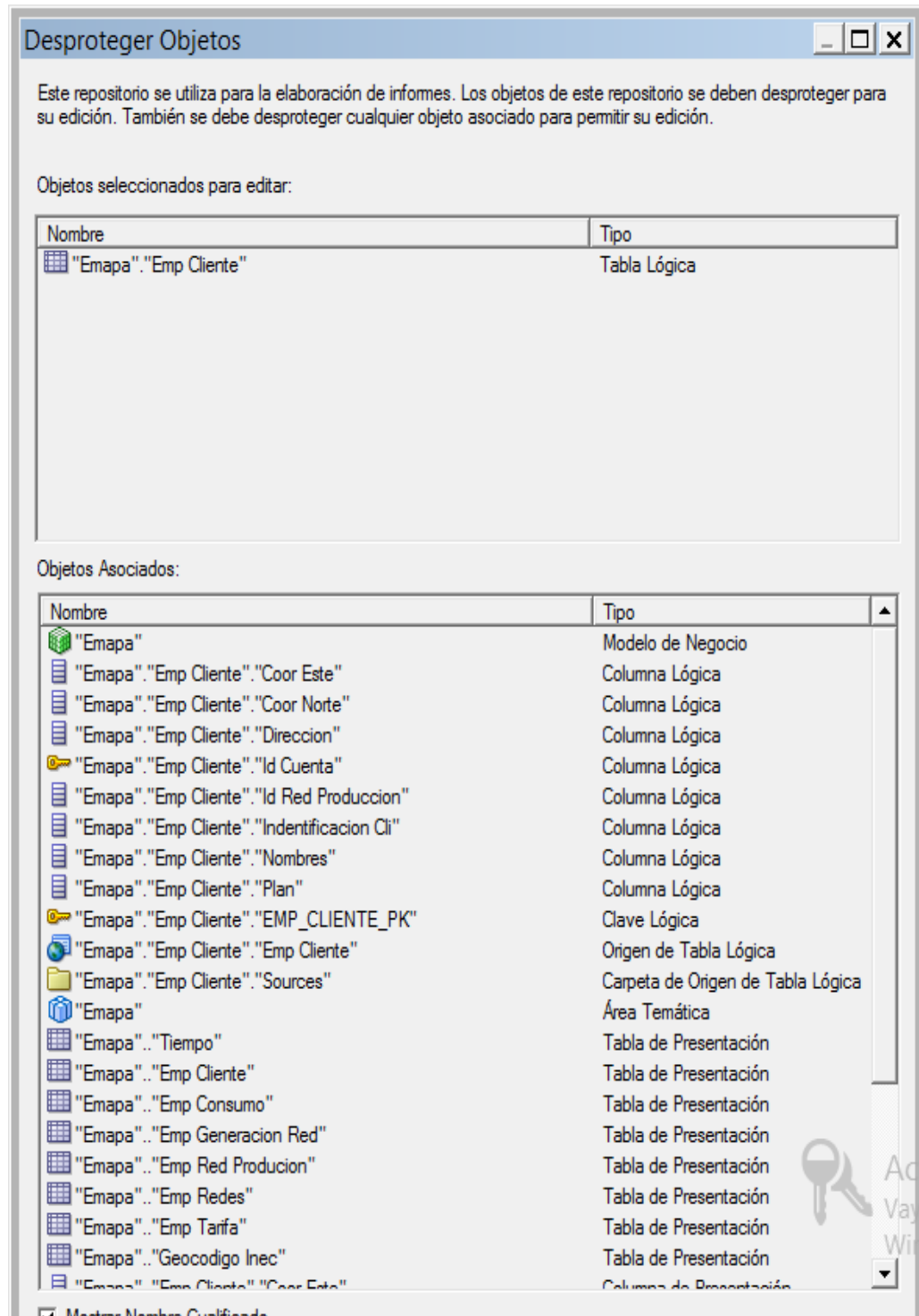


Ilustración 38: Dimensión Clientes

Fuente: Propia

En la Ilustración 39, se observa la información perteneciente a la dimensión Red Producción, en la cual contiene la información para los análisis de producción de agua potable.

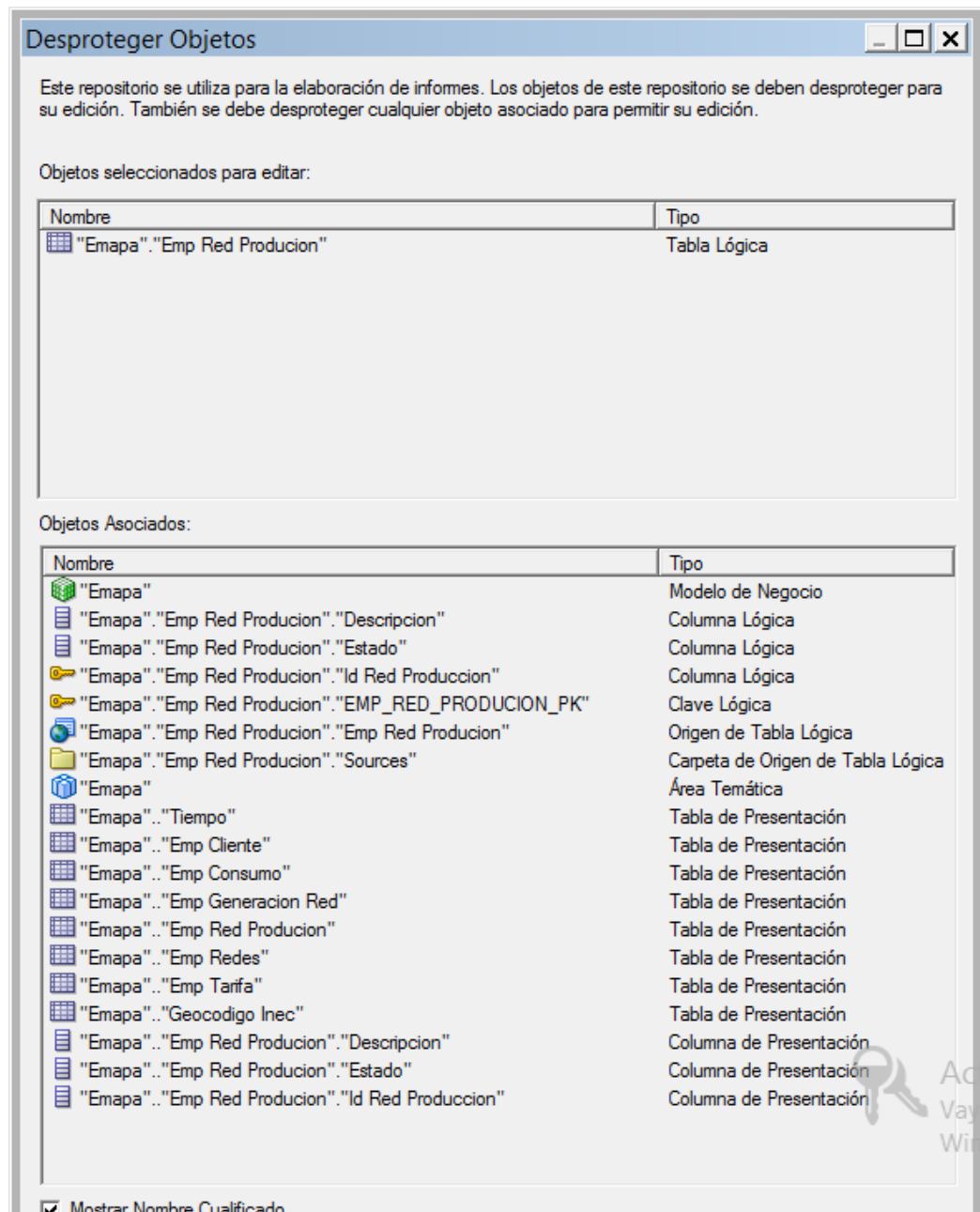


Ilustración 39: Dimensión Redes Producción

Para realizar los respectivos análisis se debe tener la tabla de hechos, como se muestra en la Ilustración 40, se tiene la tabla de hechos de consumo, en la cual se tiene una función de agregado tanto en consumo como en el id cuenta.

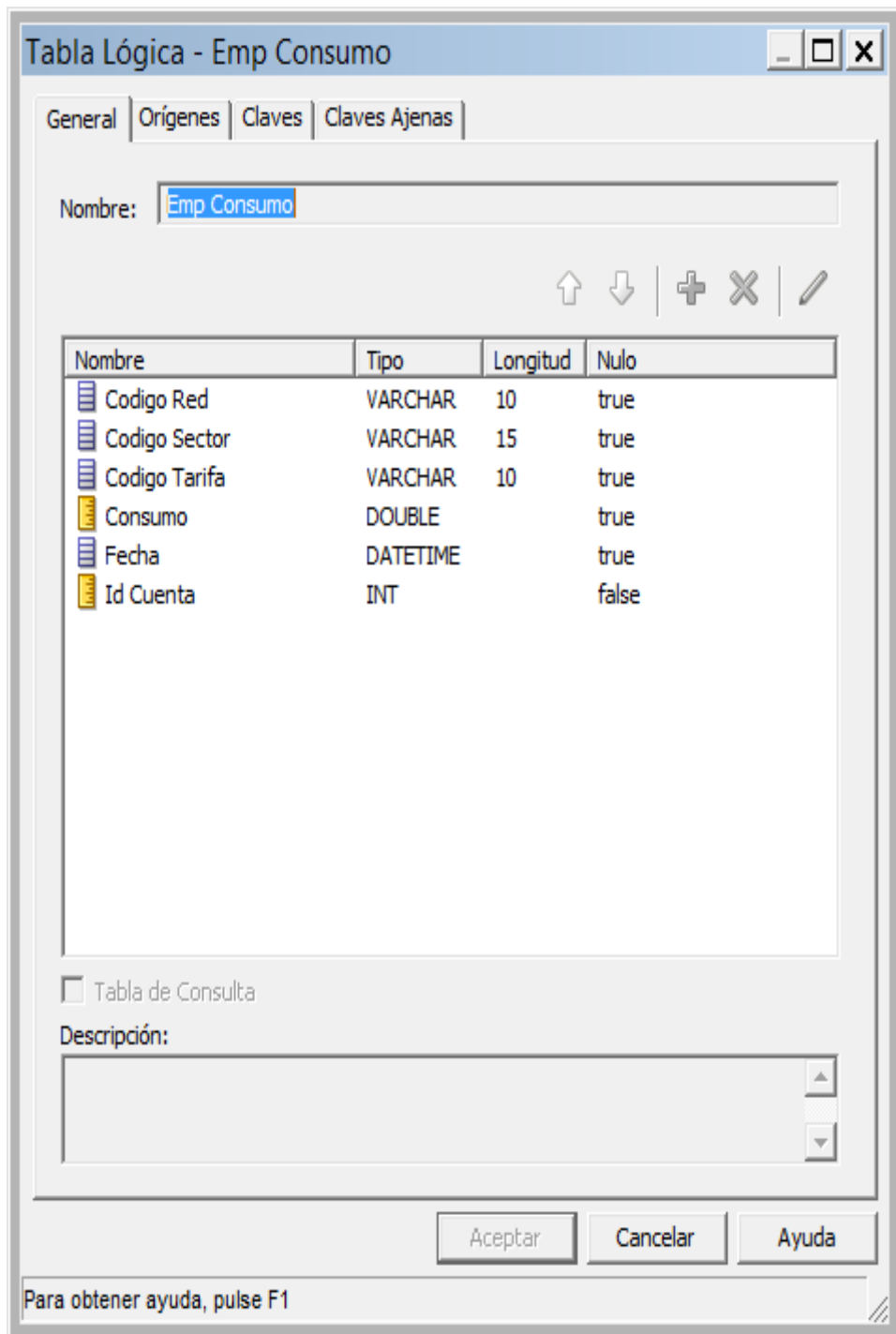


Ilustración 40: Tabla de hechos Consumo

Fuente: Propia

Para el análisis de producción, se creó a tabla de hechos Generación Red como se muestra en la Ilustración 41, en la cual se utilizó la función de agregado para el campo producción.

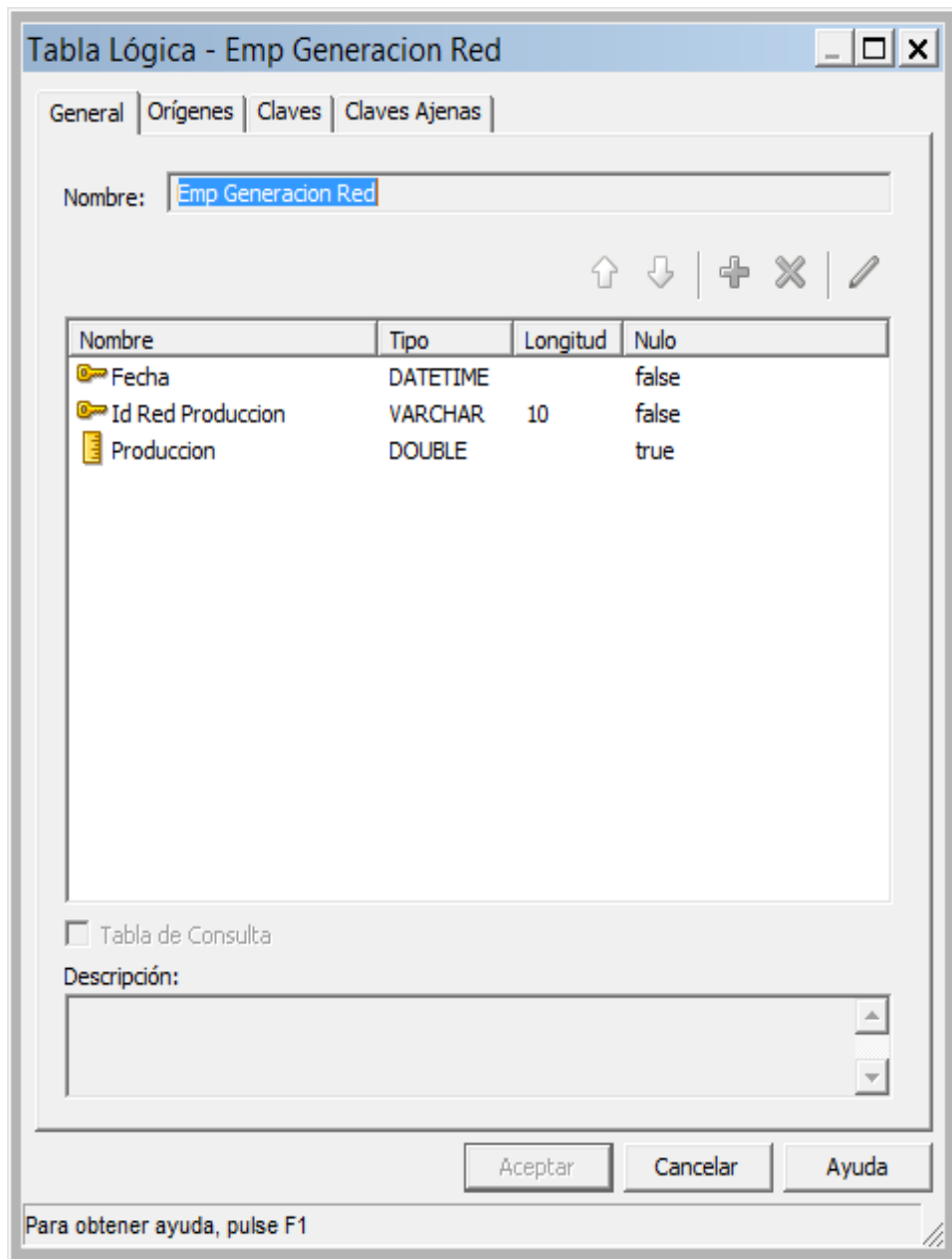


Ilustración 41: Tabla de hechos Generación

Fuente: Propia

3.7.6 Modelo físico

Presentamos el modelo de la base de datos relacional que se implementó para el desarrollo de esta propuesta, como se detalla en la Ilustración 42.

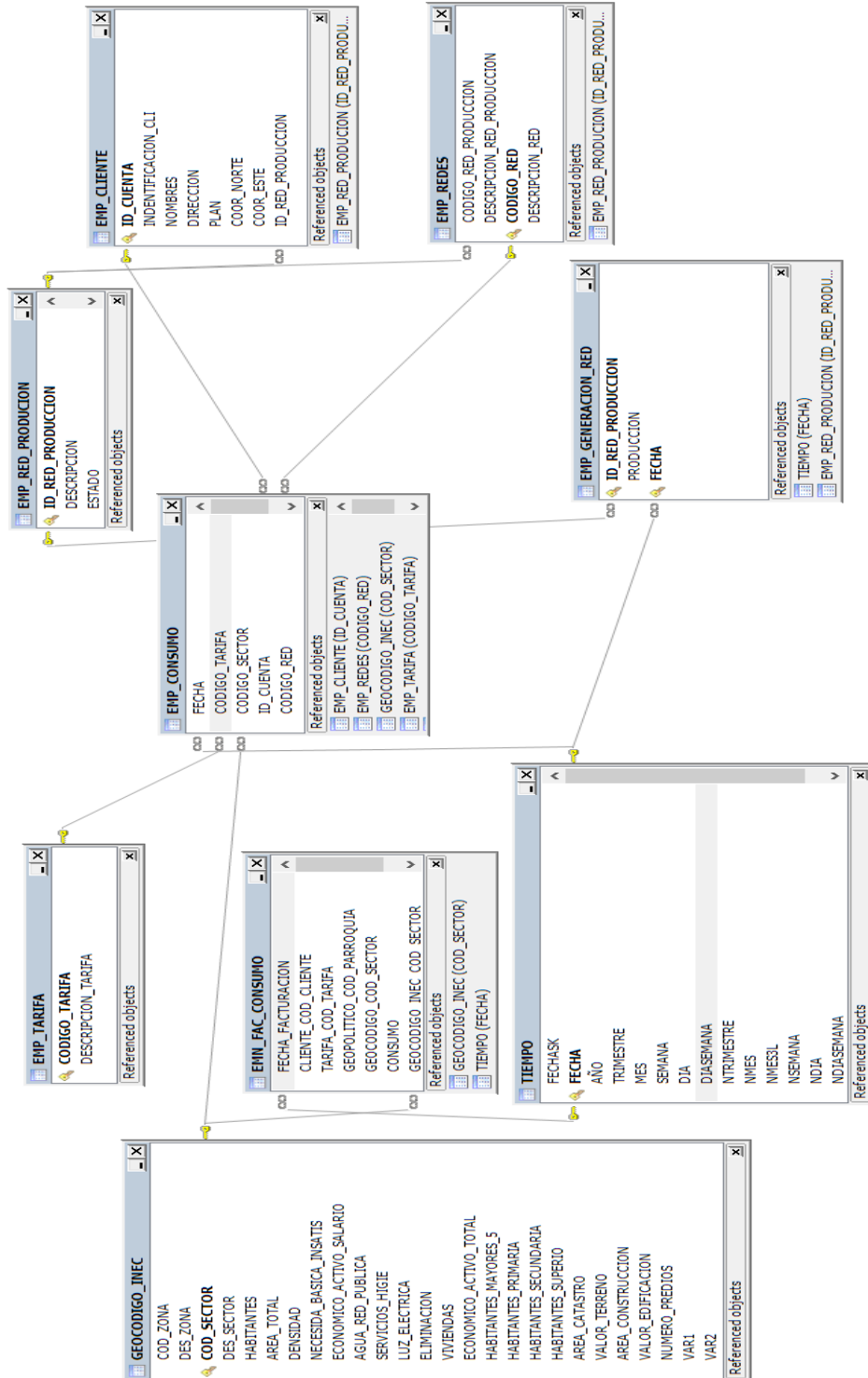


Ilustración 42: Modelo relacional

Fuente: PropiaCapa OLAPAdministración Data Mart

Para realizar administración correspondiente de las entidades perteneciente a nuestro repositorio (Data Mart), utilizamos la herramienta de administración de Oracle BI, la cual permite unificar los procesos de presentación de datos, a continuación presentamos el modelo de metadatos implementado, como se muestra en la Ilustración 43.

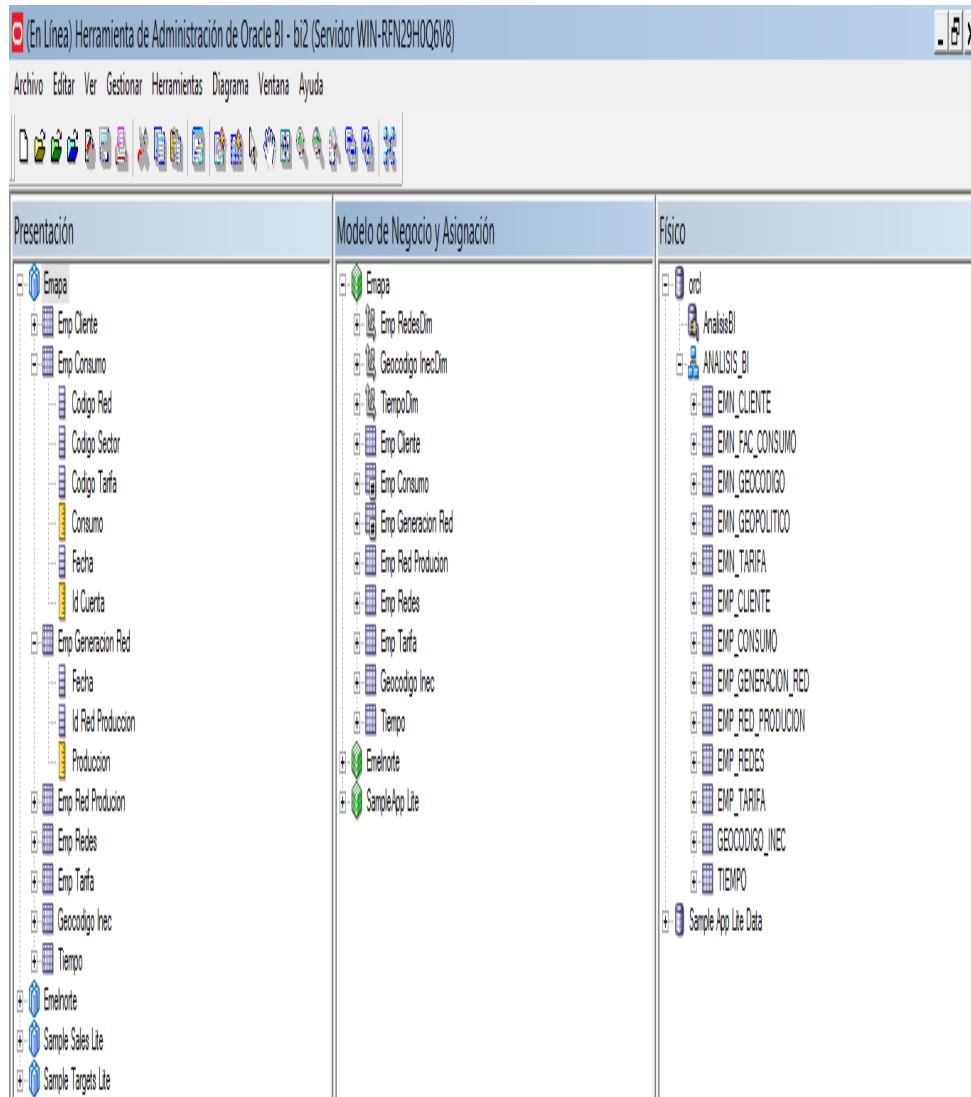


Ilustración 43: Herramienta de Administración Oracle BI

Fuente: Propia

3.7.7 Implementación de la solución BI

En la Ilustración 44 se presenta el modelo estrellas utilizados para el análisis de información con respecto al consumo de agua potable.

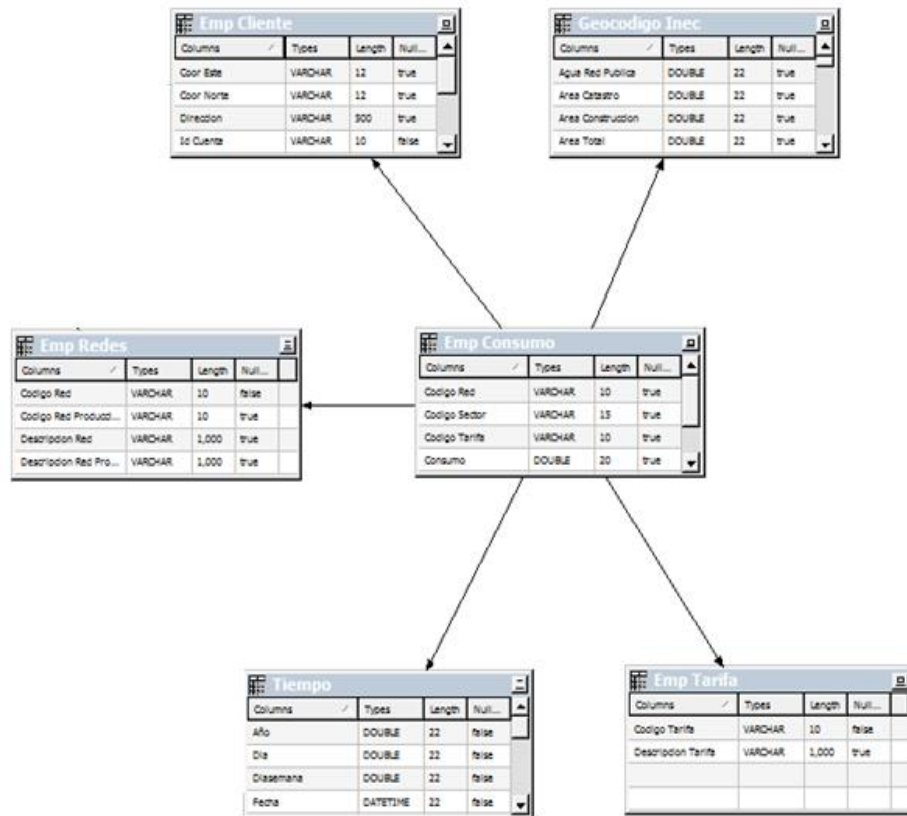


Ilustración 44: Modelo estrella Consumo

Fuente: Propia

En la Ilustración 45, se muestra el modelo estrella para el análisis de información con respecto a la producción a agua potable, al unir los dos modelos dimensionales se puede analizar datos de pérdidas de agua potable.



Ilustración 45: Modelo estrella Generación

Fuente: Propia

3.7.8 Especificación de aplicaciones BI

Se desarrolló varios reportes interactivos los cuales fueron incorporados posteriormente a un Dashboard para facilitar el manejo de los usuarios finales, en los cuales la presentación de información será de fácil comprensión.

En la Ilustración 46, se aprecia el análisis de pérdidas de agua potable, en la cual se basa en la comparación del agua distribuida por cada una de las plantas o tanques de distribución, con el total comercializado.

En el análisis mostrado se tiene como resultado que el promedio de pérdida total de los tres años de estudio fue de 35% es decir que de los 51 millones de metros cúbicos producidos tan solo 33'150.000 fueron comercializados, el restante es considerado como perdidas ya sea por hurtos o fallas en el sistema de

distribución. Se debe considerar y realizar una investigación específica de las plantas que superan el promedio, y sobre todo la planta Guaraczapas que muestra un nivel de pérdida mayor al 60%, es decir que de cada 100 m³ menos de 40m³ son comercializados.

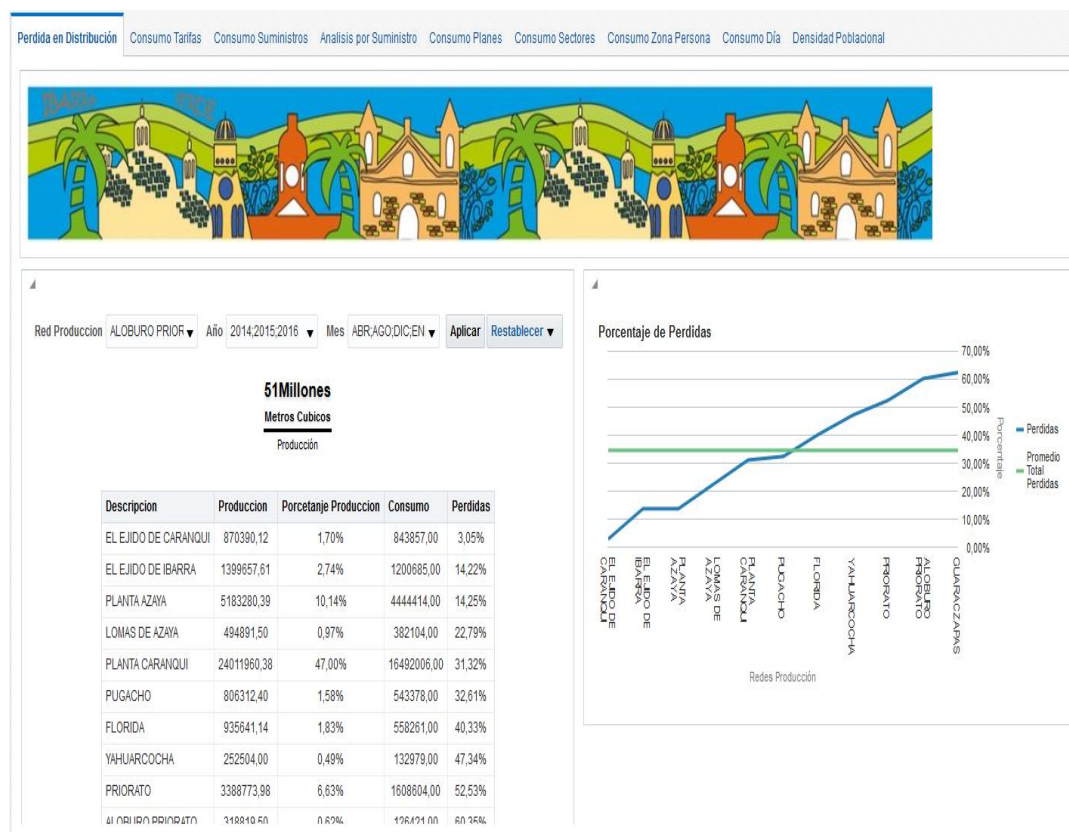


Ilustración 46: Análisis de pérdidas

Fuente: Propia

En la Ilustración 47, se puede visualizar el análisis de datos por rango de consumo, que nos permite identificar los suministros que están dentro de un rango de consumo definidor por el usuario.

En el caso específico citado se puede observar que se encontraron más de 4 suministros que su consumo es superior a 500 m³ al mes con tarifa residencial cuidada, este consumo se puede considerar alto en relación al sugerido por la OMS que es de 100 litros diarios por persona, en el caso específico se recomendó

realizar una inspección a los suministros citados para realizar una comprobación de la utilización de agua potable.

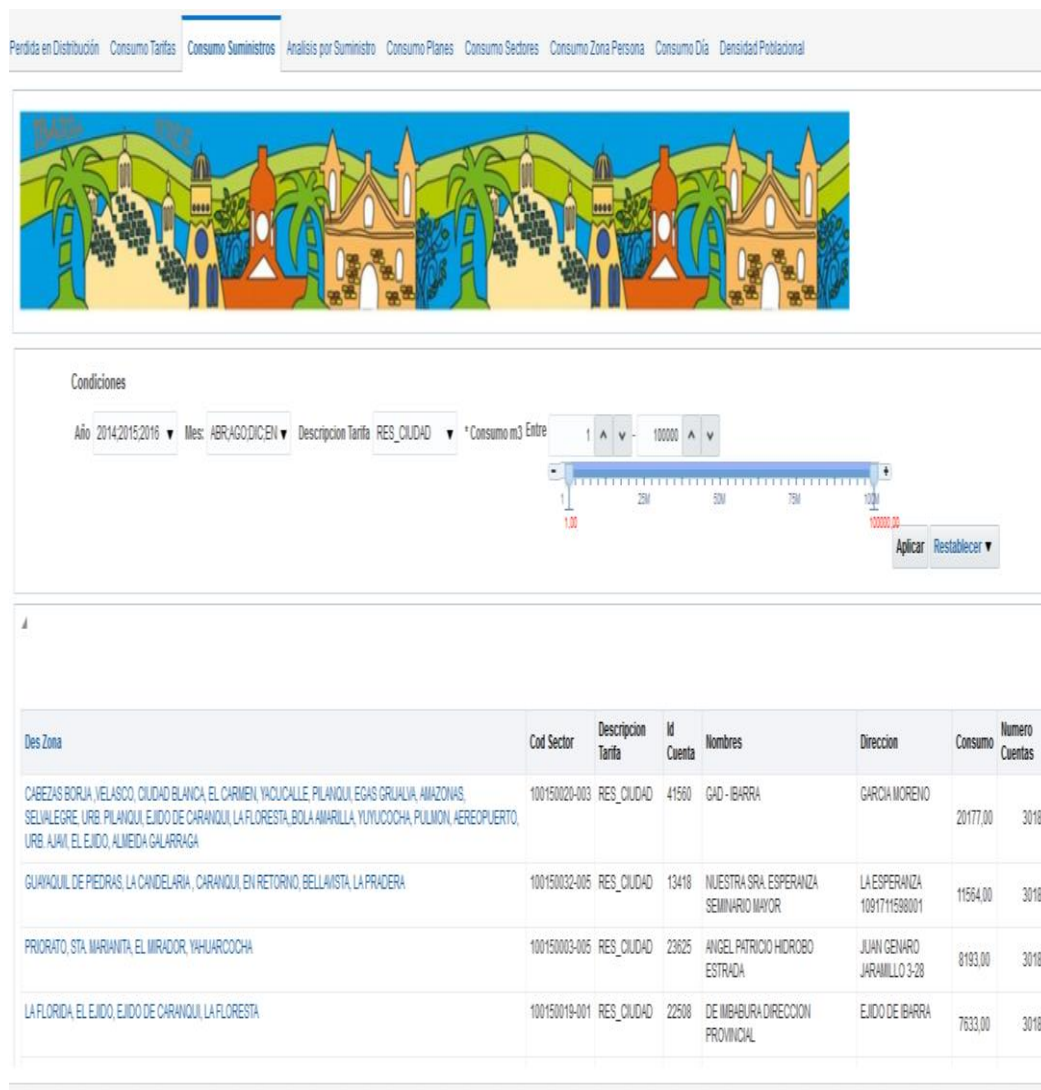


Ilustración 47: Análisis por rango de consumo

Fuente: Propia

En la Ilustración 48, se visualiza el análisis de densidad poblacional, en la cual se toma el promedio mensual de consumo y la densidad poblacional (hec/hab).

En la cual se puede visualizar cada una de las zonas de estudio, realizando un cruce de información entre el consumo promedio residencial y la densidad poblacional, se puede identificar que a mayor densidad poblacional el promedio del consumo de agua potable aumenta, esta información nos será útil para la realización de un análisis más profundo con la utilización de herramientas

estadísticas como es el SPSS. Como conclusión del análisis se recomendó analizar los casos especiales que no se ajustan a la correlación, para identificar posibles fallas en la información utilizada.

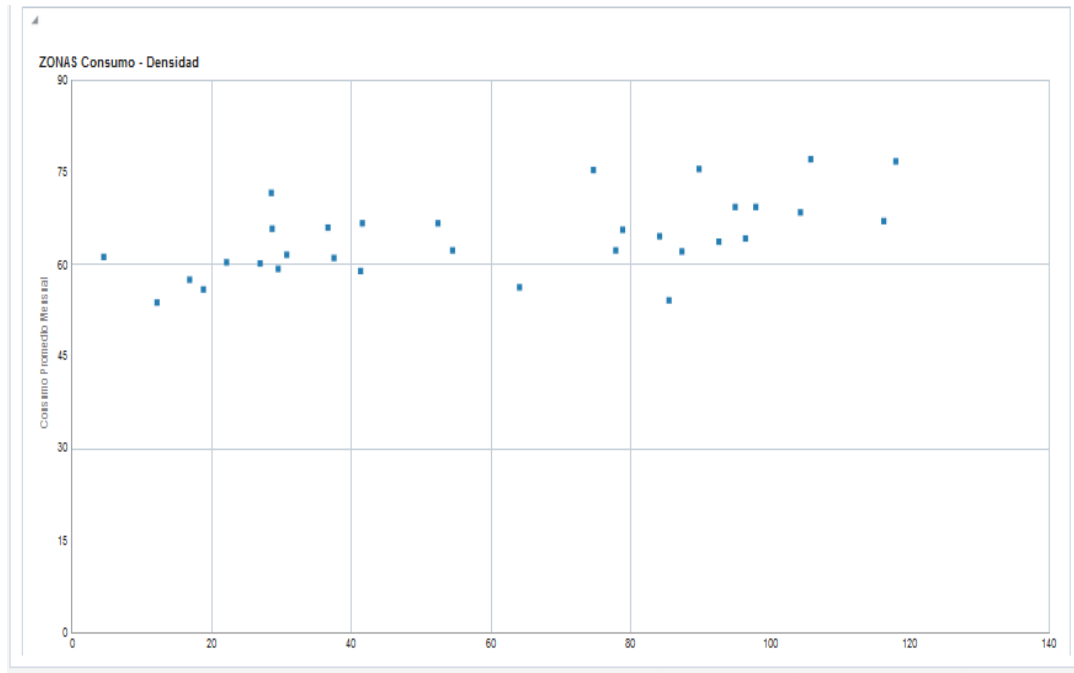


Ilustración 48: Análisis de Densidad Poblacional

Fuente: Propia

3.7.9 Mantenimiento y Crecimiento del DW

Al ser una herramienta de gran impacto y utilidad a nivel institucional tanto para EMAPA como para Ibarra Verde, los trabajos de mejora serán continuos siempre y cuando se siga contando con el apoyo fundamental de las entidades que proveen de los datos parte esencial de los análisis.

CAPITULO IV

RESULTADOS, CONCLUSIONES Y RECOMENDACIONES

4 Introducción

En este capítulo detallaremos los resultados de un análisis de cruce de información entre variables, variables independientes como son consumo de agua potable, densidad poblacional y habitantes hogar de cada sector, con la variable dependiente de toma de decisiones con respecto a la información obtenida, para luego utilizar una herramienta de especializada en estadística y análisis de datos como es el SPSS, información que será clave y muy útil para el personal del proyecto Ibarra Verde.

Además se entregarán las conclusiones y recomendaciones del desarrollo de las herramientas enfocadas en los análisis realizados para cada una de las instituciones participantes.

4.1 Presentación de Resultados

Analizaremos un caso concreto desarrollado en conjunto con el grupo de trabajo del proyecto Ibarra Verde, en el cual desean conocer qué grado de correlación tiene la densidad poblacional de cada zona y sector urbano con respecto al consumo de agua potable, tomando en cuenta solo la tarifa residencial.

Para este análisis necesitamos el uso de información de tres instituciones, Municipio de Ibarra que proveerá la información del catastro en km² de cada zona y sector; INEC que proveerá la información de los habitantes, en este caso tenemos la población 2010 que pertenece al último censo nacional, para proyectar la población futura se aplicara la formula detallada en la tabla 23; Por ultimo tenemos la información de EMAPA de los consumos en m³.

Tabla 23: Formula de proyección poblacional

Formula de Población Futura	
$Pf = N * (1 + 0,0186)^a$	
Pf	Población Futura
N	Número de población 2010 INEC
0,0186	Constante de tasa de crecimiento
a	Años de proyección
^	Elevación de potencia

Fuente: Propia

4.1.1 Aplicación de Business Intelligence en el Análisis de Consumo y Distribución de Agua Potable.

En el país actualmente las empresas de Agua Potable y Alcantarillado no poseen sistemas Business Intelligence que ayuden a tomar decisiones para mejorar la distribución y reducir el consumo con la presente investigación y la aplicación de la herramienta Business Intelligence se pudo identificar varios factores que son contra productores para conservar el líquido vital.

Al analizar los datos solo de EMAPA, se pudo identificar que las pérdidas en distribución (agua no contabilizada) superaba el 30% en algunos casos, igualmente se pudo identificar suministros que reciben beneficios de consumos residenciales que superaban los 500m³ de consumo mensuales es decir 500.000 litros a mes, estos resultados para consumo residenciales era demasiado altos, ya que tomando en cuenta lo enunciado por la OMS el consumo diario de agua potable de una persona debería ser 100 litros diarios o como base de 3 a 4m³ al mes.

En el análisis conjunto con las variables de las otras entidades, se pudo identificar que el consumo promedio habitante día es de 136 litros es decir 4m³ por persona al mes, este resultado se pudo obtener cruzando la información de EMAPA con la del INEC 2010 proyecta a los años de estudio mediante la fórmula que se detalla en la tabla 23 de la presente investigación.

4.1.2 Definición de la muestra de estudio.

Lo primero que definimos son las reglas para tomar en cuenta solo los que cumplen las condiciones para este análisis, las reglas son las siguientes:

- Suministros que sean de tarifas consideradas residenciales, en este caso 3EDAD_CUIDAD, DISCAPACIDAD_CUIDAD, RES_CUIDAD, PRIORATO.
- Suministros que se encuentren en las zonas y sectores censales amanzanados urbanos, en este caso que se encuentren dentro de las zonas 100150001 a 1001500032, descripción de las zonas se muestra en la tabla 24.
- Suministros que se encuentra fuera de las áreas de mercados.
- Suministros que cumplan con el mínimo promedio de consumo propuestos por la OMS que es de 100 litros diarios, que transformados en consumo anual es de 3m³ mensuales.

Al aplicar todos los filtros detallados anteriormente ya tenemos la muestra de suministros a utilizar en los análisis propuestos.

Tabla 24: Sectores INEC

Codificación	Descripción Zona
100150001	COLINAS DE IBARRA, HUERTOS FAMILIARES, AZAYA NORTE
100150002	PRIORATO
100150003	PRIORATO, STA. MARIANITA, EL MIRADOR, YAHUARCOCHA
100150004	EL OLIVO, LA QUINTA, STA. MARIANITA, EL MIRADOR, OLIVO ALTO, LA VICTORIA
100150005	SAN MIGUEL, SIMON BOLIVAR, URB. 7 DE ABRIL, AZAYA, COLINAS DE IBARRA, HUERTOS FAMILIARES, 16 DE FEBRERO, 28 DE ABRIL, COLINAS DE IBARRA, EL CAMAL, EN OBRER, GALO LARREA, STO. DOMINGO

100150006	AZAYA NORTE, AZAYA, MIRADOR DE AZAYA, COLINAS DE IBARRA, HUERTOS FAMILIARES, COLINAS DE IBARRA
100150007	AZAYA, MIRADOR DE AZAYA,, 16 DE FEBRERO, 28 DE ABRIL
100150008	LOS PINOS COOP. 15 DE DICIEMBRE, MIRADOR DE AZAYA,LA PRIMAVERA, LAS PALMAS, PANECILLO, ALPACHACA, 16 DE FEBRERO, 28 DE ABRIL
100150009	EL CHOFER, LAS PALMAS AUXILIOS MUTUOS, EL CHOFER PANECILLO, 16 DE FEBRERO, 28 DE ABRIL, AJAVI GRANDE, LA MERCED, URB. RIVADENEIRA, NUEVO HOGAR, URB. JARDIN, URB. JOSE M. LEORO, CIUDAD BLANCA
100150010	16 DE FEBRERO, 28 DE ABRIL, ALPACHACA, AZAYA, MIRADOR DE AZAYA, AJAVI CHICO, SAN MIGUEL, SIMON BOLIVAR, URB. 7 DE ABRIL, AUXILIOS MUTUOS, EL CHOFER
100150011	AJAVI CHICO, STO. DOMINGO, AJAVI GRANDE, LA MERCED, SAN MIGUEL, SIMON BOLIVAR, URB. 7 DE ABRIL
100150012	MOLINOS, SAN MARTIN, STO. DOMINGO, AJAVI GRANDE, LA MERCED
100150013	EL LIBERTADOR, EL OLIVO, MOLINOS, SAN MARTIN, SAN AGUSTIN, TAHUANDO, SAN JUAN CALLE, B.E.V. ,CATOLICA, SAN FRANCISCO, LA VICTORIA
100150014	SAN AGUSTIN, TAHUANDO, MOLINOS ,SAN MARTIN, EL CARMEN,EL LIBERTADOR,SAN FRANCISCO, EGAS GRIJALVA , AMAZONAS
100150015	CIUDAD BLANCA, AJAVI GRANDE, LA MERCED, MOLINOS, SAN MARTIN, CABEZAS BORJA,

	VELASCO, URB. JARDIN, URB. JOSE M. LEORO, CARMEN
100150016	URB. RIVADENEIRA, NUEVO HOGAR, URB. JARDIN, URB. JOSE M. LEORO, URB. PILANQUI, CIUDAD BLANCA, ALMEIDA GALARRAGA
100150017	NUEVOS HORIZONTES, ZOILA GALARRAGA, CANANVALLE, COOP. FLOTA IMBABURA, EL EMPREDRADO, LAS PALMAS, EL CHOFER, URB. RIVADENEIRA, NUEVO HOGAR, ALMEIDA GALARRAGA, URB. JARDIN, URB. JOSE M. LEORO, EL EJIDO
100150018	EL MILAGRO, URB. COLINAS DEL SUR, PUGACHO BAJO, CANANVALLE, NUEVOS HORIZONTES, ZOILA GALARRAGA, EL EJIDO, URB. RIVADENEIRA, NUEVO HOGAR, LA FLORIDA
100150019	LA FLORIDA, EL EJIDO, EJIDO DE CARANQUI, LA FLORESTA
100150020	CABEZAS BORJA, VELASCO, CIUDAD BLANCA, EL CARMEN, YACUCALLE, PILANQUI, EGAS GRIJALVA, AMAZONAS, SELVALEGRE, URB. PILANQUI, EJIDO DE CARANQUI, LA FLORESTA, BOLA AMARILLA, YUYUCOCHA, PULMON, AEREOPUERTO, URB. AJAVI, EL EJIDO, ALMEIDA GALARRAGA
100150021	EGAS GRIJALVA, AMAZONAS, BASILICA, YACUCALLE, DOMINGO ALBUJA , TEODORO GOMEZ, DON BOSCO, SAN JUAN CALLE
100150022	DON BOSCO, EGAS GRIJALVA, AMAZONAS, EL CARMEN, SAN JUAN CALLE, SAN FRANCISCO, B.E.V., CATOLICA, BASILICA, YACUCALLE
100150023	B.E.V., CATOLICA, LA VICTORIA, SAN JUAN CALLE, LA CAMPINA

100150024	LA CAMPINA, LOS CEIBOS, DOMINGO ALBUJA ,TEODORO GOMEZ, LA PRIMAVERA
100150025	DOMINGO ALBUJA, TEODORO GOMEZ, LOS CEIBOS, FAUSTO ENDARA
100150026	YACUCALLE, DOMINGO ALBUJA , TEODORO GOMEZ
100150027	BOLA AMARILLA , YUYUCOCHA, YACUCALLE, CUATRO ESQUINAS, DOMINGO ALBUJA, TEODORO GOMEZ, LA CANDELARIA, CARANQUI
100150028	FAUSTO ENDARA, DOMINGO ALBUJA, TEODORO GOMEZ, LOS CEIBOS, CUATRO ESQUINAS, EN RETORNO, LA CANDELARIA, CARANQUI
100150029	LOS CEIBOS, EN RETORNO, SAN LUIS, FAUSTO ENDARA, LA CANDELARIA, CARANQUI
100150030	BOLA AMARILLA , YUYUCOCHA, LA CANDELARIA, CARANQUI, GUAYAQUIL DE PIEDRAS
100150031	10 DE AGOSTO, BOLA AMARILLA, ,YUYUCOCHA, PULMON, EJIDO DE CARANQUI, LA FLORESTA, LA CANDELARIA, CARANQUI
100150032	GUAYAQUIL DE PIEDRAS, LA CANDELARIA , CARANQUI, EN RETORNO, BELLAVISTA, LA PRADERA

Fuente: INEC


4.1.2.1 Densidad Poblacional

Con los datos de suministros base para el análisis del estudio, procedemos a realizar los cálculos, primero transformamos los km² a hectáreas, procedemos a calcular la densidad poblacional hab/hectáreas y luego procedemos a calcular el promedio del consumo mensual, en la Ilustración 49 se puede observar el análisis a nivel total de todas las zonas de estudio.

AnalisisEmpaTesis

Página Inicial Catálogo Favoritos Paneles de Control Nuevo Abrir Conectado como weblogic

Perdida en Distribución Consumo Tarifas Consumo Suministros Analisis por Suministro Consumo Planes Consumo Sectores Consumo Zona Persona Consumo Día Densidad Poblacional



Página 1

Año 2014 2015 2016

Zona	Sector	Area Catastro(hectareas)	Habitantes 2010	% Habitantes Menores 5	Densidad 2010	Habitantes Año Estudio	Densidad Año Estudio	Consumo Promedio Mensual Año Estudio
100150001 - COLINAS DE IBARRA, HUERTOS FAMILIARES, AZUÑA NORTE	ZONA-50001 - SECTOR 001	11.47	454	11.23%	39.59	489	42.82	15.63
		11.47	454	11.23%	39.59	488	43.41	16.73
		11.47	454	11.23%	39.59	507	44.22	16.58
	ZONA-50001 - SECTOR 002	47.22	497	11.27%	10.52	535	11.33	17.89
		47.22	497	11.27%	10.52	545	11.54	18.56
		47.22	497	11.27%	10.52	555	11.76	19.75
	ZONA-50001 - SECTOR 003	45.07	463	9.94%	10.27	498	11.06	17.88
		45.07	463	9.94%	10.27	508	11.27	19.74
		45.07	463	9.94%	10.27	517	11.47	19.69

Ilustración 49: Datos Análisis Densidad Poblacional

Fuente: Elaboración Propia

En la Ilustración 50 se puede visualizar en análisis en un diagrama de dispersión de todas las zonas en estudio.

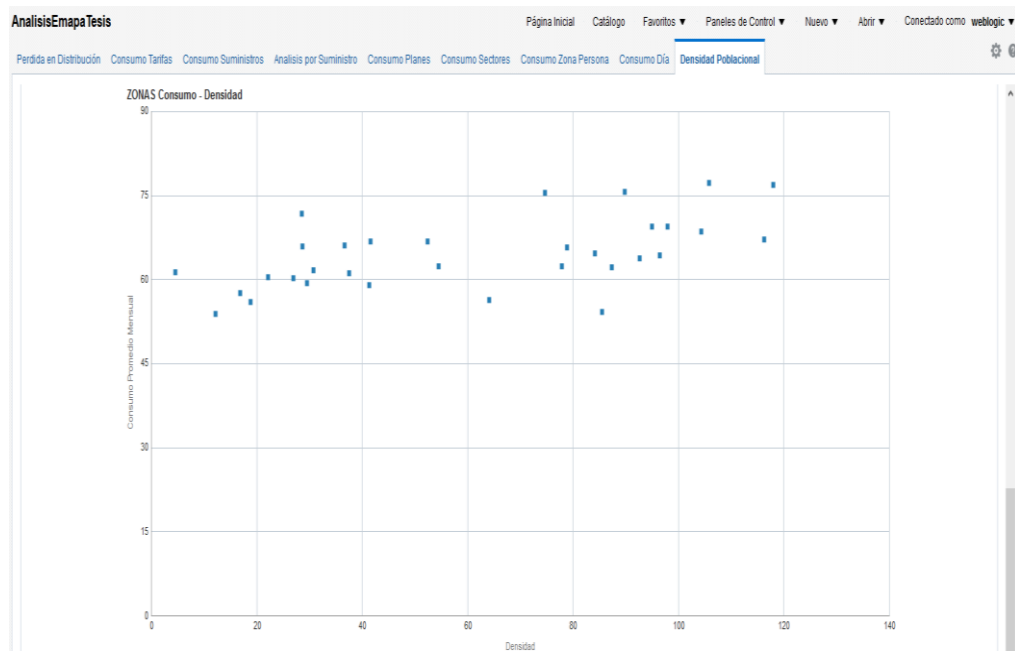


Ilustración 50: Gráfico Análisis Densidad Poblacional Zonas

Fuente: Elaboración Propia

En la Ilustración 51 se observa el análisis en un diagrama de dispersión de todos los sectores de una zona específica.

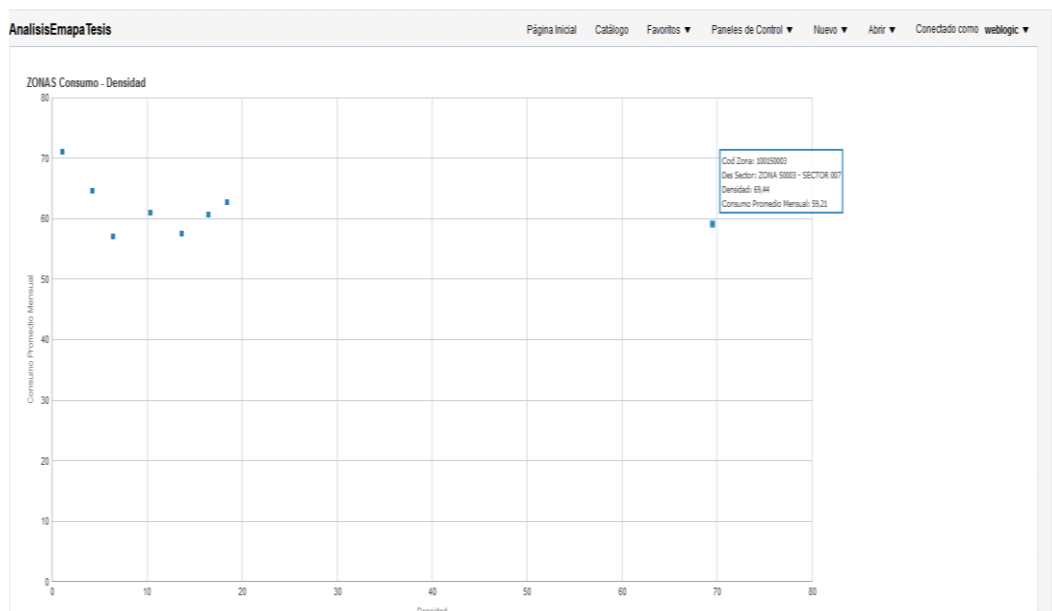


Ilustración 51: Análisis Densidad Poblacional Sector

Fuente: Elaboración Propia

Para la obtención de una mejor interpretación de la correlación de los datos, utilizamos los valores de las dos variables, Densidad Año Estudio y Consumo Promedio Mensual Año de Estudio, estas variables las analizamos con una herramienta especializada en estadística y proyecciones como es SPSS, la cual con el siguiente análisis, en la cual tenemos dos variables consideradas, la primera la densidad poblacional promedio de los tres años de estudio y la segunda el consumo promedio mensual de los tres años de estudio.

Densidad Poblacional habitantes por hectárea (hab/ha)

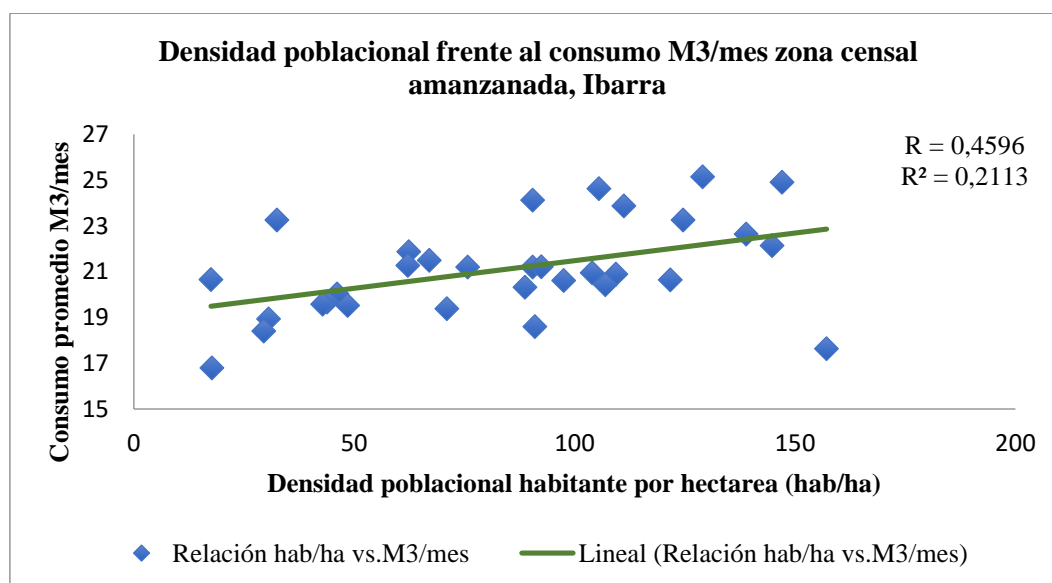


Ilustración 52: Densidad poblacional habitante por hectárea (hab/ha) con el consumo de agua residencial M3/mes por zona censal.

Fuente: Datos BI 2017, basado en información INEC 2010 y EMAPA-I periodo de consumo de agua residencial (periodo 2014 - 2017)

Tabla 25: Nivel de significación del coeficiente de correlación R relación hab/ha vs M3/mes .

Nº	VARIABLES	Valor de coeficiente de correlación R	Niveles de significación		Significado
			5%	1%	
			0,3442	0,4428	
1	Densidad Hab/ha	0,45963	*	**	Correlación positiva significativa al 95% y 99%

Fuente: Resultados de correlación lineal 2018

Nota: 1 Niveles de significación de 5% y 1% calculados en base a la tabla 16, tomando en cuenta que se tiene 32 datos.

Por el número de datos analizados existe una correlación positiva significativa entre la densidad poblacional y el consumo promedio m3/mes, es decir que a mayor número de habitantes por hectárea existe un incremento en el consumo del agua, existen datos que están muy alejados de la lineal de correlación, esto se debe a que en esta zona censal existe un bajo consumo de agua potable pero alta densidad poblacional, esto podría deberse a malas tomas de lecturas o analizarlos con otros factores, se recomienda realizar una verificación de cada uno de los suministros que se encuentran en este sector .

Para mejor visualización del análisis la Ilustración nos 53 muestra con colores específicos la densidad poblacional y el consumo de agua potable.

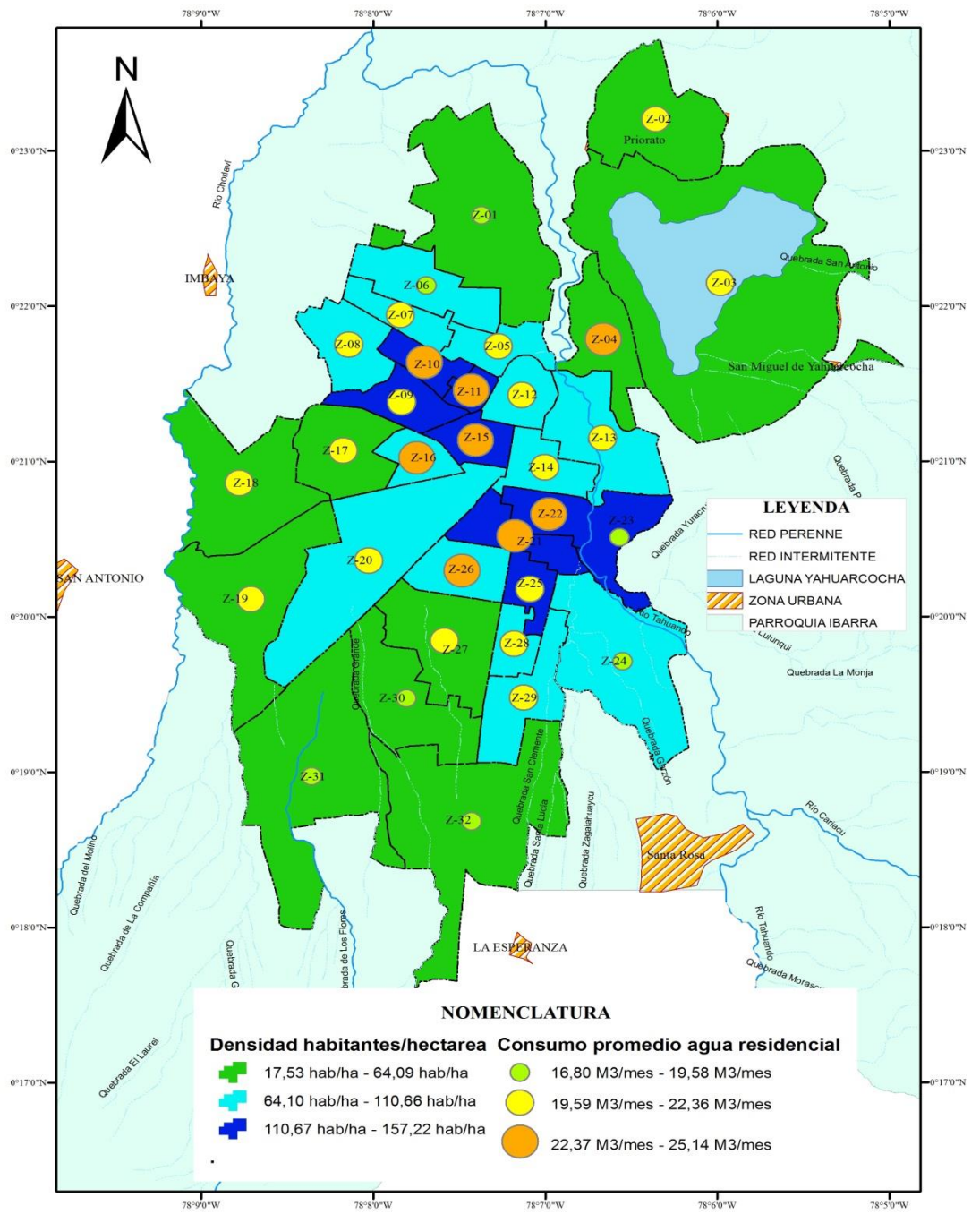


Ilustración 53: Zonificación de la Densidad Poblacional (hab/ha) con el consumo promedio de agua residencial m3/mes por zona censal amanzanada, parroquia urbana de la ciudad de Ibarra.

Fuente: Basado en información INEC 2010 Y EMAPA-I periodo de consumo de agua residencial (periodo 2014-2017).

4.1.2.1 Tamaño de vivienda

El tamaño de la vivienda está fuertemente asociado con el requerimiento del líquido vital en la Ilustración 54 se aprecia que las viviendas de mayor tamaño tienen a gastar más agua que aquellas viviendas con menor superficie. También se muestra el rango de viviendas más frecuente entre 100 y 250 m² que serían el tamaño de casas más representativas para la ciudad. Se puede considerar que las viviendas con mayor tamaño dispondrían de más superficies para su mantenimiento incrementando el gasto del agua.

Para Mendoza (2010), un cuarto adicional en la vivienda aumenta el gasto en consumo de agua en 3,72%, consecuentemente si se analiza el tamaño de casa mayor a 250 m² se tiene la tendencia a incrementar el consumo mensual. En conclusión existe una correlación positiva significativa del tamaño de la vivienda con los M³/mes.

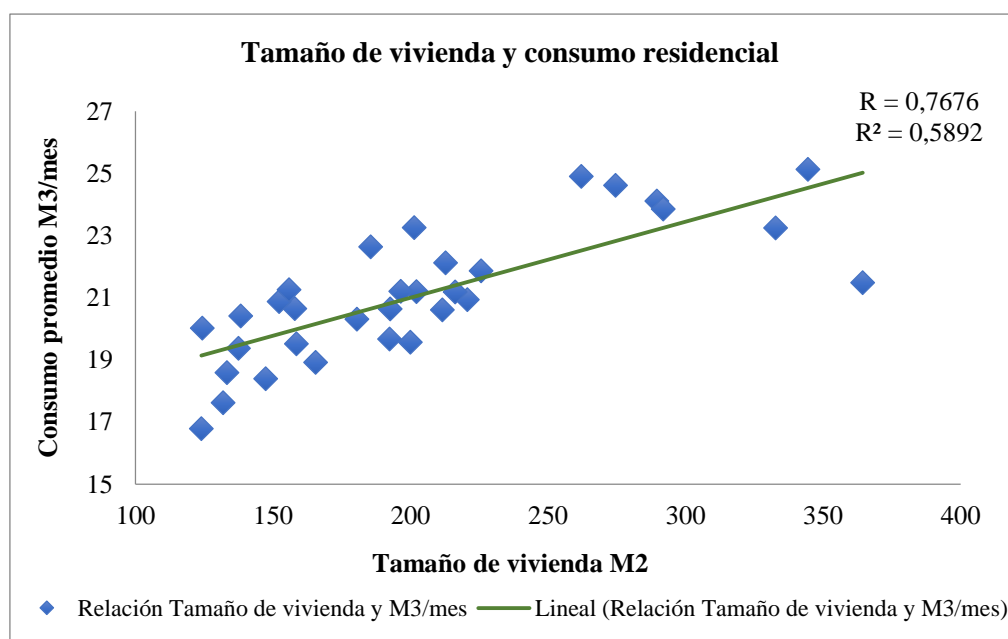


Ilustración 54: Tamaño de vivienda (M²) y el consumo residencial M³/mes

Fuente: Elaboración propia con base en Datos BI- Proyecto Ibarra Verde, 2017. EMAPA-I consumos fracturados (2014-2016) y adaptado GADI catastro 2016

En la Ilustración 55, se aprecia una distribución de las viviendas de mayor tamaño concentradas en el centro de la ciudad y a medida que se alejan del centro va disminuyendo la superficie de las edificaciones. Las zonas con presencia de casas más grandes se localizan en las zonas Z-14, Z-15, Z-16, Z-21, Z-22, que corresponde a las localidades de La Merced, San Agustín, Ciudad Blanca, urbanizaciones: Pílanqui, José Leoro, Jardín; Basílica; Amazonas y parte del barrio de Yacucalle.

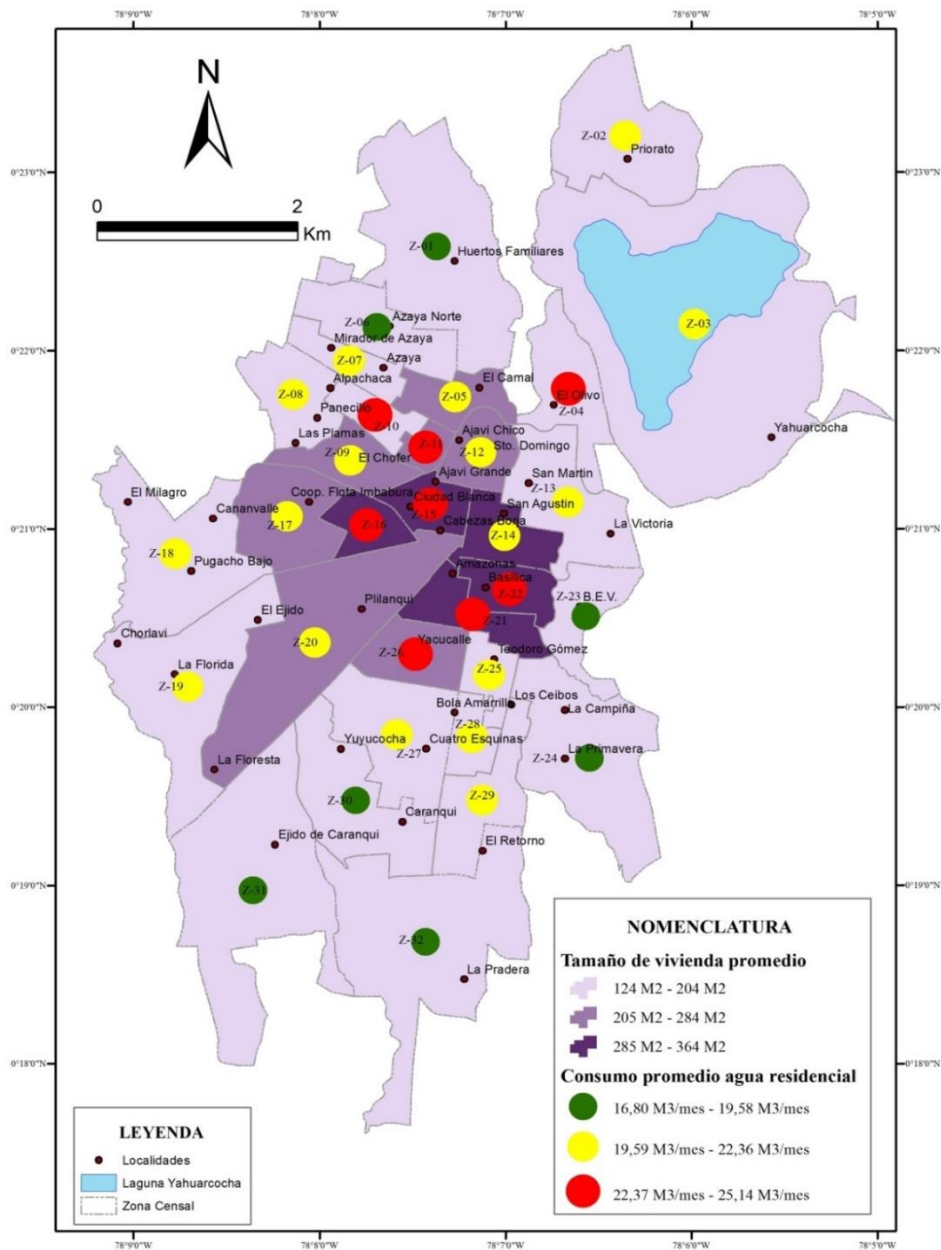


Ilustración 55: Distribución espacial tamaño de vivienda (M2) y consumo residencial M3/mes

Fuente: Elaboración propia con base en Datos BI- Proyecto Ibarra Verde, 2017. EMAPA-I consumos fracturados (2014-2016) y adaptado GADI catastro 2016.

A continuación en la Tabla 26 se detalla la tabla de significancia utilizada para el análisis de los datos.

Tabla 26: Significancia de los coeficientes de correlación del Libro Aplicación de estimadores estadísticos y diseño experimentales en investigaciones forestales.

Grados de libertad	5%	1%	Grados de libertad	5%	1%
1	.997	1.000	24	.388	.496
2	.950	.990	25	.381	.487
3	.878	.959	26	.374	.478
4	.811	.917	27	.367	.470
5	.754	.874	28	.361	.463
6	.707	.834	29	.355	.456
7	.666	.798	30	.349	.449
8	.632	.765	35	.325	.418
9	.602	.735	40	.304	.393
10	.576	.708	45	.288	.372
11	.553	.684	50	.273	.354
12	.532	.661	60	.250	.325
13	.514	.641	70	.232	.302
14	.497	.623	80	.217	.283
15	.482	.606	90	.205	.267
16	.468	.590	100	.195	.254
17	.456	.575	125	.174	.228
18	.444	.561	150	.159	.208
19	.433	.549	200	.138	.181
20	.423	.537	300	.113	.148
21	.413	.526	400	.098	.128
22	.404	.515	500	.088	.115
23	.396	.505	1000	.062	.081

Fuente: Aguirre y Vizcaino 2010

4.1.3 Resultados Generales de la Investigación

Al realizar análisis de los datos en cuanto al consumo de agua potable, se pudo identificar factores por los cuales el consumo no era el adecuado, se pudo evidenciar suministros que tenían tarifas con beneficios de uso residencial y que en realidad el consumo eran para otras actividades muy alejadas a la realidad. Al realizar los cruces con variables como número de población se identificó que los consumos tomando en cuenta el número de habitantes era demasiado alto al promedio sugerido por la OMS, que es de 100 litros diarios por persona.

Se pudo identificar que los índices de pérdida de agua potable superaban el 30% es decir que de cada 100 litros de agua potable distribuidos por cada planta solo 70 son facturados, los 30 restantes son pérdidas ya sean por fallas en el sistema de distribución o por utilización de agua de forma fraudulenta.

La plataforma BI está desarrollada en base a los requerimientos de cada una de las entidades requirentes, con lo cual los resultados han sido evaluados por parte del personal de las entidades involucradas, se ha enfocado el análisis de resultados a cada área de negocio, se han realizado análisis tanto de resultados obtenidos en la herramienta BI, así como también el análisis estadístico a partir de la información generada por el BI mediante la utilización de software especializado con es el PSP.

Para facilitar el análisis por parte de los usuarios finales se desarrolló un Dashboard el cual cuenta con informes interactivos de cada una de las necesidades de cada empresa, en el cual se puede acceder a información necesaria para la toma de decisiones.

Al generar conocimiento a partir de información de varias instituciones nos ayuda a comprender que el desarrollo de las zonas urbanas y por ende de la utilización de recursos, depende de varios factores socioeconómicos y sobre todo de la organización que se lleve al momento de planificar la expansión de las ciudades.

4.1.4 Desarrollo de Hipótesis

La realización de esta investigación contribuyo a identificar factores que afectan el abastecimiento de agua potable, entre los cuales están, en primer lugar

se identificó que el 30% de agua potable que es distribuida por las plantas de tratamiento no llegan a ser comercializada, es decir que este porcentaje se pierde ya sea por fallas en el sistema de distribución o por hurtos, esto está catalogado como agua no contabilizada.

En segundo lugar se pudo identificar que existían varios medidores que por su consumo no cumplían los requisitos para estar en ciertas tarifas, existían medidores que consumían más de 500 m³ con tarifa residencial y todos sus beneficios, al igual que medidores con tarifa residencial que pertenecían a entidades públicas, mercados y asociaciones.

Al identificar estos problemas se puede corregir de manera más clara, ya que con el análisis de información se tiene la información exacta donde están sucediendo este tipo de problemas, al corregir estos problemas el recurso agua potable podrá ser distribuido de mejor manera, ya que al tener mayor disponibilidad de agua potable se podrá tener mayor cobertura y sobre todo disponibilidad, o que ayudara a mejorar el estilo de vida de los habitantes.

4.1.4.1 ¿Cómo se realiza actualmente el análisis de consumo de agua potable?

Actualmente el consumo de agua potable se lo realiza en áreas por separado, es decir sin tener una integridad de datos para poder relacionarlos, se maneja en dos fuentes de datos separadas los consumos y la producción, sin tomar en cuenta que no se puede realizar análisis tomando en cuenta factores externos como crecimiento poblacional, aumento de infraestructura, etc.

4.1.4.2 ¿Qué tipo de análisis requiere en cuestión del consumo de agua potable?

Los análisis que se requiere con los de tendencias de consumos, tomando en cuenta los parámetros propios de la empresa así como los externos, se debe enfocar los análisis a reducir las pérdidas y mejorar la distribución del agua potable.

4.1.4.3 ¿Cuáles son los parámetros o estándares para tomar decisiones?

Los parámetros que se deben considerar para los análisis de agua potable son los propios de la empresa, como por ejemplo cual sería el máximo de consumo por tarifa para considerar un mal uso del agua, al igual que los externos que son los recomendados por la OMS en el cual especifica que el consumo ideal por persona de agua potable es de 100 m³ al día. Toda esta información debe ser directamente relacionada con la disponibilidad de agua potable existente en cada una de sus plantas de distribución, con el objetivo de mantener una disponibilidad de agua potable para la distribución en cada uno de los sectores.

4.2 Conclusiones

Se ha desarrollado la herramienta BI bajo la utilización de la metodología de Kimball, la cual nos ayudó a realizar una implementación sistematizada de la herramienta, con la cual se pudo aplicar para realizar análisis de consumos en diversos aspectos, al obtener resultados se pudo identificar varios factores por los cuales los consumos de agua potable son altos, entre los cuales están la utilización de agua en actividades que son diferentes a los de consumos residenciales así como el desperdicio ya sean estos por consumo no registrado hurto o fallas en los sistemas de distribución.

Se utilizó técnicas propias de la inteligencia de negocios como son OLAP y Data Mining, lo que permitió realizar análisis detallados y sobre todo de fácil comprensión por los usuarios finales, se realizó análisis basados en tendencias de consumo, máximos, mínimos, así como de correlaciones entre datos de diferentes instituciones.

Al realizar la integración de datos de EMAPA con las diferentes instituciones, se consolidó información valiosa para los análisis, en los cuales se obtuvo conocimiento sobre las tendencias de consumo de agua potable, así como los índices de pérdidas o desperdicio de agua potable que nos permiten tomar medidas para mejorar la distribución y disminuir las pérdidas.

La implementación de esta herramienta mejoró notablemente el tiempo en el cual se puede acceder a información importante para la toma de decisiones, aunque en la actualidad existe una resistencia en la implementación de herramientas gerenciales de toma de decisiones en instituciones públicas, esto generado por la falta de visión y sobre todo por los costos de inversión que representa, tanto para la adquisición como tal y falta de personal capacitado.

La aplicación de controles de mandos permitió a los usuarios finales acceder a la información mediante informes interactivos fácilmente con el objetivo de obtener resultados de forma rápida y de fácil comprensión.

En la actualidad las empresas privadas son las que están a la vanguardia de la implementación de herramientas BI, sobre todo las empresas que están el mercado de la banca y venta de productos, en la parte pública las empresas pioneras que tiene implementado esta infraestructura son las de generación y comercialización de energía eléctrica.

4.3 Recomendaciones

Para mantener actualizada la información pertinente a las diferentes instituciones, se propone elaborar un proceso en el cual consten todas las actividades realizadas por cada responsable, con el objetivo de mantener los Procesos ETL activos para tener un Almacén de Datos actualizado con la información de los diferentes almacenes de datos.

Los investigadores deben conocer los procesos de cada una de las instituciones, para poder obtener un análisis de datos claros y de fácil comprensión por terceros.

Para el levantamiento de información y análisis de requisitos se debe identificar a los stakeholder, con el fin de obtener resultados acordes a las necesidades de cada institución interesada.

Se debe tener conocimiento de técnicas de inteligencia de negocios para el análisis de información ya que esto ayuda a realizar análisis de la información de manera rápida y enfocada en nuestras necesidades.

En cuanto al desarrollo de proyectos relacionados con BI, se recomienda la utilización de la metodología propuesta por Ralph Kimball, metodología con tendencia de desarrollo ágil y con facilidades de obtener resultados rápidos gracias a su modelo multidimensional.

Para obtener mejores resultados en la distribución de agua potable lo óptimo sería la aplicación de redes de agua inteligente, lo cual nos permite detectar en tiempo real averías, atascos o fuga de agua potable. Esta tecnología se está aplicando en los proyectos pilotos de ciudades inteligentes en países como España

y Alemania, el análisis del costo beneficio es significativo ya que precautelar el líquido vital óptimo para el consumo humano es indispensable.

5 Referencias

- Abreu, J. L. (2014). *El Método de la Investigación*.
- Abril Frade, D. O., & Pérez Castillo, J. N. (2017). *Estado actual de las tecnologías de bodega de datos y OLAP aplicadas a bases de datos espaciales*. REVISTA INGENIERÍA E INVESTIGACIÓN.
- Azuaje, A. (2014). *Metodología de Kimball*. CARACAS.
- Brito, D. (26 de 01 de 2014). *Inteligencia de Negocios*. Obtenido de <http://inteligenciadenegociosdiegobrito.blogspot.com/2014/01/la-metodologia-de-kimball.html>
- Castro Rozo, F. E. (2013). *Indicadores de gestión para la toma de decisiones basada en Inteligencia de Negocios*. Caldas: TIA.
- Cota , M. (12 de Julio de 2011). *BI-SPAIN*. Obtenido de BI-LATINO: <http://www.bi-spain.com/articulo/71397/oracle/energia-utilities/mexico-df-ahorra-11-millones-de-litros-de-agua-por-hora-gracias-al-business-intelligence-de-oracle-video-entrevista-exclusiva>
- Cueva Andrade, C., Jerez Cevallos, S., Díaz Zuñiga, P., & Ron Egas, M. (2014). *DETERMINACIÓN DE LA LÍNEA BASE PARA LA APLICACIÓN DE SISTEMAS DE BI (BUSINESS INTELLIGENCE) EN EL ECUADOR, DENTRO DE UN PROGRAMA PARA EL USO DE TECNOLOGÍA DE AVANZADA EN LA ADMINISTRACIÓN PÚBLICA Y PRIVADA*. Quito: ESPE.
- DATAPRIX. (s.f.). *Dataprix, el portal profesional sobre Software y Servicios IT para la empresa*. Obtenido de <http://www.dataprix.com/24-or-genes-antecedentes-las-bases-datos>
- emapa. (3 de 10 de 2017). *Emapa Ibarra - Filosofía*. Obtenido de Emapa Ibarra: <http://www.emapaibarra.gob.ec/index.php/filosofia>
- Esri. (s.f.). *ArcGis Resources*. Obtenido de <http://resources.arcgis.com/es/help/getting-started/articles/026n00000014000000.htm>

- Frolich, L. M., Caraguay Procel, J. A., Rosales, O., Cárdenas Pepinós, A. F., Castro, P., Llanga Ochoa, P. I., . . . Sarmiento, F. (2017). Ecología humana y la urbe inteligente: Utilizando mapeo interactivo para el análisis socio uso de agua y de energía eléctrica en Ibarra. *CIENCIA*, 227-240.
- GAMBOA CRUZADO, J. (2015). *DESARROLLO DE BUSINESS INTELLIGENCE, BASADO EN LA METODOLOGÍA DE RALPH KIMBALL, PARA MEJORAR EL PROCESO DE TOMA DE DECISIONES EN EL ÁREA DE ADMISIÓN DE LA UNIVERSIDAD AUTÓNOMA DEL PERÚ*. Lima.
- Gómez Morales, A. J. (2013). *Inteligencia de negocios, una ventaja competitiva para las organizaciones*. Trujillo.: Revista CIENCIA Y TECNOLOGÍA.
- GUEVARA VEGA, C. P. (2015). *DESARROLLO DE UNA PLATAFORMA DE BUSINESS INTELLIGENCE PARA FACILITAR EL ANÁLISIS DE DATOS DE LAS COMPETENCIAS GENERALES DE FORMACIÓN APLICADAS EN EL DESEMPEÑO LABORAL DE LOS EGRESADOS DE LA UNIVERSIDAD TÉCNICA DEL NORTE*. LATACUNGA.
- Ibarra, G. A. (12 de 11 de 2017). *Gobierno Autónomo Descentralizado de San Miguel de Ibarra*. Obtenido de <https://www.ibarra.gob.ec/web/>
- Instituto Nacional de Estadística y Censos. (15 de 11 de 2017). *Instituto Nacional de Estadística y Censos*. Obtenido de Instituto Nacional de Estadística y Censos: <http://www.ecuadorencifras.gob.ec/institucional/home/>
- Lascano Sumbana, K. V. (2016). *DISEÑO DE UN MODELO PARA DETERMINAR LA DEMANDA DE AGUA POTABLE PARA JUNTAS DE AGUA POTABLE DE LA PROVINCIA DE TUNGURAHUA: CASO DE APLICACIÓN PARROQUIA HUAMBALO*. Ambato.
- Oracle. (1 de 04 de 2014). *Technetwork*. Obtenido de <http://www.oracle.com/technetwork/es/articles/database-performance/pdbs-backup-recovery-dbpit-2192211-esa.html>
- Oracle. (2016). *Guía del Usuario de Oracle Business Intelligence Enterprise Edition*. Oracle.

- PowerData. (19 de 09 de 2015). *PowerData Solutions S.L.* Obtenido de <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/406542/qu-son-las-bases-de-datos-multidimensionales>
- Quest. (2014). *Quest*. Obtenido de <https://www.quest.com/toad/>
- Rodriguez Almeida, A. M., & Da Silva Camargo, S. (2015). *ACADEMIC ANALYTICS: APLICANDO TÉCNICAS DE BUSINESS INTELLIGENCE SOBRE DATOS DE PERFORMANCE ACADÉMICA EN ENSEÑANZA SUPERIOR*. Aracaju: EXATAS E TECNOLÓGICAS.
- Sinnexus. (2016). *Sinnexus - Business Intelligence Informática estratégica*. Obtenido de http://www.sinnexus.com/business_intelligence/arquitectura.aspx
- Universidad Poliecnica de Valencia. (s.f.). *Museo Informatica*. Obtenido de <http://histinf.blogs.upv.es/author/>
- WEB 2.0 - MediaWiki. (30 de 01 de 2014). *Inteligencia de Negocios*. Obtenido de <http://inteligenciadenegociosval.blogspot.com/2014/01/metodologia-de-kimball.html>