



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE CIENCIAS
ESCUELA DE COMPUTACIÓN
CENTRO DE INVESTIGACIÓN EN SISTEMAS DE INFORMACIÓN

DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA OBTENCIÓN DE INDICADORES RELACIONADOS CON LOS OBJETIVOS DE DESARROLLO SOSTENIBLE PARA VENEZUELA

**Trabajo Especial de Grado presentado ante la ilustre
Universidad Central de Venezuela por el
Br. Abelardo José Moreno Carballeda
Para optar al título de Licenciado en Computación**

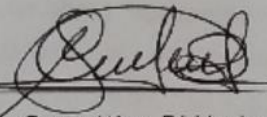
**Tutora: Profa. Concettina Di Vasta
Caracas, septiembre 2019**

ACTA

Quienes suscriben, miembros del Jurado designado por el Consejo de Escuela de Computación, para examinar el Trabajo Especial de Grado presentado por el Br. Abelardo José Moreno Carballada C.I. 18.002.106, con el título "DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA OBTENCIÓN DE INDICADORES RELACIONADOS CON LOS OBJETIVOS DE DESARROLLO SOSTENIBLE PARA VENEZUELA" a los fines de optar al título de **Licenciado en Computación**, dejan constancia de lo siguiente:

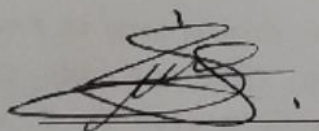
Dicho trabajo, leído por cada uno de los miembros del jurado, se fijó el día 10 de Octubre de 2019, a las 8 a.m., para que su autor lo defendiera en forma pública en el Laboratorio Groupware de la Escuela de Computación, mediante una presentación oral de su contenido, luego de lo cual se respondieron las preguntas formuladas. Finalizada la defensa pública del Trabajo Especial de Grado, el jurado decidió aprobarlo con la nota de 20 puntos.

En fe de lo cual se levanta la presente Acta, en Caracas a los 10 días del mes de Octubre del año 2019.



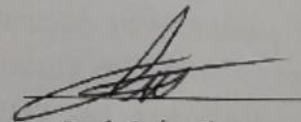
Prof. Concettina Di Vasta

Tutora



Prof. Jaime Parada

Jurado Principal



Prof. Iván Flores

Jurado Principal

AGRADECIMIENTOS

Agradezco a mi tutora Tina por su guía y ayuda durante el desarrollo de este TEG, a Estefanía, Jorge, Iván y Manuel por su invaluable colaboración durante el desarrollo del proyecto, a mi amigo German por apoyarme y aconsejarme a lo largo de mi carrera universitaria y a mi esposa, Daniela, por su amor, paciencia y compañía.

“Everything not saved will be lost.”

- ***Nintendo “Quit Screen” message***

UNIVERSIDAD CENTRAL DE VENEZUELA

FACULTAD DE CIENCIAS

ESCUELA DE COMPUTACIÓN

CENTRO DE INVESTIGACIÓN EN SISTEMAS DE INFORMACIÓN

**DESARROLLO DE UNA SOLUCIÓN DE INTELIGENCIA DE NEGOCIO PARA LA
OBTENCIÓN DE INDICADORES RELACIONADOS CON LOS OBJETIVOS DE
DESARROLLO SOSTENIBLE PARA VENEZUELA**

Autor: Br. Abelardo José Moreno Carballada

Tutora: Profa. Concettina Di Vasta

Fecha: Caracas, septiembre 2019

RESUMEN

En el año 2015 las Naciones Unidas proponen una serie de indicadores que deberán ser cumplidos por todos los países del mundo llegados el 2030 para lograr que el mundo sea un lugar mejor, aumentando la calidad de vida de billones de personas y sin dejar a nadie atrás. Se crearon un total de 231 indicadores que abarcan diferentes ámbitos tanto sociales, económicos y ambientales.

Al contar con esta gran cantidad de indicadores que deben ser medidos, analizados y publicados por cada país, se vuelve un gran reto realizar el tratamiento, manipulación y administración de los recursos necesarios para lograr esta tarea. Es por esto que surge la necesidad de desarrollar una solución de Inteligencia de Negocio que proporcione las herramientas necesarias para extraer, gestionar y visualizar los datos que permiten dar respuesta a estos indicadores. En la actualidad existen algunos países, como México y Perú, que ya cuentan con soluciones de Inteligencia de Negocio que dan respuesta a esta problemática, pero muchos otros aun no, entre ellos Venezuela.

Es por esto que, en el presente Trabajo Especial de Grado, se desarrolla una solución de Inteligencia de Negocio, adaptada para Venezuela, que permite la medición y despliegue de un subconjunto de los indicadores propuestos por las Naciones Unidas. Las tecnologías utilizadas para el desarrollo de la solución de Inteligencia de Negocio fueron las siguientes: Pentaho Data Integration, herramienta que permite extraer y transformar diversas fuentes de datos para luego consolidarlas y cargarlas en un único repositorio de datos; como Sistema Manejador de Base de Datos se utiliza PostgreSQL, para la creación de los almacenes de datos y área de preparación de datos, y se emplea Tableau como herramienta de visualización de los datos que se encuentran en los almacenes de datos creados. Como método de desarrollo de una solución de Inteligencia de Negocio se utiliza el Ciclo de Vida Dimensional de Ralph Kimball.

Palabras Claves: Objetivos de Desarrollo Sostenible, Agenda 2030, Solución de Inteligencia de Negocio, Indicadores, Venezuela, Pentaho Data Integration, PostgreSQL, Tableau.

ÍNDICE GENERAL

AGRADECIMIENTOS	iii
RESUMEN	iv
ÍNDICE GENERAL	v
ÍNDICE DE TABLAS	viii
ÍNDICE DE FIGURAS	ix
INTRODUCCIÓN	1
CAPÍTULO 1: PROBLEMA DE INVESTIGACIÓN	4
1.1 Planteamiento del problema	4
1.2 Objetivos	6
1.2.1 General.....	6
1.2.2 Específicos.....	6
1.3 Solución propuesta	7
1.4 Justificación de la solución	8
1.5 Alcance de la solución	9
CAPÍTULO 2: MARCO CONCEPTUAL	10
2.1 Desarrollo sostenible	10
2.1.1 Definición	10
2.1.2 Objetivos de Desarrollo Sostenible	11
2.1.3 Antecedentes	13
2.1.3.1 Perú: Sistema de Monitoreo y Seguimiento de los Indicadores de los Objetivos de Desarrollo Sostenible (ODS)	14
2.1.3.2 México: Sistema de Información de los Objetivos de Desarrollo Sostenible	19
2.2 Sistema de información	22
2.2.1 Definición	22
2.2.2 Tipos de sistemas de información.....	23
2.2.2.1 Basados en los niveles organizacionales	23
2.2.2.4 Basados en el procesamiento de datos	24
2.2.3 Beneficios de los sistemas de información	26
2.3 Inteligencia de Negocio	27
2.3.1 Definición	27
2.3.2 Características.....	28
2.3.3 Arquitectura de una solución de Inteligencia de Negocio	28
2.3.3.1 Fuentes de datos	29
2.3.3.2 Procesos de extracción, transformación y carga	29
2.3.3.3 Área intermedia.....	30
2.3.3.4 Almacén de datos	30

2.3.3.5 Herramientas de análisis	37
2.3.4 Beneficios de la Inteligencia de Negocio.....	37
2.4 Indicadores.....	38
2.4.1 Definición	38
2.4.2 Características.....	38
2.4.3 Tipos.....	39
2.4.4 Beneficios de los indicadores	39
2.5 Herramientas tecnológicas.....	40
2.5.1 Herramientas para el almacenamiento de datos.....	40
2.5.1.1 PostgreSQL versión 9.4.....	40
2.5.2 Herramientas para los procesos ETL.....	41
2.5.2.1 Pentaho Data Integration (Kettle) versión 8.2.....	41
2.5.3 Herramientas de visualización	42
2.5.3.1 Tableau versión 2019.2	42
CAPÍTULO 3: MARCO METODOLÓGICO.....	44
3.1 Ciclo de vida dimensional del negocio	44
3.1.1 Planificación del proyecto.....	45
3.1.2 Definición de requerimientos del negocio.....	46
3.1.3 Flujo de tecnología.....	46
3.1.4 Flujo de datos.....	47
3.1.5 Flujo de aplicaciones de Inteligencia de Negocio	47
3.1.6 Implementación	48
3.1.7 Mantenimiento y crecimiento	48
3.1.8 Administración del proyecto.....	48
CAPÍTULO 4: MARCO APLICATIVO	49
4.1 Planificación del proyecto	49
4.2 Definición de requerimientos del negocio.....	49
4.3 Flujo de tecnología	51
4.3.1 Diseño de la arquitectura técnica	51
4.3.2 Selección de productos e implementación	51
4.4 Flujo de datos	51
4.4.1 Modelado de base de datos.....	51
4.4.1.1 Área de preparación	51
4.4.1.2 Modelado dimensional.....	55
4.4.2 Diseño físico	57
4.4.2.1 Diseño físico del área de preparación	57
4.4.2.2 Diseño físico del modelo dimensional	60
4.4.3 Diseño e implementación del subsistema de ETL.....	67
4.4.3.1 Sistema ETL del área de preparación.....	68
4.4.3.2 Sistema ETL del almacén de datos.....	72
4.5 Flujo de aplicaciones de Inteligencia de Negocio.....	78
4.5.1 Especificación de aplicaciones BI	79
4.5.2 Desarrollo de aplicaciones BI	81

4.6 Implementación	94
4.7 Mantenimiento y crecimiento	94
4.8 Administración del proyecto	95
CONCLUSIONES	101
TRABAJOS FUTUROS.....	102
REFERENCIAS BIBLIOGRÁFICAS	103
ANEXOS.....	106
A. Contabilización de Indicadores	106
B. Correspondencia entre archivos fuentes y área de preparación.....	111
C. Estructura física de las tablas pertenecientes al área de preparación	114
D. Estructura física de las tablas pertenecientes al almacén de datos.....	127

ÍNDICE DE TABLAS

Tabla 1. Comparación entre sistemas OLTP y OLAP.....	25
Tabla 2. Comparación entre almacén de datos y bodega de datos.....	31
Tabla 3. Propiedades S.M.A.R.T. de indicadores.....	39
Tabla 4. Planificación del proyecto.....	49
Tabla 5. Resumen de contabilización de Indicadores.....	50
Tabla 6. Resumen de productos seleccionados.....	51
Tabla 7. Correspondencia de algunos archivos fuente con área de preparación.....	55
Tabla 8. Estructura de la tabla SA_DASH_INDICADOR_332.....	58
Tabla 9. Estructura de la tabla SA_DASH_INDICADOR_461.....	58
Tabla 10. Estructura de la dimensión fuente.....	60
Tabla 11. Estructura de la dimensión tiempo semestral.....	61
Tabla 12. Estructura física de la tabla de hechos FACT_DASH_IND121.....	63
Tabla 13. Estructura física de la tabla de hechos FACT_DASH_IND333.....	64
Tabla 14. Correspondencia de áreas temáticas y bodega de datos creadas.....	67
Tabla 15. Prueba de Aceptación.....	96

ÍNDICE DE FIGURAS

Figura 1. Arquitectura propuesta de la solución de Inteligencia de Negocio.....	7
Figura 2. Ciclo de vida dimensional del negocio.....	8
Figura 3. Esquema de los tres pilares del desarrollo sostenible.	10
Figura 4. Objetivos de Desarrollo Sostenible.	12
Figura 5. Ventana principal Sistema ODS Perú.....	14
Figura 6. Menú de selección de ODS.....	14
Figura 7. Menú de indicadores ODS 1 - Fin de la Pobreza.....	15
Figura 8. Cuadro estadístico de indicadores.	15
Figura 9. Visualizaciones de indicadores.	16
Figura 10. Mapas de indicadores.	16
Figura 11. Información acerca de metodología de cálculo de indicadores.....	17
Figura 12. Errores muestrales de indicadores.....	17
Figura 13. Selección de sub-indicadores.....	18
Figura 14. Selección de programas presupuestales.....	18
Figura 15. Ventana principal del Sistema ODS México.	19
Figura 16. Ventana emergente de selección de indicadores.	19
Figura 17. Visualización asociada a un indicador.....	20
Figura 18. Cuadro estadístico asociado a una visualización.	20
Figura 19. Información sobre el cálculo y obtención de un indicador.....	21
Figura 20. Datos utilizados para el cálculo de un indicador.....	21
Figura 21. Triángulo de sistemas de información.	22
Figura 22. Pirámide de los niveles organizacionales.....	23
Figura 23. OLTP y OLAP trabajando juntos.	26
Figura 24. Relación entre datos, información y conocimiento.....	27
Figura 25. Arquitectura de una solución de Inteligencia de Negocio.	29
Figura 26. Implementación de bodegas de datos.....	31
Figura 27. Características de los almacenes de datos.	33
Figura 28. Tabla de Hechos.	35
Figura 29. Tabla Dimensional.	36
Figura 30. Nivel Jerárquico	36
Figura 31. Modelo Dimensional.....	37
Figura 32. Interfaz gráfica de PostgreSQL.	41
Figura 33. Interfaz de Pentaho Data Integration.....	42
Figura 34. Interfaz de Tableau.	43
Figura 35. Ciclo de vida dimensional.....	45
Figura 36. Formato de archivo fuente.	50
Figura 37. Archivo Excel sin tratamiento.	52
Figura 38. Archivo Excel transpuesto.....	53
Figura 39. Archivo Excel transpuesto correspondiente al indicador 3.1.1 con fuente INN. ..	53
Figura 40. Archivo Excel transpuesto correspondiente al indicador 3.1.1 con fuente MPPPS.	54
Figura 41. Modelo Dimensional Conceptual.....	57
Figura 42. Datos de la tabla SA_DASH_INDICADOR_332.....	59
Figura 43. Datos de la tabla SA_DASH_INDICADOR_461.....	59
Figura 44. Datos de la dimensión DIM_UNIDAD_MEDIDA.....	61

Figura 45. Datos de la dimensión DIM_TIEMPO_SEMESTRAL	62
Figura 46. Datos de la tabla de hechos FACT_DASH_IND_121	65
Figura 47. Datos de la tabla de hechos FACT_DASH_IND_333	65
Figura 48. Modelo Dimensional correspondiente al indicador 1.2.1.....	66
Figura 49. Modelo Dimensional correspondiente al indicador 3.3.3.....	66
Figura 50. Estructura del repositorio “ODS Intermedia” para las transformaciones	68
Figura 51. Estructura del repositorio “ODS Intermedia” para los Jobs	68
Figura 52. ETL para la carga de la SA del indicador 1.2.1	69
Figura 53. Campos correspondientes al Excel del indicador 1.2.1	69
Figura 54. Paso de carga hacia la tabla SA_DASH_INDICADOR_121	70
Figura 55. Job para la carga de las SA del ODS 1	71
Figura 56. Job final que agrupa el resto de los Jobs y control de errores	71
Figura 57. Transformaciones para las dimensiones del almacén de datos.....	72
Figura 58. Transformaciones para las tablas de hechos del almacén de datos.....	72
Figura 59. Jobs que llenan el almacén de datos	73
Figura 60. ETL para la carga de la dimensión DIM_UBICACION_GEO	73
Figura 61. Paso de inserción de SK inicial para la dimensión	74
Figura 62. Consulta que obtiene los datos a ser cargados de la SA hacia la dimensión	74
Figura 63. Paso de carga hacia la tabla DIM_UBICACION_GEO	75
Figura 64. ETL para la carga de la tabla de hechos FACT_DASH_IND_121	75
Figura 65. Paso de inserción de SK inicial para la tabla de hechos.....	76
Figura 66. Consulta que obtiene los datos a ser cargados en la tabla de hechos	76
Figura 67. Paso de carga hacia la tabla FACT_DASH_IND_121	77
Figura 68. Job para la carga de las tablas de hechos del ODS 1	78
Figura 69. Job final que agrupa el resto de los Jobs y control de errores	78
Figura 70. Estándar de visualizaciones para cuadros estadísticos.....	80
Figura 71. Estándar de visualizaciones para gráficos.....	80
Figura 72. Estándar de visualizaciones para mapas	81
Figura 73. Capa lógica de Tableau	81
Figura 74. Representación de los campos crudos del almacén de datos en Tableau.....	82
Figura 75. Representación de los campos limpios del almacén de datos en Tableau	82
Figura 76. Cuadro estadístico del porcentaje de personas en hogares con pobreza extrema	83
Figura 77. Cuadro estadístico del porcentaje de hogares pobres.....	83
Figura 78. Gráfico del indicador complementario Tasa Neta de Escolaridad Inicial.....	84
Figura 79. Gráfico de la esperanza de vida nacional.....	85
Figura 80. Gráfico comparativo entre la esperanza de vida de hombres y mujeres.....	86
Figura 81. Gráfico comparativo entre la esperanza de vida nacional, de hombres y de mujeres	87
Figura 82. Gráfico de la población desocupada	88
Figura 83. Mapa perteneciente al indicador 8.10.1	88
Figura 84. Inicio del cuadro de mando	89
Figura 85. Listado de los 17 ODS	90
Figura 86. Listado de metas del ODS 1	91
Figura 87. Icono en gris indicando metas que no poseen datos	91
Figura 88. Listado de indicadores de la meta 1.2	92
Figura 89. Indicador principal asociados al indicador 1.2.1	92
Figura 90. Navegación entre los indicadores	93

Figura 91. Menú de Navegación de Indicadores	93
Figura 92. Menú de Miga de Pan	94
Figura 93. Prueba de aceptación – Resultados Pregunta 1	96
Figura 94. Prueba de aceptación – Resultados Pregunta 2	97
Figura 95. Prueba de aceptación – Resultados Pregunta 3	97
Figura 96. Prueba de aceptación – Resultados Pregunta 4	98
Figura 97. Prueba de aceptación – Resultados Pregunta 5	98
Figura 98. Prueba de aceptación – Resultados Pregunta 6	99
Figura 99. Prueba de aceptación – Resultados Pregunta 7	99
Figura 100. Prueba de aceptación – Resultados Pregunta 8.....	100

INTRODUCCIÓN

Existen grandes problemas que afectan a todas las personas a nivel mundial, tales como la gran desigualdad social que se encuentra latente en cada uno de nuestros países y el incremento de la degradación ambiental que atenta con el futuro del planeta. Estos problemas son característicos de nuestra realidad actual y presentan un gran desafío para la comunidad internacional. La opción de continuar de la misma manera que lo hemos hecho en las últimas décadas ya no es viable, lo que hace necesario transformar el paradigma de desarrollo actual en uno que nos lleve por la vía del desarrollo sostenible, inclusivo y con visión de largo plazo.

Es por esto que, frente a estos desafíos, los 193 Estados Miembros de las Naciones Unidas, junto con un gran número de actores de la sociedad civil, el mundo académico y el sector privado, entablaron un proceso de negociación abierto, democrático y participativo, que resultó en la proclamación de la Agenda 2030 para el Desarrollo Sostenible, con sus 17 Objetivos de Desarrollo Sostenible (ODS), en septiembre de 2015.

Esta nueva Agenda es la expresión de los deseos, aspiraciones y prioridades de la comunidad internacional para los próximos 15 años. La Agenda 2030 es una agenda transformadora, que pone la igualdad y dignidad de las personas en el centro y llama a cambiar nuestro estilo de desarrollo, respetando el medio ambiente. Es un compromiso universal adquirido tanto por países desarrollados como en desarrollo, en el marco de una alianza mundial reforzada, que toma en cuenta los medios de implementación para realizar el cambio y la prevención de desastres por eventos naturales extremos, así como la mitigación y adaptación al cambio climático.

La Agenda 2030 para el Desarrollo Sostenible cuenta con 17 ODS, cada uno de estos ODS posee una cantidad de metas, totalizando un número de 169 metas distribuidas entre todos los ODS. De la misma manera cada una de las metas posee una cantidad de indicadores, llegando a un número de 231 indicadores distribuidos entre todas las 169 metas.

Al contar con una gran cantidad de indicadores a ser medidos, desplegados y publicados por cada país, no es tarea fácil el tratamiento, manipulación y administración de los recursos necesarios para lograr esta tarea. Es por esto que resulta necesario contar con una Solución de Inteligencia de Negocio que proporcione las herramientas necesarias para la extracción, manejo y visualización de los datos que respondan a todos y cada uno de los 231 indicadores de la Agenda 2030. En la actualidad existen algunos países, como México y Perú, que ya cuentan con Soluciones de Inteligencia de Negocio que permiten dar respuesta a los indicadores ODS, pero muchos otros, entre ellos Venezuela aun no cuentan con una.

Es por esto que este trabajo especial de grado se enfoca en el desarrollo de una solución de Inteligencia de Negocio, adaptada para Venezuela, que implemente una solución automatizada, para la obtención de los indicadores especificados en la Agenda 2030 que permitan la medición y despliegue de los ODS. Dicha solución será utilizada por una institución pública del país, el Instituto Nacional de Estadística (INE) para poder hacer un seguimiento y control de la Agenda 2030.

El INE se encarga de proveer los datos necesarios, tanto aquellos generados por ellos, como los que son generados por otras instituciones públicas del país. Muchos de estos datos no se encuentran disponibles para dar respuesta a los 17 ODS, por lo que se enfoca la solución de Inteligencia de Negocio hacia aquellos ODS de los cuales el INE posee mayor información, los cuales son:

ODS 1.- Fin de la Pobreza

ODS 3.- Salud y Bienestar

ODS 4.- Educación de Calidad

ODS 8.- Trabajo Decente y Crecimiento Económico

Por otro lado, aun dentro de esos cuatro ODS seleccionados, muchas veces es difícil obtener respuesta para un indicador directo relacionado a un ODS. En estos casos se utiliza un indicador complementario para medir el objetivo. Un indicador complementario es aquel que es capaz de dar respuesta de forma indirecta a los indicadores ODS, ya sea respondiéndolos de forma parcial, o respondiendo aspectos relacionados a lo que se busca obtener. En total entonces, se manejan 123 indicadores de eficacia pertenecientes a 4 ODS, y esos 123 indicadores se dividen a su vez en 91 complementarios y 32 principales.

Este trabajo consta de cuatro capítulos. A continuación, se hace un resumen de cada uno de ellos:

Capítulo 1. Problema de Investigación: en este capítulo, se presenta la situación actual acerca de los requerimientos que posee la organización para dar respuesta a los ODS; además, se establece el objetivo general y cada uno de los objetivos específicos del presente Trabajo Especial de Grado (TEG), así como la solución propuesta, la justificación de la solución y el alcance de la solución.

Capítulo 2. Marco Conceptual: en este capítulo, se detallan conceptos asociados al origen, motivación, enfoque y objetivo de los ODS, continuando con conceptos relacionados a los sistemas de información, siguiendo con la definición de Inteligencia de Negocio y precisando acerca de la arquitectura de una solución de Inteligencia de Negocio, así como a qué tipo de usuarios está dirigida. Se define lo que son indicadores, sus características y tipos. También, se definen las herramientas tecnológicas que facilitan la construcción de la solución de inteligencia de negocio desarrollada.

Capítulo 3. Marco Metodológico: en este capítulo, se presenta el enfoque de desarrollo propuesto por Ralph Kimball, denominado “Ciclo de Vida Dimensional del Negocio” la cual es la metodología que más se adapta al desarrollo de la solución de Inteligencia de Negocio para la obtención de Indicadores relacionados a los ODS.

Capítulo 4. Marco Aplicativo: en este capítulo, se explican cada una de las actividades que se llevaron a cabo siguiendo el ciclo de vida dimensional de Ralph Kimball para el desarrollo de la solución de Inteligencia de Negocio propuesta para la obtención de Indicadores relacionados con los ODS.

Por último, se presentan las conclusiones y beneficios obtenidos tras la elaboración de este TEG. También se presentan los posibles trabajos futuros, anexos y referencias bibliográficas.

CAPÍTULO 1: PROBLEMA DE INVESTIGACIÓN

1.1 Planteamiento del problema

En septiembre del 2015, las Naciones Unidas acordaron la creación de la Agenda 2030 para el Desarrollo Sostenible, la cual cuenta con 17 Objetivos de Desarrollo Sostenible (ODS), 169 metas y 231 indicadores, que llaman a cambiar nuestro estilo de desarrollo, respetando el medio ambiente y colocando la igualdad y dignidad de las personas en el centro.

La Agenda 2030 para el Desarrollo Sostenible es un compromiso universal adquirido por todos los países del mundo, sin distinción, desde los más desarrollados a los menos desarrollados. Cada gobierno de cada país debe dar respuestas para el seguimiento y medición sobre el cumplimiento de la Agenda 2030, basados en mecanismos de planificación existentes, y estrategias nacionales de desarrollo con el apoyo de la ayuda internacional en la formación de capacidad. Los Estados presentan informes periódicos de sus progresos, que incluyan las contribuciones nacionales y subnacionales, y de todos los sectores de la población.

Ante esta situación, los países deben estar en la capacidad de evaluar el progreso de cada uno de los 231 indicadores de la Agenda 2030 de forma precisa y oportuna para la toma de decisiones y dar respuesta al requisito de informes periódicos solicitados por las Naciones Unidas. En muchas instancias, los datos necesarios para lograr obtener respuesta de cada uno de los 231 indicadores se encuentran almacenados en archivos planos o en hojas de cálculo, lo que retrasa la obtención oportuna de la información. Es por esto que se hace necesario una solución de Inteligencia de Negocio que permita realizar la carga de los datos en una estructura que sea eficaz para su consulta y permita la visualización oportuna de la información para poder conocer de forma rápida los avances que se han realizado en cada uno de los aspectos considerados en los ODS y tomar decisiones oportunas para seguir avanzando hacia el logro de estos objetivos.

En la actualidad, Venezuela no cuenta con una solución de Inteligencia de Negocio que permita la evaluación del progreso realizado hacia el logro de los indicadores propuestos en la Agenda 2030, lo que motiva al desarrollo de dicha solución de Inteligencia de Negocio para Venezuela.

La entidad gubernamental encargada de publicar los datos de los indicadores de la Agenda 2030 es el INE, por lo que el desarrollo de la solución se encuentra enfocada hacia las necesidades y disponibilidad de los datos de dicha institución.

El INE se encarga de proveer los datos necesarios, tanto aquellos generados por ellos, como los que son generados por otras instituciones públicas del país. Muchos de estos datos no se encuentran disponibles para dar respuesta a los 17 ODS, por lo que se enfoca la solución de Inteligencia de Negocio hacia aquellos ODS de los cuales el INE posee mayor información, los cuales son:

ODS 1.- Fin de la Pobreza
ODS 3.- Salud y Bienestar
ODS 4.- Educación de Calidad
ODS 8.- Trabajo Decente y Crecimiento Económico

Por otro lado, aun dentro de esos cuatro ODS seleccionados, muchas veces es difícil obtener respuesta para un indicador directo relacionado a un ODS. En estos casos se utiliza un indicador complementario para medir el objetivo. Un indicador complementario es aquel que es capaz de dar respuesta de forma indirecta a los indicadores ODS, ya sea respondiéndolos de forma parcial, o respondiendo aspectos relacionados a lo que se busca obtener. En total entonces, se manejan 123 indicadores pertenecientes a 4 ODS, y esos 123 indicadores se dividen a su vez en 91 complementarios y 32 principales.

1.2 Objetivos

A continuación, se presenta el objetivo general y los objetivos específicos propuestos para el TEG.

1.2.1 General

Desarrollar una solución de Inteligencia de Negocio que permita la obtención, análisis y visualización de los Indicadores asociados a los Objetivos de Desarrollo Sostenible para apoyar a la toma de decisiones por parte del INE.

1.2.2 Específicos

- Analizar los requerimientos funcionales y no funcionales de la solución de Inteligencia de Negocio.
- Diseñar el modelo dimensional de acuerdo a los indicadores de los ODS.
- Diseñar e implementar el almacén de datos que contendrá los datos históricos de un país que dan respuesta a los indicadores de los ODS.
- Crear los procesos de extracción, transformación y carga de datos en el almacén de datos, con el fin de unificar los datos provenientes de distintas fuentes para ser usados por las herramientas de visualización.
- Construir las visualizaciones y cuadros de mando asociados a los indicadores de los ODS.
- Realizar las pruebas de aceptación y de calidad de datos.

1.3 Solución propuesta

La solución propuesta, dada la problemática planteada, consta de una arquitectura de dos ambientes, como se presenta en la Figura 1.

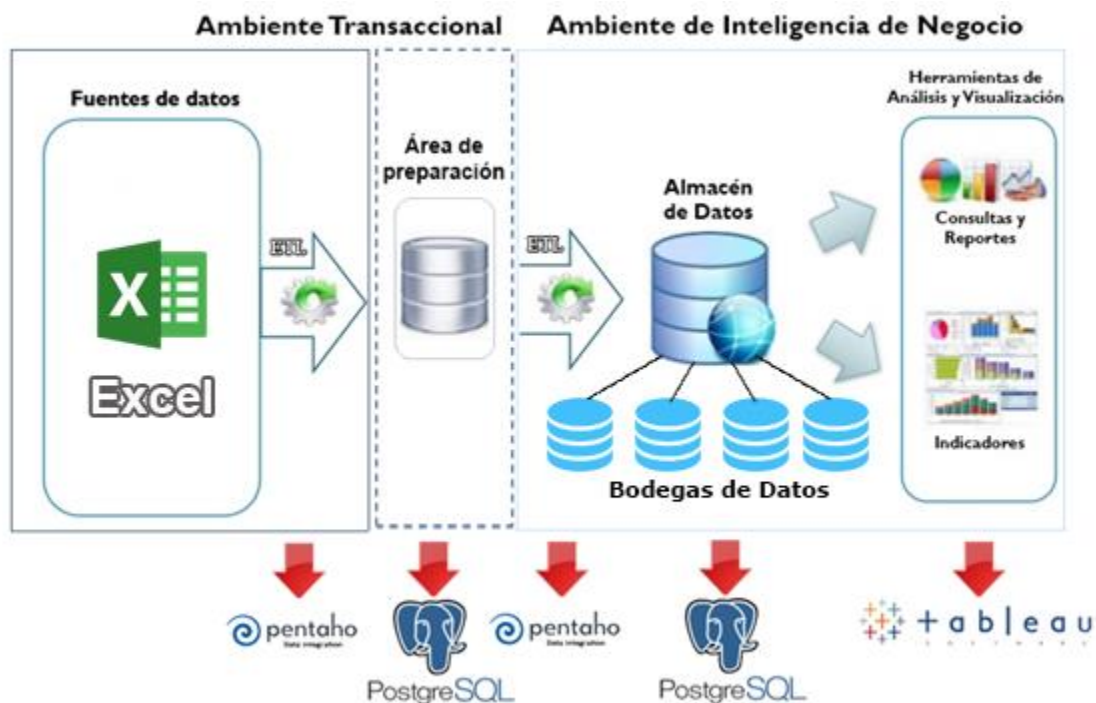


Figura 1. Arquitectura propuesta de la solución de Inteligencia de Negocio.

El primer ambiente es el transaccional, donde se realiza el análisis de las fuentes de datos, estas fuentes son suministradas en formato Excel, y contienen los datos que dan respuesta, de forma directa o indirecta, a los indicadores ODS. Al ser datos diversos, y provenientes de diferentes fuentes, es necesario utilizar un área de preparación de datos en donde se realiza la integración de estos por medio de procesos ETL, para los cuales se utiliza la herramienta Pentaho Data Integration versión 8.2, por ser una herramienta de software libre que brinda los requerimientos necesarios para poder gestionar la extracción, transformación y carga de datos, tanto para el área de preparación como para el almacén de datos.

Una vez cargados los datos en el área de preparación, se pasa al segundo ambiente, el cual es el ambiente de Inteligencia de Negocio que consta de dos componentes: el almacén de datos conformado por la unión de las bodegas de datos y las herramientas de acceso o análisis de datos. El almacén de datos es cargado con los datos provenientes del área de preparación y usando la herramienta de ETL antes mencionada. Este almacén de datos es diseñado mediante un modelo dimensional, siguiendo las mejores prácticas en el área.

Por último, se elaboran visualizaciones y cuadros de mando que den respuesta a los indicadores de los ODS, utilizando la herramienta Tableau versión 2019.2, con licencia gratuita para estudiantes. Al contar con estos indicadores se pueden monitorear los avances hacia los Objetivos de la Agenda 2030 de Venezuela.

Esta solución se implementa usando el Sistema Manejador de Base de Datos (SMBD), PostgreSQL versión 9.4, para crear las estructuras que contienen los datos utilizados en la solución, tanto para el área de preparación como para el almacén de datos.

La metodología utilizada, como se observa en la Figura 2, es la propuesta por Ralph Kimball, llamada Ciclo de Vida Dimensional del Negocio (Kimball y Margy, 2013), ya que establece los estándares necesarios para el análisis, diseño, construcción e implementación de una solución de Inteligencia de Negocio.

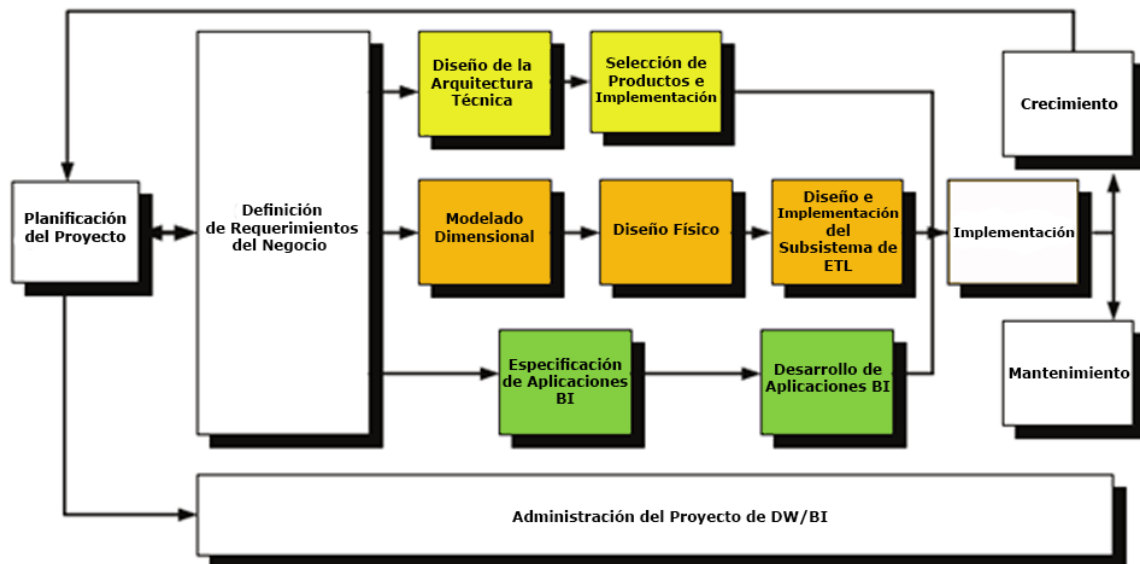


Figura 2. Ciclo de vida dimensional del negocio.

1.4 Justificación de la solución

Una solución de Inteligencia de Negocio que permita obtener la información de manera precisa y oportuna para la toma de decisiones estratégicas en cuanto a los indicadores concernientes a 4 de los ODS conlleva a una gran variedad de beneficios, tales como: tener los datos almacenados, integrados y en un mismo formato, centralización de “la verdad”, confianza en la información, así como crear y manejar diversas visualizaciones para los indicadores propuestos.

Con el desarrollo de la arquitectura anterior, se ofrecen unos módulos tecnológicos los cuales pueden ayudar a la confiabilidad de los datos, automatizar el procesamiento de los datos, y guardarlos finalmente en un almacén de datos que puede dar respuesta a las consultas analíticas que se realizan para el despliegue de los indicadores planteados.

Por último, se tienen las visualizaciones y cuadros de mando en donde se publican los indicadores y consultas requeridas. El uso de estas herramientas permite la generación de contenidos dinámicos tales como gráficos, tablas dinámicas, mapas, filtros y parámetros.

1.5 Alcance de la solución

La solución de Inteligencia de Negocio propuesta que apoya a la obtención, visualización y análisis de los ODS es desarrollada específicamente para el INE, y abarca los indicadores asociados a los siguientes ODS:

ODS 1.- Fin de la Pobreza

ODS 3.- Salud y Bienestar

ODS 4.- Educación de Calidad

ODS 8.- Trabajo Decente y Crecimiento Económico

Los datos a utilizados son aquellos suministrados por el INE, los cuales son considerados de libre acceso para toda la población. Si no se poseen los datos necesarios para dar respuesta a algún indicador de forma directa, se utilizan indicadores complementarios que estén asociados con lo que se quiere responder con el indicador principal.

CAPÍTULO 2: MARCO CONCEPTUAL

2.1 Desarrollo sostenible

2.1.1 Definición

El término de desarrollo sostenible fue utilizado por primera vez por la Comisión Mundial de Medio Ambiente y Desarrollo de Naciones Unidas (1987) donde lo definen como aquel que permite satisfacer las necesidades de las generaciones presentes sin comprometer las posibilidades de las generaciones del futuro para atender sus propias necesidades.

Según Gómez, C. (2006) el objetivo del desarrollo sostenible es definir proyectos viables y reconciliar los aspectos económico, social, y ambiental de las actividades humanas; se trata de progresar en estos ámbitos sin tener que destruir el medio ambiente. En la Figura 3 se pueden ver los "tres pilares" que deben ser tenidos en cuenta tanto por las empresas, como por las comunidades y las personas son:

- **Sostenibilidad económica:** se da cuando la actividad que se mueve hacia la sostenibilidad ambiental y social, y es financieramente posible y rentable.
- **Sostenibilidad social:** basada en el mantenimiento de la cohesión social y de su habilidad para trabajar en la persecución de objetivos comunes. Implica la mitigación de impactos sociales negativos causados por la actividad que se desarrolla, así como la potencialización de los impactos positivos.
- **Sostenibilidad ecológica:** compatibilidad entre la actividad considerada y la preservación de la biodiversidad y de los ecosistemas, evitando la degradación de las funciones fuente y sumidero. Incluye un análisis de los impactos derivados de la actividad considerada en términos de flujos, consumo de recursos difícil o lentamente renovables, así como en términos de generación de residuos y emisiones. Este último pilar es necesario para que los otros dos sean estables.



Figura 3. Esquema de los tres pilares del desarrollo sostenible.

La Comisión Mundial de Medio Ambiente y Desarrollo de Naciones Unidas en 1987 plantea que la justificación del desarrollo sostenible proviene del hecho de que el hombre habita en un planeta finito, pero tiene patrones de consumo desmedidos. En la Tierra se tienen recursos naturales limitados (nutrientes en el suelo, agua potable, minerales, entre otros.), susceptibles de agotarse. Otro factor es el hecho de la creciente actividad económica sin más criterio que el económico mismo, tanto a escala local como planetaria. El impacto negativo en el planeta puede producir graves problemas medioambientales que resulten incluso irreversibles.

Por ello, se proponen tres reglas básicas en relación con el ritmo que se debe de seguir en dicho desarrollo:

1. Ningún recurso renovable deberá utilizarse a un ritmo superior al de su generación.
2. Ningún contaminante deberá producirse a un ritmo superior al que pueda ser reciclado, neutralizado o absorbido por el medio ambiente.
3. Ningún recurso no renovable deberá aprovecharse a mayor velocidad de la necesaria para sustituirlo por un recurso renovable utilizado de manera sostenible.

Es decir, al momento de utilizar un recurso, hay que asegurarse de que se está reponiendo al mismo ritmo del que se está utilizando, y de igual forma neutralizar los agentes contaminantes que se expulsan al medio ambiente al mismo ritmo en el que se generan.

2.1.2 Objetivos de Desarrollo Sostenible

La evaluación del grado de desarrollo sostenible alcanzado por un país o una región plantea Gómez, C. (2006), resulta un tema complejo y no totalmente resuelto por varias razones. De una parte, el concepto de desarrollo sostenible en sí mismo resulta algo ambiguo y susceptible a diversas interpretaciones, en dependencia de quien realiza la evaluación. Por otro lado, integrar aspectos económicos, sociales y ecológico-ambientales en una evaluación no es una tarea fácil, pues entraña valorar aspectos que se miden en unidades distintas y cuya importancia relativa depende también del criterio del observador.

Debido a esta dificultad las Naciones Unidas decidieron crear en el 2015 una serie de objetivos, conocidos como los Objetivos de Desarrollo Sostenible (ODS) en donde se fijan metas de los tres ámbitos, el económico, social y ecológico, que pueden ser representadas mediante indicadores que permiten su evaluación.

Según el Programa de Naciones Unidas para el Desarrollo (2019), los ODS son un llamado universal a la adopción de medidas para poner fin a la pobreza, proteger el planeta y garantizar que todas las personas gocen de paz y prosperidad. Son 17 Objetivos de Desarrollo Sostenible con 169 metas planteando el horizonte 2015-2030 para ser aplicados por todos los países del mundo ya sean ricos o pobres.

Los ODS son los siguientes (Ver Figura 4):



Figura 4. Objetivos de Desarrollo Sostenible.

- **Objetivo 1 – Fin de la Pobreza:** poner fin a la pobreza en todas sus formas y en todo el mundo.
- **Objetivo 2 – Hambre Cero:** poner fin al hambre, lograr la seguridad alimentaria y la mejora de la nutrición y promover la agricultura sostenible.
- **Objetivo 3 – Salud y Bienestar:** garantizar una vida sana y promover el bienestar de todos a todas las edades.
- **Objetivo 4 – Educación de Calidad:** garantizar una educación inclusiva y equitativa de calidad y promover oportunidades de aprendizaje permanente para todos.
- **Objetivo 5 – Igualdad de Género:** lograr la igualdad de género y empoderar a todas las mujeres y las niñas.
- **Objetivo 6 – Agua Limpia y Saneamiento:** garantizar la disponibilidad y la gestión sostenible del agua y el saneamiento para todos.
- **Objetivo 7 – Energía Asequible y No Contaminante:** garantizar el acceso a una energía asequible, fiable, sostenible y moderna para todos.
- **Objetivo 8 – Trabajo Decente y Crecimiento Económico:** promover el crecimiento económico sostenido, inclusivo y sostenible, el empleo pleno y productivo y el trabajo decente para todos.
- **Objetivo 9 – Industria, Innovación e Infraestructura:** construir infraestructuras resilientes, promover la industrialización inclusiva y sostenible y fomentar la innovación.

- **Objetivo 10 – Reducción de las Desigualdades:** reducir la desigualdad en los países y entre ellos.
- **Objetivo 11 – Ciudades y Comunidades Sostenibles:** lograr que las ciudades y los asentamientos humanos sean inclusivos, seguros, resilientes y sostenibles.
- **Objetivo 12 – Producción y Consumo Responsables:** garantizar modalidades de consumo y producción sostenibles.
- **Objetivo 13 – Acción por el Clima:** adoptar medidas urgentes para combatir el cambio climático y sus efectos.
- **Objetivo 14 – Vida Submarina:** conservar y utilizar sosteniblemente los océanos, los mares y los recursos marinos para el desarrollo sostenible.
- **Objetivo 15 – Vida de Ecosistemas Terrestres:** proteger, restablecer y promover el uso sostenible de los ecosistemas terrestres, gestionar sosteniblemente los bosques, luchar contra la desertificación, detener e invertir la degradación de las tierras y detener la pérdida de biodiversidad.
- **Objetivo 16 – Paz, Justicia e Instituciones Sólidas:** promover sociedades pacíficas e inclusivas para el desarrollo sostenible, facilitar el acceso a la justicia para todos y construir a todos los niveles instituciones eficaces e inclusivas que rindan cuentas.
- **Objetivo 17 – Alianzas para lograr los Objetivos:** fortalecer los medios de implementación y revitalizar la alianza mundial para el desarrollo sostenible.

Para asegurarse que los diferentes países cumplan con cada uno de los ODS planteados, es necesario utilizar una solución de Inteligencia de Negocio que permita evaluar, medir y conocer el estado actual en el que se encuentra dichos países en cada una de las categorías correspondientes a cada ODS.

2.1.3 Antecedentes

En los últimos años, distintos países han realizado sus propios sistemas que ayudan a conocer el estado actual en que se encuentran con respecto a los ODS propuestos en la Agenda 2030. Aun cuando no existe un sistema genérico que deba ser seguido por todos y cada uno de los países, es importante conocer acerca de los sistemas ya existentes para ser tomados como base en el presente trabajo de investigación. A continuación, se detallan los antecedentes que se utilizan como referencia.

2.1.3.1 Perú: Sistema de Monitoreo y Seguimiento de los Indicadores de los Objetivos de Desarrollo Sostenible (ODS)

El Sistema de Monitoreo y Seguimiento de los Indicadores de los ODS realizado por el Instituto Nacional de Estadística e Informática (INEI) de Perú en 2018 es un sistema donde la información es completamente de libre acceso y publicada en el Internet. Al ingresar al sistema, como se puede ver en la Figura 5, se da una breve introducción a lo que son los ODS y la motivación de crear el sistema de monitoreo.



Figura 5. Ventana principal Sistema ODS Perú.

Una vez dentro del sistema, se encuentra el menú para seleccionar el ODS que se quiera visualizar, haciendo uso de los logos y nombres propuestos por las Naciones Unidas, como se puede ver en la Figura 6.



Figura 6. Menú de selección de ODS.

Al seleccionar uno de los ODS, como se aprecia en la Figura 7, el sistema cuenta con un menú donde se despliegan todas las metas asociadas a ese ODS, y al seleccionar una de ellas, se despliegan los indicadores asociados a esa meta. Si el sistema no cuenta con los datos para mostrar el estado de Perú en ese indicador, se muestra con texto rojo.

PERÚ: SISTEMA DE MONITOREO Y SEGUIMIENTO DE LOS INDICADORES DE LOS OBJETIVOS DE DESARROLLO SOSTENIBLE

1 FIN DE LA POBREZA

OBJETIVO 1: PONER FIN A LA POBREZA EN TODAS SUS FORMAS Y EN TODO EL MUNDO

En el Perú los índices de pobreza extrema se han reducido de manera considerable hasta ubicarse en 3,8% en el año 2016. Si bien se trata de un logro notable, 1 de cada 25 peruanos aún vive en situación de pobreza extrema.

La pobreza va más allá de la falta de ingresos y recursos para garantizar unos medios de vida sostenibles. Entre sus manifestaciones se incluyen el hambre y la malnutrición, el acceso limitado a la educación y a otros servicios básicos, la discriminación y la exclusión sociales y la falta de participación en la adopción de decisiones. El crecimiento económico debe ser inclusivo con el fin de crear empleos sostenibles y promover la igualdad.

RESUMEN DE OBJETIVO N°1

OBJETIVOS DE DESARROLLO SOSTENIBLE		INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA		
N° Metas	N° Indicadores	N° Indicadores	N° Sub - Indicadores	N° Indicadores en Construcción
7	14	9	30	5

Figura 7. Menú de indicadores ODS 1 - Fin de la Pobreza.

Una vez seleccionado un indicador en particular, podemos ver diferente información asociada a él, como lo son cuadros estadísticos, visualizaciones, mapas, información acerca del cálculo y responsables del indicador, los errores asociados al cálculo del indicador, los subindicadores y programas presupuestales, tal como se muestran en las Figuras 8, 9, 10, 11, 12, 13 y 14. Al ser información de libre acceso y de uso público, también existe la opción de descargar los datos.

1.2.1 Incidencia de la pobreza monetaria total
(Porcentaje)

Nacional Área de residencia Región natural Dominio Geográfico

Departamento

[Descargar en excel](#)

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
NACIONAL														
Nacional	58.7	55.6	49.1	42.4	37.3	33.5	30.8	27.8	25.8	23.9	22.7	21.8	20.7	20.2

Fuente: Instituto Nacional de Estadística e Informática - Encuesta Nacional de Hogares (ENAHOG)

Figura 8. Cuadro estadístico de indicadores.

1.2.1 Incidencia de la pobreza monetaria total
(Porcentaje)

- Nacional
 Área de residencia
 Región natural
 Dominio Geográfico
 Departamento

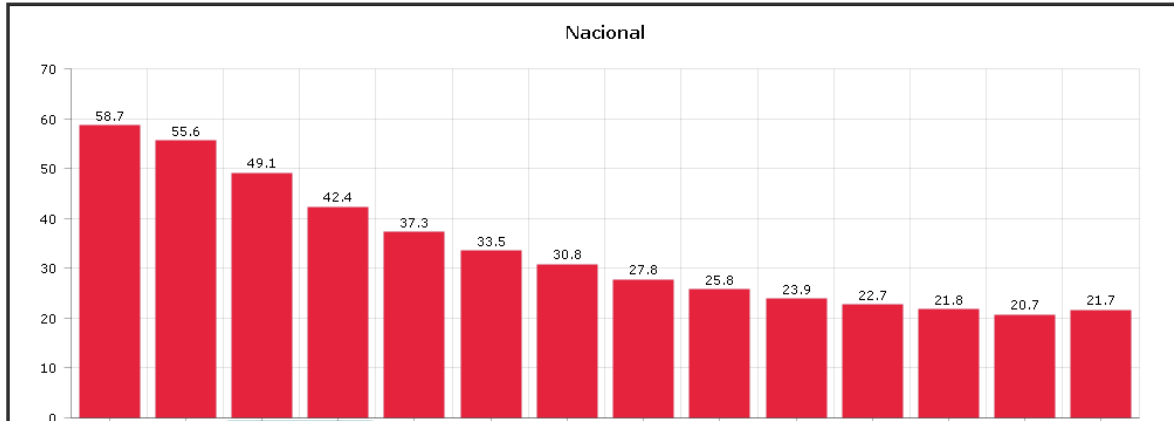


Figura 9. Visualizaciones de indicadores.

1.2.1 Incidencia de la pobreza monetaria total

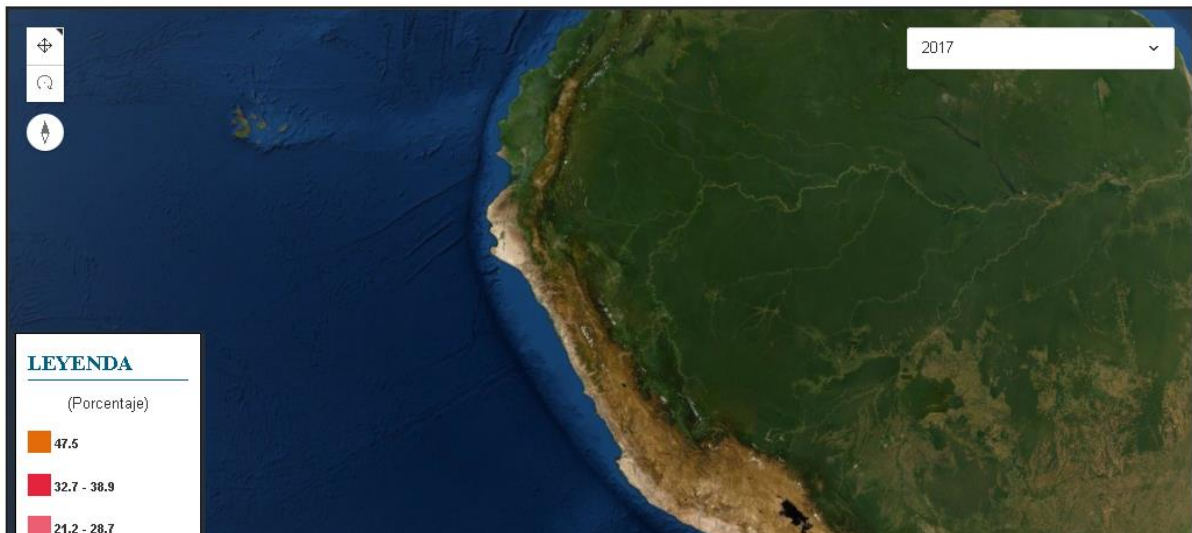


Figura 10. Mapas de indicadores.

1.2.1 Incidencia de la pobreza monetaria total

[Descargar en excel](#)

Nombre del Indicador INEI	Incidencia de la pobreza monetaria total
Definición	Es la proporción de la población con gasto per cápita por debajo del valor de la Línea de Pobreza (LP).
Algoritmo	$H = \frac{Q}{N} \times 100$
Variables	H: Proporción de la población en situación de pobreza Q: Población con gasto per cápita inferior al valor de la línea de pobreza (LP) N: Total de población
Unidad de Medida	Porcentaje.
Fuente de los datos	Instituto Nacional de Estadística e Informática - Encuesta Nacional de Hogares (ENAHOG)

Figura 11. Información acerca de metodología de cálculo de indicadores.

1.2.1 Incidencia de la pobreza monetaria total

(Porcentaje)

[Descargar en Excel](#)

Año	Variable	Estimador	Intervalo inferior	Intervalo superior	Coficiente de variación
2004	Nacional	58.7	57.3	60.1	1.2
2005	Nacional	55.6	54.1	57.1	1.4
2006	Nacional	49.1	47.6	50.7	1.6
2007	Nacional	42.4	41.2	43.7	1.5
2008	Nacional	37.3	36.1	38.5	1.7
2009	Nacional	33.5	32.2	34.8	1.9
2010	Nacional	30.8	29.6	32.0	2.0
2011	Nacional	27.8	26.8	28.8	1.9

Figura 12. Errores muestrales de indicadores.

CUADROS GRAFICOS MAPAS METADATO ERRORES MUESTRALES **SUB INDICADORES** PROGRAMAS PRESUPUESTALES

ELEGIR UN SUB INDICADOR...

- 1.2.1.1 Incidencia de la pobreza monetaria total según sexo
- 1.2.1.2 Incidencia de la pobreza monetaria total según grupos de edad
- 1.2.1.3 Incidencia de la pobreza monetaria total según lengua materna
- 1.2.1.4 Incidencia de la pobreza monetaria total según auto percepción étnica
- 1.2.1.5 Grupos de departamentos con niveles de pobreza monetaria total estadísticamente semejantes
- 1.2.1.6 Incidencia de la pobreza monetaria total según condición de discapacidad
- 1.2.1.7 Incidencia de la pobreza monetaria total de la población joven de 15 a 24 años de edad ocupada
- 1.2.1.8 Incidencia de la pobreza monetaria total de la población adulto mayor de 60 a más años de edad ocupada
- 1.2.1.9 Incidencia de la pobreza monetaria total de las mujeres jefas de hogar monoparental
- 1.2.1.10 Incidencia de la pobreza monetaria total de jefatura de hogar monoparental según tipo de hogar

INGRESAR

Figura 13. Selección de sub-indicadores.

CUADROS GRAFICOS MAPAS METADATO ERRORES MUESTRALES SUB INDICADORES **PROGRAMAS PRESUPUESTALES**

ELEGIR UN INDICADOR DE UN PROGRAMA PRESUPUESTAL...

- 0098-288 Porcentaje de cobertura de la población objetivo que reciben el servicio de Acompañamiento a Familias
- 0098-289 Porcentaje de cobertura de la población objetivo que reciben el servicio de cuidado diurno
- 0049-121 Porcentaje de mujeres de hogares usuarios Juntos que en último nacimiento en los 5 años antes de la encuesta recibieron 6 o más controles prenatale
- 0049-122 Porcentaje de niños de hogares usuarios JUNTOS que cuentan con 6 años de edad y cursan por primera vez el primer grado del nivel primario de Educaci
- 0049-123 Porcentaje de niños en edad pre-escolar (3-5 años) de hogares usuarios JUNTOS que asisten a servicios de educación inicial
- 0049-124 Porcentaje de niños menores de 36 meses de hogares usuarios JUNTOS que asisten al menos al 80% de los Controles de Crecimiento y Desarrollo se
- 0049-125 Proporción de deserción escolar de niños y adolescentes de hogares usuarios Juntos que culminaron el nivel primario de educación básica regular el añ
- 0049-126 Proporción de deserción escolar del nivel secundario de adolescentes de hogares usuarios Juntos

INGRESAR

Figura 14. Selección de programas presupuestales.

El sistema de ODS de Perú es bastante completo, aun cuando no cuente con los datos para dar respuesta a todos los indicadores, está diseñado de tal manera que puede ser modificado para agregar datos a los indicadores que actualmente no los posean. Se puede considerar como un buen sistema para utilizar de base para el desarrollo de sistemas ODS de cualquier país.

2.1.3.2 México: Sistema de Información de los Objetivos de Desarrollo Sostenible

El Sistema de Información de los ODS desarrollado por el Instituto Nacional de Estadística y Geografía (INEGI) de México cumple con las mismas funcionalidades presentes en el sistema anterior, tales como la posibilidad de acceder a todos los datos y visualizaciones de los ODS, metas e indicadores de forma libre, pero con una interfaz y forma de navegación diferente. En la Figura 15 se puede ver la ventana principal del sistema, la cual nos permite acceder a los ODS de forma directa.



Figura 15. Ventana principal del Sistema ODS México.

Al seleccionar cualquiera de los ODS, se presenta una ventana emergente, como se ven en la Figura 16, en donde podemos seleccionar el indicador que se quiera visualizar, agrupados por sus respectivas metas, cabe acotar que no se muestran todos los indicadores ni metas que están asociados a ese ODS, solo se presentan aquellos para los cuales el sistema posee datos.

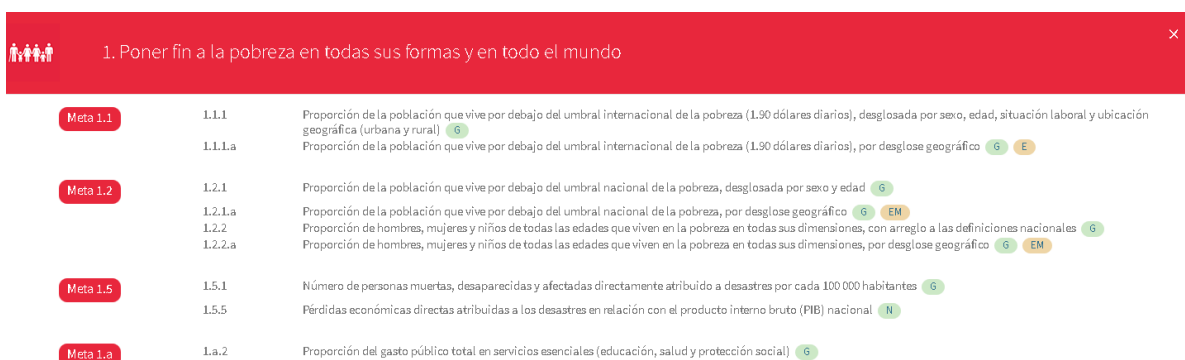


Figura 16. Ventana emergente de selección de indicadores.

Cuando se selecciona un indicador en particular, el sistema redirige a otra ventana, donde se puede ver información referente a este indicador, como se muestra en las Figuras 17, 18, 19 y 20.

1.1.1 Proporción de la población que vive por debajo del umbral internacional de la pobreza (1.90 dólares diarios), desglosada por sexo, edad, situación laboral y ubicación geográfica (urbana y rural) [G](#)

Indicador Metadato Datos para el cálculo

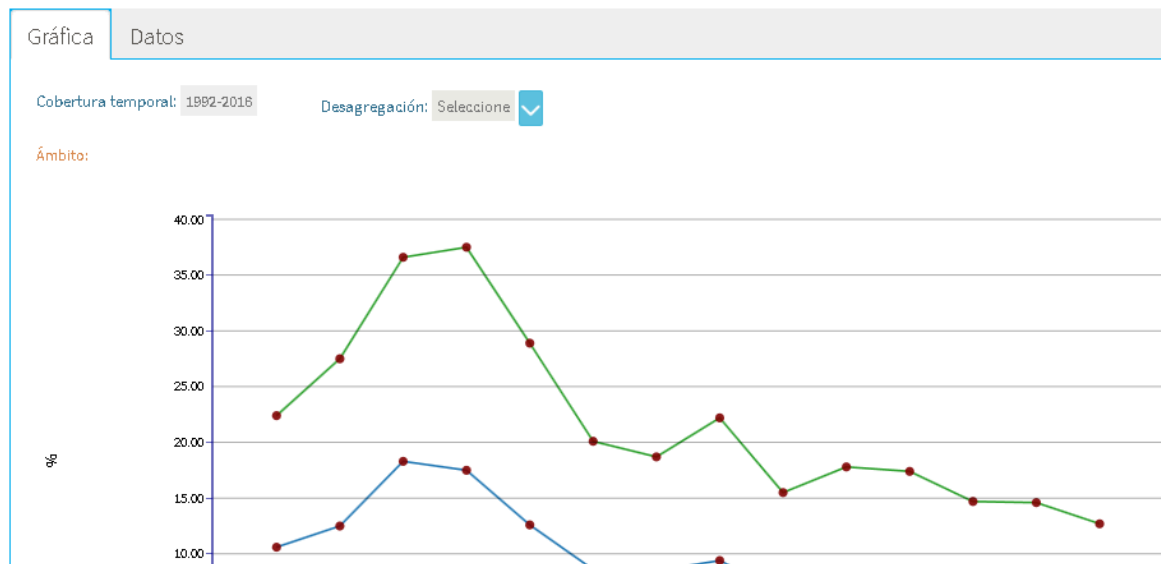


Figura 17. Visualización asociada a un indicador.

1.1.1 Proporción de la población que vive por debajo del umbral internacional de la pobreza (1.90 dólares diarios), desglosada por sexo, edad, situación laboral y ubicación geográfica (urbana y rural) [G](#)

Indicador Metadato Datos para el cálculo

Gráfica Datos

(Porcentaje) Cobertura temporal: 1992-2016

Periodo	Ámbito		
	Total	Urbana	Rural
2016	5.8	1.7	12
2014	6.4	1.6	14
2012	6.5	1.6	14
2010	7.9	2.2	17
2008	7.9	1.9	17
2006	6.6	1.3	18
2005	9.4	1.8	22
2004	8.5	2.4	18
2002	8.6	1.5	20
2000	12.6	2.1	28
1998	17.5	4.7	37
1996	18.3	6.5	36
1994	12.5	2.5	27
1992			

Figura 18. Cuadro estadístico asociado a una visualización.

1.1.1 Proporción de la población que vive por debajo del umbral internacional de la pobreza (1.90 dólares diarios), desglosada por sexo, edad, situación laboral y ubicación geográfica (urbana y rural) ⁶

Indicador	Metadato	Datos para el cálculo
Objetivo	Poner fin a la pobreza en todas sus formas y en todo el mundo	
Meta	De aquí a 2030, erradicar para todas las personas y en todo el mundo la pobreza extrema (a partir de octubre de 2015, se considera que sufren pobreza extrema las personas que viven con menos de 1.90 dólares de los Estados Unidos al día)	
Nombre del indicador	Proporción de la población que vive por debajo del umbral internacional de la pobreza (1.90 dólares diarios), desglosada por sexo, edad, situación laboral y ubicación geográfica (urbana y rural)	
Definición	Es el porcentaje de la población que en el año del cálculo vive con menos del equivalente a 1.90 dólares diarios a precios de 2011, según la paridad del poder adquisitivo.	
Tipo de indicador	Global	
Algoritmo	$PPID_t = \left(\frac{Q_t}{N_t} \right) 100$ <p>Significado de las siglas o abreviaturas: PPID_t: Proporción de la población de referencia que vive por debajo del umbral internacional de la pobreza en el año t Q_t: Población de referencia con ingreso per cápita inferior a 1.90 dólares en el año t N_t: Población de referencia total en el año t</p>	
Descripción narrativa del cálculo del indicador	Este indicador resulta de dividir el número de personas con ingreso per cápita inferior a 1.90 dólares diarios a precios de 2011, entre la población de referencia total, multiplicado por 100.	
Unidad de medida	Porcentaje	
Cobertura geográfica	Nacional	
Referencia temporal	1992-2016	

Figura 19. Información sobre el cálculo y obtención de un indicador.

1.1.1 Proporción de la población que vive por debajo del umbral internacional de la pobreza (1.90 dólares diarios), desglosada por sexo, edad, situación laboral y ubicación geográfica (urbana y rural) ⁶

Indicador	Metadato	Datos para el cálculo
Serie para el cálculo		
<input checked="" type="radio"/> INPC (2010=100, agosto de cada año) <input type="radio"/> Consumo privado a precios corrientes <input type="radio"/> Deflactor (base 2011=100) <input type="radio"/> Propensión marginal a consumir (ConsPr/PIB) <input type="radio"/> Canasta mensual (1.90 dls) <input type="radio"/> Población total <input type="radio"/> PIB a precios corrientes <input type="radio"/> Población en pobreza		
INPC (2010=100, agosto de cada año)		
Datos Cobertura temporal: 1992-2016		
Periodo	INPC (2010=100, agosto de cada año)	
2016	119.547	
2014	113.438	
2012	105.279	
2010	97.347	
2008	89.355	
2006	81.250	

Figura 20. Datos utilizados para el cálculo de un indicador.

Estos dos sistemas que se han analizado, aunque presenten ciertas diferencias en cuanto a funcionalidades e información desplegada, tienen algunos puntos claves en común, como lo son la presentación de los indicadores en formas de cuadros y gráficos. También se le da importancia a la presentación de la metodología de cálculo y obtención de los datos utilizados para obtener los indicadores. Estos aspectos sirven de base para futuros sistemas destinados al monitoreo y presentación de los ODS.

2.2 Sistema de información

Una vez analizada la problemática que se intenta resolver y conociendo bien el contexto del problema, debemos comenzar entonces a estudiar sobre aquellos elementos y/o herramientas que permiten desarrollar una solución de Inteligencia de Negocio, empezando desde lo más general, en este caso los sistemas de información.

2.2.1 Definición

Los sistemas de información, según Frost, Kenyo, Pels y Pike (2012) son la combinación de personas, tecnologías de información y procesos de negocio que logran un objetivo del negocio. Podemos representar un sistema de información como un triángulo conformado por personas, procesos y tecnología de información en cada uno de sus vértices, dicha pirámide puede ser observada en la Figura 21.



Figura 21. Triángulo de sistemas de información.

Por otro lado, se puede complementar la definición anterior con la propuesta por Devece, Guiral y Lapiedra (2011) que los define como el conjunto formal de procesos que, operando sobre una colección de datos estructurada de acuerdo con las necesidades de una empresa, recopila, elabora y distribuye la información necesaria para la operación de dicha empresa.

Dada la complejidad de los procesos de tratamiento de la información y los diferentes grados o niveles en los que, según los problemas, es posible estructurar datos y procesos, se hace necesaria la existencia de diferentes tipos de sistemas de información, capaces de abarcar la totalidad de la información que la organización precisa.

2.2.2 Tipos de sistemas de información

Existen diferentes tipos de sistemas de información, cada uno enfocado a resolver distintas necesidades de la organización, a continuación, se verán algunos de estos tipos.

2.2.2.1 Basados en los niveles organizacionales

Tal como lo dicen Devece, Guiral y Lapiedra (2011), toda organización puede ser dividida verticalmente según su estructura jerárquica, y es denominada como pirámide organizacional, la cual cuenta con tres niveles básicos como se muestra en la Figura 22.

A continuación, se mencionan diferentes sistemas de información enfocados a cada uno de los niveles de la pirámide organizacional.



Figura 22. Pirámide de los niveles organizacionales.

- **Nivel estratégico**

Este nivel, siendo el más alto, es donde se elaboran las políticas y estrategias a seguir por la organización. Se determinan los objetivos a largo plazo y se toman decisiones que afectan a toda la organización. Entre los sistemas de información que son utilizados en este nivel se encuentran:

- **Sistemas de apoyo para ejecutivos:** también conocidos como ESS, de sus siglas en inglés, *Executive Support Systems*. Son aquellos que están dirigidos a la alta dirección puesto que permiten agilizar la obtención de datos ya sean internos como externos de la empresa, más principales y estratégicos para la organización, presentándolos de la forma más clara posible con el propósito de poder dar alcance a los factores críticos de éxito (Mos, 2015).

- **Sistemas de apoyo a la toma de decisiones:** también conocidos como DSS, de sus siglas en inglés, *Decision Support Systems*. Son sistemas de información basados en un computador interactivo, flexible y adaptable, especialmente desarrollado para apoyar la solución de un problema de gestión no estructurado para mejorar la toma de decisiones. Utiliza datos, proporciona una interfaz amigable y permite la toma de decisiones en el propio análisis de la situación (Turban & Aronson, 2005).

- **Nivel táctico**

En este nivel se coordinan las actividades que se desarrollan en el nivel operacional, así como las decisiones que se toman y que afectan a un sector, área o departamento específico. En este nivel se utilizan los siguientes sistemas de información:

- **Sistemas de información gerencial:** también conocidos como MIS, de sus siglas en inglés, *Management Information Systems*. Son el resultado de interacción colaborativa entre personas, tecnologías y procedimientos, orientados a solucionar problemas corporativos y/o empresariales (Laudon y Laudon, 2014).

- **Nivel operacional**

Este nivel tiene como función la de realizar en forma eficaz las tareas que se ejecutan en la organización. Se realizan tareas rutinarias programadas previamente por el nivel táctico. Los siguientes sistemas de información son empleados en este nivel:

- **Sistemas de procesamiento de transacciones:** también conocidos como TPS, de sus siglas en inglés, *Transaction Processing Systems*. Son aquellos sistemas de información creados para procesar grandes cantidades de datos referentes operaciones diarias de negocios, como lo son nóminas, inventarios. Estos sistemas eliminan la tediosidad al momento de realizar las transacciones operativas de la empresa.

2.2.2.4 Basados en el procesamiento de datos

Este tipo de sistemas de información se basan en los datos y en cómo pueden ser procesados los mismos, ya sea para su uso en el día a día o para el apoyo para la toma de decisiones, los siguientes sistemas de información forman parte de esta clasificación:

- **Procesamiento transaccional en línea:** también conocidos como OLTP, de sus siglas en inglés, *OnLine Transaction Processing*, son aquellos sistemas que se encargan de almacenar y gestionar los datos relevantes a las operaciones del día a día de una organización. Usualmente este término se asocia a las bases de datos operacionales, las cuales tienen como objetivo principal la recolección de datos de lo que ocurrió en algún momento en específico (Franco, 2019).

Los sistemas OLTP están caracterizados por un gran número de transacciones de inserción, actualización o borrado. Su mayor énfasis está en el procesamiento veloz

de las consultas, manteniendo la integridad de los datos.

- **Procesamiento analítico en línea:** también conocidos como OLAP, de sus siglas en inglés, *OnLine Analytical Processing*, son frecuentemente utilizados para almacenar y gestionar los datos relevantes para el análisis y la toma de decisiones. Este tipo de sistemas ofrecen la capacidad de generar reportes de forma rápida, evitando la necesidad de contactar al departamento de tecnología para generar un reporte a la medida (Franco, 2019).

Los sistemas OLAP están caracterizados por un volumen bajo de transacciones, las consultas generalmente son muy complejas e involucran funciones de agregación.

A continuación, se presenta la Tabla 1, en donde se puede ver una comparación, basada en diferentes criterios, entre los sistemas OLTP y OLAP.

Tabla 1. Comparación entre sistemas OLTP y OLAP según Wang en 2008.

Criterios	OLTP	OLAP
Definición	Es un sistema transaccional en línea y gestiona modificaciones de bases de datos.	Es un sistema en línea para la recolección y el análisis de datos.
Enfoque	Insertar, actualizar y borrar información de las bases de datos.	Extraer los datos para análisis que ayuda a la toma de decisiones.
Datos	OLTP y sus transacciones son la fuente original de los datos (Datos operacionales).	Diferentes bases de datos OLTP son usadas como fuente de datos OLAP (Datos consolidados).
Transacciones	OLTP posee transacciones cortas.	OLAP posee transacciones largas.
Tiempo	El tiempo de procesamiento de una transacción es comparativamente menor.	El tiempo de procesamiento de una transacción es comparativamente mayor.
Consultas	Consultas simples.	Consultas complejas.
Normalización	Las tablas en una base de datos OLTP están normalizadas.	Las tablas en una base de datos OLAP no se encuentran normalizadas.
Propósito de los datos	Controlar y ejecutar tareas fundamentales del negocio.	Ayudar a la planificación, resolución de problemas y toma de decisiones.

Se puede ver que los sistemas OLTP y OLAP no compiten entre sí, y en vez de preguntarse cuál es mejor, es más conveniente hacerse la pregunta, “¿Cómo se complementan el uno al otro?”. Sabemos que los sistemas OLTP son usados para almacenar y gestionar los datos de las operaciones del día a día, y los sistemas OLAP son utilizados para analizar esos datos. En la Figura 23 se muestra la manera en que estos sistemas pueden trabajar en conjunto, los datos son extraídos de las bases de datos OLTP y almacenados en las bases de datos OLAP, para así lograr el análisis de los mismos.

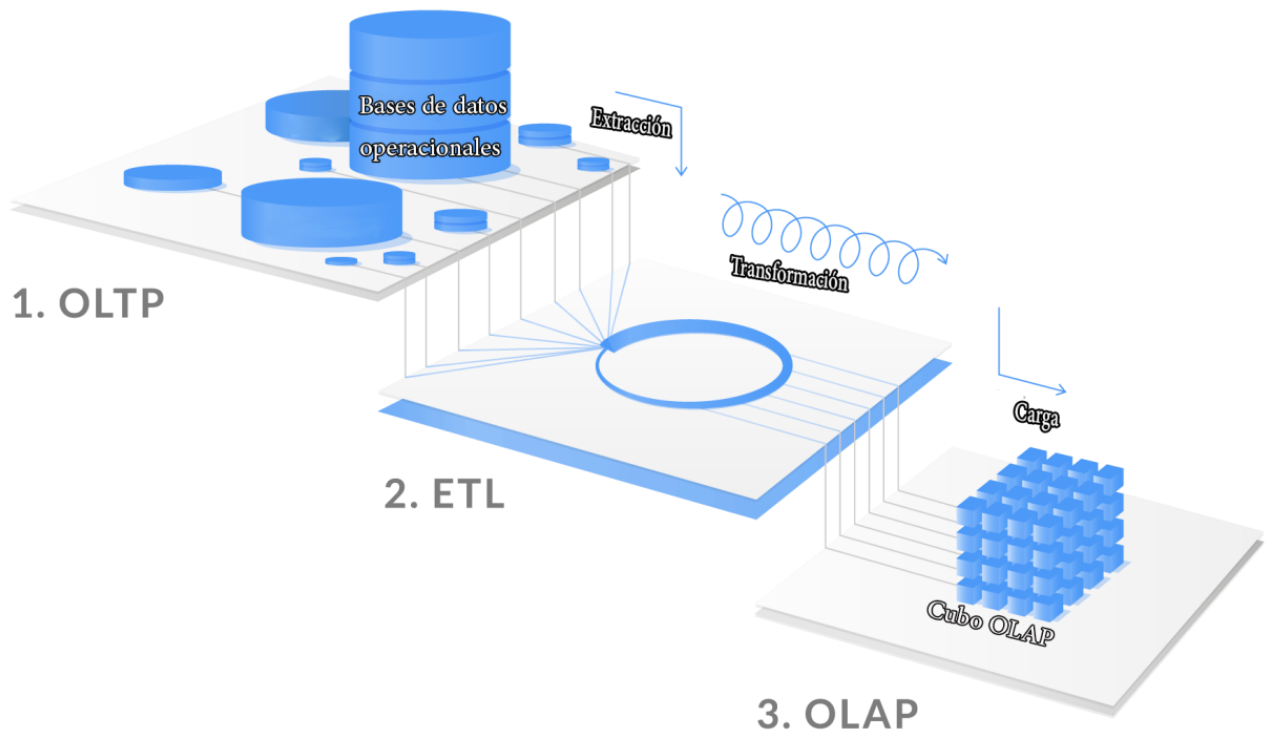


Figura 23. OLTP y OLAP trabajando juntos.

2.2.3 Beneficios de los sistemas de información

Franco, en 2019, comenta que se puede ver que los Sistemas de información son importantes en la gestión de la organización. Hoy en día, el entorno global cambia constantemente y la competencia comercial es alta. Organizaciones con sistemas de gestiona eficaces y acceso rápido a la información, utilizan este tipo de sistemas de manera efectiva debido a que la introducción de la informática en la industria ayuda a reducir los costos de producción.

A continuación, se listan algunos de los beneficios propuestos por Franco en 2019 de los sistemas de información:

- Acceso rápido a la información y por ende mejora en la atención a los usuarios.
- Mayor motivación en los mandos medios para anticipar los requerimientos de las directivas.
- Generación de informes e indicadores, que permiten corregir fallas difíciles de detectar y controlar con un sistema manual.
- Posibilidad de planear y generar proyectos institucionales soportados en sistemas de información que presentan elementos claros y sustentados.
- Evitar pérdida de tiempo recopilando información que ya está almacenada en bases de datos que se pueden compartir.
- Impulso a la creación de grupos de trabajo e investigación debido a la facilidad para encontrar y manipular la información.
- Soluciona el problema de falta de comunicación entre las diferentes instancias. A nivel directivo se hace más efectiva la comunicación

- Organización en el manejo de archivos e información clasificada por temas de interés general y particular.
- Generación de nuevas dinámicas, utilizando medios informáticos como el correo electrónico, multimedia, teleconferencia, acceso directo a bases de datos y redes nacionales e internacionales.
- Aumento de la productividad gracias a la liberación de tiempos en búsqueda y generación de información repetida.

2.3 Inteligencia de Negocio

Después de haber estudiado con respecto a los sistemas de información de forma general, podemos hacer un estudio más específico de un tipo de sistema de información que se encarga del procesamiento analítico, los conocidos como Inteligencia de Negocio.

2.3.1 Definición

El término inteligencia de negocio, más reconocido por su nombre inglés *Business Intelligence* (BI), es definida por Sinnexus (2017) como 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. En la Figura 24 se puede ver esta relación entre datos, información y conocimiento.

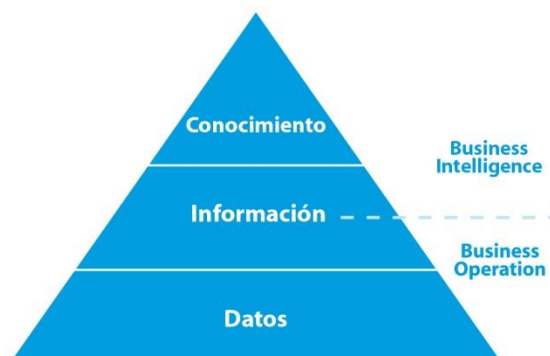


Figura 24. Relación entre datos, información y conocimiento.

Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, podemos definir la Inteligencia de Negocio 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 o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio.

La finalidad de la Inteligencia de Negocio es observar qué está ocurriendo en la organización para poder comprender por qué ocurre, predecir qué puede pasar, indicar cuáles son las posibilidades, y con base a esto poder decidir qué decisiones deben tomarse en la organización.

2.3.2 Características

Para (Loshin, 2003), toda solución de Inteligencia de Negocio debe cumplir con las siguientes características:

- **Accesibilidad a la información:** significa permitir el acceso de los usuarios finales a los datos, con independencia de la procedencia de estos.
- **Apoyo a la toma de decisiones:** implica garantizar que los usuarios finales tengan acceso a las herramientas de análisis, las cuales permiten seleccionar y manipular los datos de la empresa.
- **Orientación al usuario final:** la Inteligencia de Negocio busca la independencia entre los conocimientos técnicos de los usuarios y la capacidad para utilizar las herramientas.

2.3.3 Arquitectura de una solución de Inteligencia de Negocio

Una solución de Inteligencia de Negocio se caracteriza por presentar al menos cuatro componentes importantes, contenidos en dos ambientes.

El primer ambiente es el transaccional en el que se encuentran las fuentes de datos (que pueden ser diversas y no estructuradas) y los procesos y herramientas que permiten la consolidación de los datos provenientes de distintas fuentes y son cargados en el almacén de datos.

El segundo ambiente, conocido como ambiente de Inteligencia de Negocio está compuesto por el almacén de datos y las herramientas para presentar los datos resumidos o en detalle.

Existe un componente opcional, ubicado en el ambiente transaccional, conocido como el área intermedia o área de preparación de datos, en donde los datos provenientes de las fuentes de datos son consolidados e integrados antes de ser cargados al almacén de datos.

En la Figura 25, se diferencian los dos ambientes que contiene una arquitectura de una solución de Inteligencia de Negocio, junto con sus componentes:

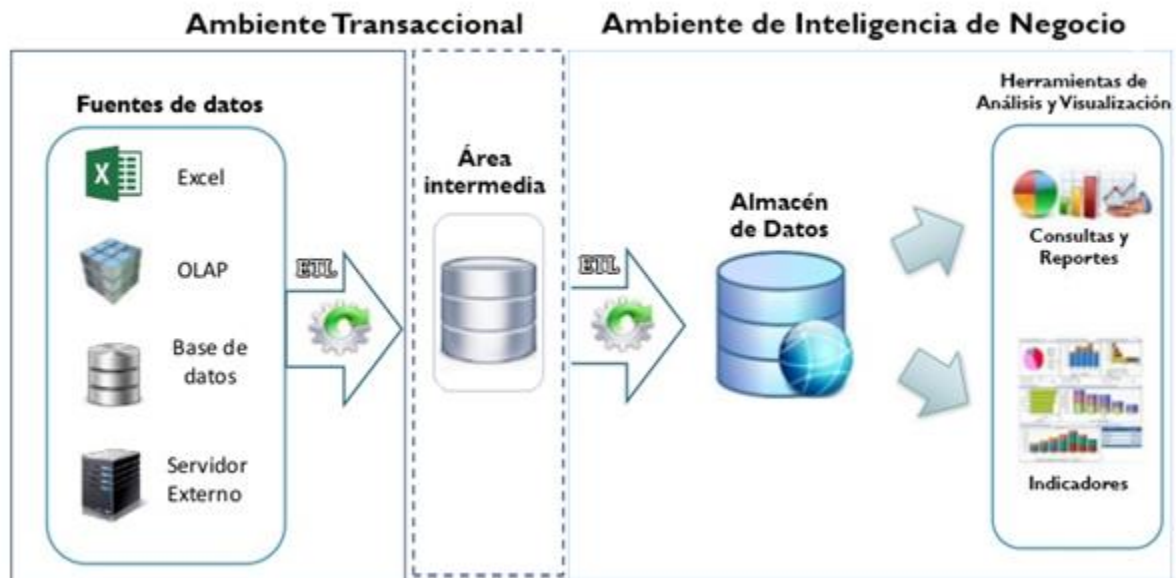


Figura 25. Arquitectura de una solución de Inteligencia de Negocio.

2.3.3.1 Fuentes de datos

En el primer componente de la arquitectura se tienen las fuentes de datos, que están representadas por aquellos sistemas que contienen datos que pueden ser explotados potencialmente por el resto de los componentes de una solución de Inteligencia de Negocio.

Generalmente, estos datos suelen presentarse en sistemas transaccionales, archivos planos, hojas de cálculos o bodegas de datos de diferentes departamentos de una organización. Estos datos no pueden ser modificados ni alterados, para mantener la integridad de estos, solo pueden ser transformados.

2.3.3.2 Procesos de extracción, transformación y carga

El segundo componente está compuesto por los procesos que se encargan de integrar los datos obtenidos de las diversas fuentes. Estos son conocidos como procesos de Extracción, Transformación y Carga, (ETL del inglés, *Extract, Transform and Load*). A continuación, se explican los 3 pasos que componen un proceso de ETL:

- **Extracción:** consiste en la recuperación de los datos físicamente de las distintas fuentes de información. Este proceso genera datos en un estado en bruto.

Según Cano (2007) la extracción de datos se puede realizar de forma manual o bien utilizando herramientas de ETL. El principal objetivo de la extracción es extraer únicamente aquellos datos de los sistemas transaccionales que son necesarios y prepararlos para el resto de los subprocesos de ETC. Para ello se deben determinar las mejores fuentes de información, las de mejor calidad. Con tal finalidad, se debe analizar las fuentes disponibles y escoger aquellas que sean adecuadas.

- **Transformación:** consiste en la recuperación de los datos limpios y de alta calidad para estructurarlos y adicionarlos en los distintos modelos de análisis de acuerdo con las reglas de negocio y los estándares que se han establecido. El proceso incluye cambios de formato, sustitución de códigos, valores derivados y agregados. Este proceso genera datos limpios, consistentes, sumariados y útiles.
- **Carga:** consiste en el almacenamiento físico de los datos transformados en el nuevo repositorio destino, este proceso incluye la integración en los distintos modelos de las áreas de negocio que se han definido.

2.3.3.3 Área intermedia

Cano en 2007, explica que normalmente, la información que se almacena en los sistemas transaccionales no está preparada para la toma de decisiones, por lo tanto, se busca almacenar los datos de una manera que expanda su flexibilidad, facilidad de acceso y administración.

El área intermedia, o también conocida como área de integración de datos, es un espacio temporal y volátil en el cual se van a ejecutar los procesos ETL. Se utiliza para hacer una primera extracción rápida de los datos fuente (sin ninguna transformación) y almacenarlos temporalmente mientras se analizan, limpian, mejoran y posteriormente se cargan en el almacén de datos.

2.3.3.4 Almacén de datos

Según Kimball (1997), un almacén de datos no es más que una copia de datos transaccionales, específicamente estructurados, utilizados para consulta y análisis.

Por su parte Inmon (1998), define un almacén de datos como un conjunto de datos Integrados orientados a materia, variantes con el tiempo y no transitorios, los cuales soportan el proceso de toma de decisiones de una administración.

Partiendo de estas definiciones, de dos de los pioneros en el área de almacenes de datos, se puede armar que ambas se centran en los datos per se. Por lo tanto, podemos definir un almacén de datos, de forma más generalizada, como un repositorio centralizado e histórico, que almacena gran cantidad de datos, utilizados para apoyar en la toma de decisiones de una organización. Generalmente, estos almacenes son caracterizados por ser la fuente que alimenta los procesos analíticos dentro de una organización.

Adicionalmente, los almacenes de datos se encargan de guardar información histórica y permiten principalmente consultas, debido a que por lo general están conectados a herramientas de análisis que extraen la información para mostrarla a gerentes o analistas para la toma de decisiones.

Es importante definir en este apartado lo que son las bodegas de datos (del inglés *Data Marts*). Las bodegas de datos, como se ve en la Figura 26, son bases de datos orientadas a consultas especializadas en un área específica del negocio y están orientadas a satisfacer necesidades específicas del negocio. Su implementación suele ser semejante a la de un almacén de datos, con la diferencia principal en que el tiempo de implementación es mucho más reducido.

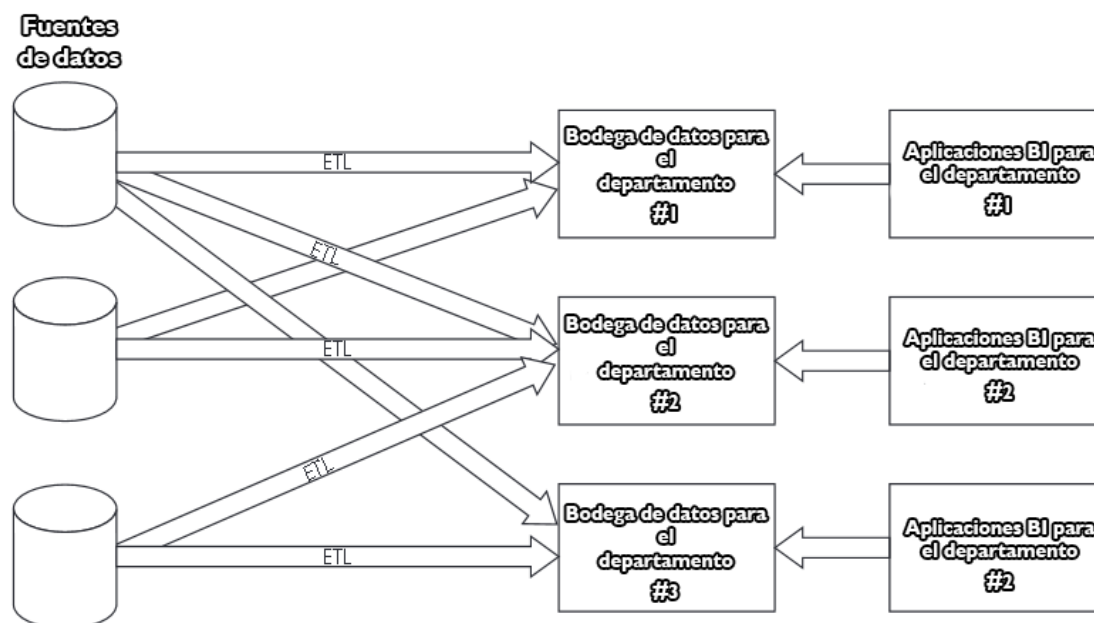


Figura 26. Implementación de bodegas de datos.

A continuación, se muestra una tabla comparativa entre los almacenes de datos y las bodegas de datos.

Tabla 2. Comparación entre almacén de datos y bodega de datos, según Ortilles en 2012.

Crterios	Almacén de datos	Bodega de datos
Definición	Es un gran repositorio de datos recolectados de diferentes departamentos dentro de una organización.	Es un subtipo de un almacén de datos. Está diseñado para satisfacer las necesidades de un departamento de la organización.
Enfoque	Se enfocan hacia todos los departamentos, e incluso a la organización entera.	Son enfocados a nivel departamental
Uso	Apoya a las decisiones estratégicas de la organización.	Ayuda a las decisiones tácticas del negocio.
Tiempo de implementación	Puede extenderse de meses a años.	El proceso de implementación es de unos cuantos meses.
Objetivo	Su objetivo principal es proporcionar un ambiente integrado y coherente del negocio.	Es utilizado en la división de negocios a nivel departamental.

- **Objetivos de un almacén de datos:**

Según Kimball en 2013, los objetivos de un almacén de datos pueden ser obtenidos rápidamente al caminar por los pasillos de una organización y escuchar lo que dicen los gerentes del negocio:

- “Recolectamos muchísima información, pero no podemos acceder a ella”
- “Los gerentes necesitan obtener los datos de forma sencilla”
- “Solo muéstrame lo que es importante”
- “Duramos toda la reunión discutiendo quien tiene los números correctos, en lugar de tomar decisiones”

Estas preocupaciones son tan universales que forman la motivación para los requerimientos de un sistema de almacén de datos e Inteligencia de Negocio, por lo que se pueden transformar en los siguientes objetivos de un almacén de datos:

- **Hacer la información fácilmente accesible:** el contenido de un sistema de almacén de datos debe ser entendible. Los datos deben ser intuitivos y obvios para el usuario del negocio, no solamente para el desarrollador. Las estructuras de datos y sus nombres deben ser un reflejo del proceso de pensamiento y vocabulario del usuario. Las herramientas de Inteligencia de Negocio y aplicaciones que acceden a los datos deben ser simples y fáciles de utilizar, también deben retornar los resultados de las consultas con tiempos de espera mínimos. Este objetivo se puede resumir con la frase “**simple y rápido**”.
- **Presentar la información de forma consistente:** los datos de un almacén de datos deben ser creíbles. Los datos deben ser formados cuidadosamente de una gran variedad de fuentes, limpiados, asegurarse de su calidad y mostrados solo cuando son aptos para el consumo de los usuarios. Consistencia también implica, que si dos elementos del almacén de datos tienen el mismo nombre, entonces deben de significar lo mismo. Alternativamente, si no significan lo mismo, deben ser nombrados de forma diferente.
- **Adaptable al cambio:** las necesidades del usuario, condiciones del negocio, datos y tecnologías están siempre sujetos a cambiar. El almacén de datos debe estar diseñado para poder lidiar con estos cambios de tal manera que no invalide los datos existentes o aplicaciones, y ellos no deben ser cambiados o alterados cuando el negocio realiza nuevas preguntas o se agregan nuevos datos al almacén de datos.
- **Presentar la información de manera oportuna:** mientras un sistema de almacén de datos es usado de forma más intensa para la toma de decisiones, se puede necesitar que los datos crudos sean transformados en información utilizable para el negocio, en solo horas, minutos o incluso segundos. El equipo encargado del almacén de datos y los usuarios del negocio deben de tener expectativas realistas del tiempo necesario para realizar esto.

- **Proteger la información:** toda la información importante de una organización esta almacenada en el almacén de datos, y que potencialmente podrían ser dañino si estuviera en las manos equivocadas. El sistema de almacén de datos debe controlar de forma efectiva el acceso a la información confidencial de la organización.
- **Debe proveer confianza en su información para la toma de decisiones:** el almacén de datos debe poseer los datos correctos para apoyar a la toma de decisiones. Las salidas más importantes de un sistema de almacén de datos son las decisiones que se hacen basadas en la evidencia analítica presentada.
- **Debe ser aceptado por sus usuarios:** no importa que tan elegante, bien construida y utilizando las mejores prácticas es la solución de Inteligencia de Negocio creada, si los usuarios de la organización no aceptan el ambiente de Inteligencia de Negocio y lo utilizan activamente. Los usuarios del sistema lo utilizaran si es una fuente rápida y sencilla de información utilizable.

- **Características de los almacenes de datos:**

En la Figura 27 se pueden ver las características principales de un almacén de datos, propuestas por Inmon en 1998.

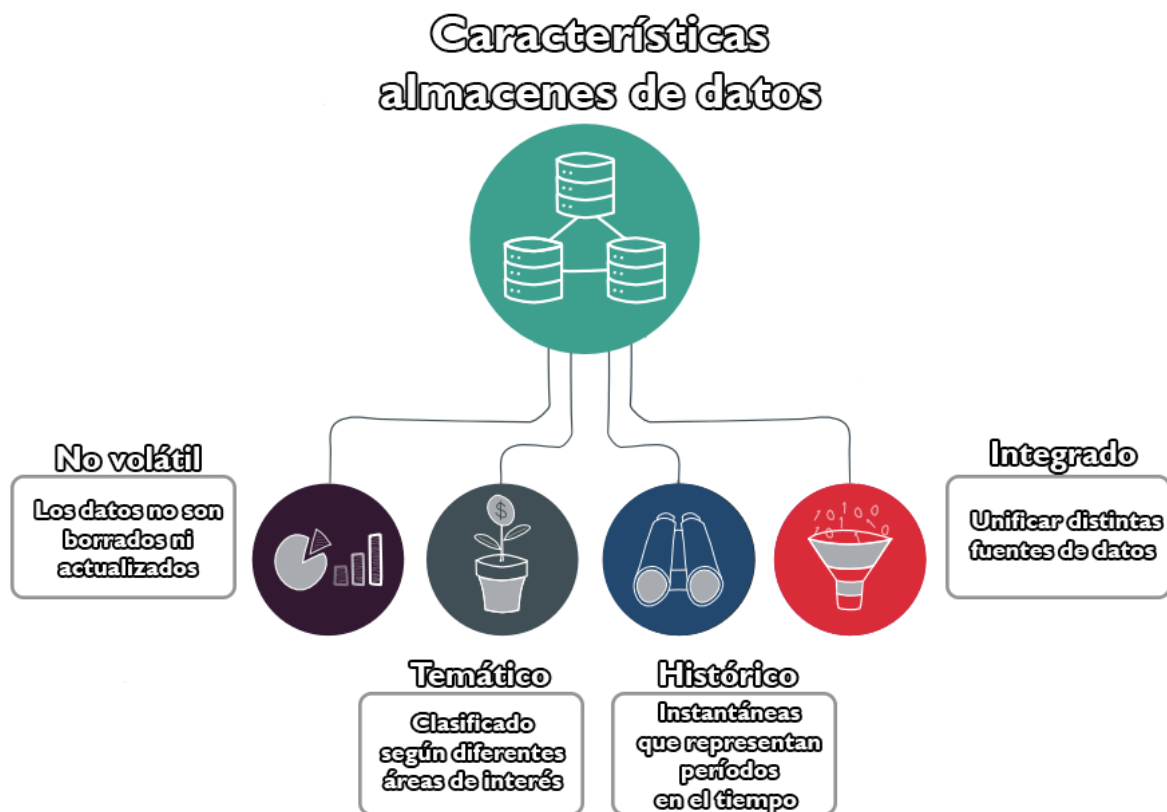


Figura 27. Características de los almacenes de datos.

- **No volátil:** la información de un almacén de datos existe para ser leída, pero no modificada. La información es por tanto permanente, significando la actualización del almacén de datos la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.
- **Temático:** sólo 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.
- **Histórico:** el tiempo es parte implícita de la información contenida en un almacén de datos. 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 almacén de datos sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el almacén de datos se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
- **Integrado:** los datos en un almacén de datos 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.

- **Modelamiento dimensional:**

Un modelo dimensional es una técnica de estructura de datos optimizada para las herramientas de almacenes de datos. Está diseñado para leer, sumarizar y analizar información numérica como valores, balances, conteos, pesos, y demás, que se encuentra en un almacén de datos. En contraste, los modelos relacionales están optimizados para inserción, actualización y borrado de los datos en un sistema OLTP.

A continuación, se nombrarán cada uno de los elementos que integran el modelo dimensional.

- **Hecho:** los hechos son métricas o medidas provenientes del proceso de negocio. Para un proceso de negocios de ventas, un hecho podría ser el número de ventas trimestrales que han ocurrido. La mayoría de las veces los hechos que se almacenan son números enteros o números reales. En la Figura 26 se muestra un ejemplo de hechos.

- **Dimensión:** las dimensiones proveen el contexto que se encuentra alrededor de un evento del proceso de negocio. En términos simples, ellos dan el “quien”, “que” y “donde” y “cuando” de un hecho. Para el proceso de negocio de ventas, y el hecho de número de ventas trimestrales, las dimensiones serían:
 - Quién -> nombre de los clientes
 - Dónde -> ubicación donde se realizó la venta.
 - Qué -> los nombres de los productos vendidos.
 - Cuando -> la fecha en que se realizó la venta.

En otras palabras, una dimensión es la ventana que permite ver información de los hechos.

- **Atributos:** los atributos son la variedad de características que componen una dimensión. En la dimensión ubicación algunos atributos podrían ser:
 - País
 - Estado
 - Código Postal
- Los atributos son utilizados para buscar, filtrar o clasificar los hechos y las tablas dimensión están compuestas de atributos. En la Figura 29 se puede ver un ejemplo de atributos dentro de una tabla dimensional.
- **Tabla de Hechos:** una tabla de hechos, como se puede ver en la Figura 28, es la tabla principal en un modelo dimensional y está compuesta de:
 - Hechos o medidas
 - Claves foráneas de las tablas dimensionales

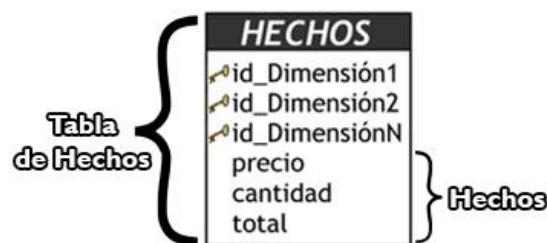


Figura 28. Tabla de Hechos.

- **Tabla Dimensional:** una tabla dimensional contiene las dimensiones con las que se describe un hecho, y están unidas a la tabla de hechos a través de claves foráneas. Este tipo de tablas se encuentran desnormalizadas y cada una de sus columnas representan los atributos de la dimensión. Las tablas dimensionales también pueden poseer uno o varios niveles de jerarquía. En la Figura 29 se puede ver un ejemplo de tabla dimensional.



Figura 29. Tabla Dimensional.

- **Niveles Jerárquicos:** una jerarquía representa una relación entre dos o más atributos pertenecientes a una tabla dimensional, siempre y cuando posean su correspondiente relación de uno a muchos. Por lo general, las jerarquías pueden identificarse fácilmente, debido a que existen relaciones de uno a muchos entre los propios atributos de una dimensión. En la Figura 30 se puede ver una tabla dimensional con niveles jerárquicos.

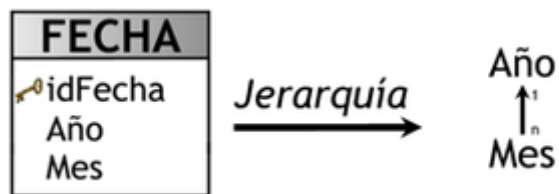


Figura 30. Nivel Jerárquico

La principal ventaja de manejar niveles jerárquicos reside en poder analizar los datos desde su nivel más general al más detallado y viceversa, al desplazarse por los diferentes niveles.

- **Granularidad:** la granularidad describe el nivel de detalle que tendrá la solución empresarial. Es el proceso de identificar el nivel más bajo de información para cualquier tabla en el almacén de datos. Por ejemplo, si una tabla contiene información de las ventas diarias, entonces debería tener una granularidad diaria, en cambio, si una tabla contiene información de las ventas mensuales, entonces tendrá una granularidad mensual.

Con cada uno de estos elementos mencionados, se puede crear un modelo dimensional, tal como se ve en la Figura 31.

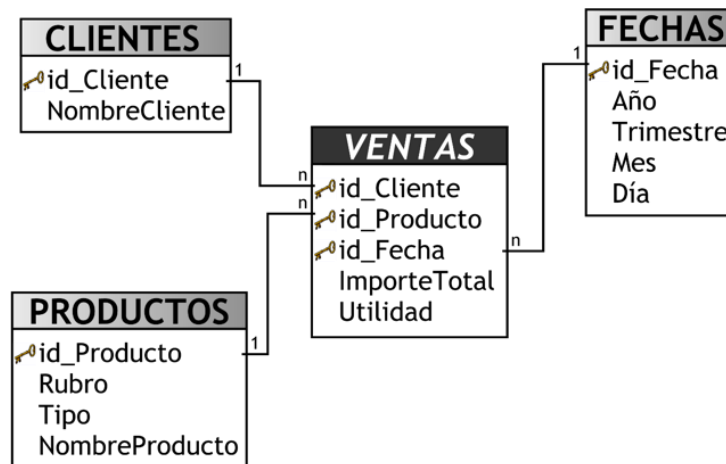


Figura 31. Modelo Dimensional.

2.3.3.5 Herramientas de análisis

Una vez que ya se ha creado el almacén de datos y se encuentra cargado de datos, se deben emplear herramientas de análisis para que el usuario final pueda darle uso a los mismos. Estas herramientas permiten visualizar y personalizar los datos, de una forma fácil e intuitiva para la generación de informes, medidores de desempeño y cuadros de mando. Más adelante se estudia con mayor detalle este componente de una solución de Inteligencia de Negocio.

2.3.4 Beneficios de la Inteligencia de Negocio

Entre los beneficios que provee la Inteligencia de Negocio a una organización, se pueden listar los siguientes, propuestos por Laudon en 2014:

- **Ayuda a aumentar la eficiencia:** muchas organizaciones desperdician buena parte de su tiempo en la búsqueda de información de departamento en departamento. Seguramente así obtendrán datos y podrán hacer sus reportes. Sin embargo, con la Inteligencia de Negocio se puede trabajar la información en una plataforma centralizada, lo que permite ahorrar tiempo y hacer que la toma de decisiones sea más eficiente.
- **Se obtiene respuestas con mayor rapidez:** a diferencia de un sistema con muchos informes separados, la centralización que fomenta la Inteligencia de Negocio permite obtener respuestas rápidas en tan solo minutos. Así, por ejemplo, un informe puede contener en forma centralizada la información referente a las ventas, el desempeño del mercadeo, los costos, los inventarios, los canales de distribución, etc.
- **Permite contar con información precisa:** con la Inteligencia de Negocio se toman decisiones basadas en data objetiva, no en presentimientos o en la intuición. La información adecuada y bien estructurada permite tomar decisiones basadas en el conocimiento producido por la propia empresa.

- **Conocer el comportamiento de los consumidores:** la Inteligencia de Negocio permite encontrar y analizar los hábitos de compra de los consumidores. Esta información es muy útil al punto que se traduce en rentabilidad para la empresa. Además, con ella es posible hacer más eficientes las campañas de fidelización, construir modelos predictivos para facilitar la venta cruzada, la realización de promociones y otras estrategias.
- **Se puede tener un mejor control sobre las áreas funcionales de la empresa:** un sistema de Inteligencia de Negocio puede abarcar áreas como producción, inventario, mercadeo, compras, servicio postventa, pues en todas ellas se usan datos. El campo de la información a obtener y analizar es muy vasto. El hecho de tenerla centralizada para cruzarla, analizarla y tomar decisiones constituye un gran beneficio en costos y en tiempo.

2.4 Indicadores

Luego de haber estudiado lo que es la Inteligencia de Negocio, y la forma en que se pueden almacenar los datos en un almacén de datos para poder obtener información relevante para la organización, debemos estudiar entonces la manera de conocer que es lo que se quiere medir con la Inteligencia de Negocio, es allí entonces donde entran lo que son los indicadores.

2.4.1 Definición

Según Beltrán en 2005, se define un indicador como la relación entre las variables cuantitativas y cualitativas, que permite observar la situación y las tendencias de cambio generadas en el objeto o fenómeno observado con respecto a los objetivos y metas previstas e influencias esperadas.

En otras palabras, un indicador es una característica específica, observable y medible que puede ser usada para mostrar los cambios y progresos que están ocurriendo hacia el logro de un resultado específico. Los indicadores son factores para establecer el logro y el cumplimiento de los objetivos y metas de un determinado proceso. La elección de los indicadores debe realizarse cuidadosamente para controlar las áreas más importantes de la empresa, y poder corregir a tiempo lo que no esté saliendo de acuerdo con lo planificado. Los indicadores pueden ser: medidas, números, hechos, opiniones o percepciones que señalen condiciones o situaciones específicas. Un indicador debe ser definido en términos precisos, no ambiguos, que describan clara y exactamente lo que se está midiendo

2.4.2 Características

Los indicadores poseen ciertos criterios o características que deben cumplir, los cuales se conocen como criterios *S.M.A.R.T* (acrónimo en inglés) que nos dicen que deben de ser específicos, medibles, logrables, relevantes y oportunos (del inglés, *specific, measurable, achievable, relevant, timely*). A continuación, se presenta una tabla que define las propiedades SMART.

Tabla 3. Propiedades S.M.A.R.T. de indicadores, según Maniviesa en 2015.

Propiedad	Definición
Específicos	El indicador se relaciona con los resultados de forma clara y directa. Esta descrito sin ambigüedades y existe un entendimiento general del indicador.
Medibles	El indicador puede ser contado, observado, analizado, probado o retado. Si un indicador no puede ser medible, entonces no se puede determinar si ha ocurrido progreso. Cuando un indicador es claro y específico, pueden ser medidos de muchas formas, permitiendo conocer si los objetivos han sido alcanzados.
Procesables	Los indicadores deben ser alcanzables por el proyecto y, por tanto, deben ser sensibles a los cambios que el proyecto desee hacer.
Relevantes	Los indicadores deben estar acordes con los resultados que se esperan lograr. Es decir, debe existir una relación entre lo que mide el indicador y la razón por la que se quiere medir, ya sea para mejorar un proceso, negocio, etc.
Oportunos	Los indicadores deben ser oportunos en varios aspectos. Primero, se deben recolectar los datos de forma oportuna, luego esos datos deben ser procesados de tal manera que den respuesta en un tiempo adecuado, de tal manera que puedan ser utilizados para ayudar a la toma de decisiones.

2.4.3 Tipos

Los indicadores pueden ser divididos en múltiples clasificaciones, pero según Beltrán en 2005, estos se dividen en dos grandes tipos:

- **Indicadores de eficiencia:** la eficiencia es la relación entre los resultados obtenidos y los recursos utilizados: hacer la mejor manera utilizando la menor cantidad de recursos. Este tipo de indicadores relaciona los recursos invertidos o utilizados con el resultado obtenido.
- **Indicadores de eficacia:** la eficacia es la relación entre los resultados obtenidos y los resultados deseados: hacer de la mejor manera, es decir, lograr los resultados esperados. Este tipo de indicadores miden el grado en que se realizan las actividades planificadas y se logran los resultados esperados. Es una comparación entre logros versus metas.

2.4.4 Beneficios de los indicadores

Los indicadores permiten realizar una valoración de las causas y efectos, así como de sus repercusiones de los aspectos que están ocurriendo en la organización. Podría decirse que el objetivo de los sistemas de medición es aportar a la empresa un camino correcto para que ésta logre cumplir con las metas establecidas. A continuación, se nombran cinco beneficios de los indicadores, propuestos por GantaBI en 2018:

- Proporcionan la información que el gestor necesita sobre cada etapa del proceso.
- Aportan una mayor precisión en la toma de decisiones.
- Proporcionan una mayor eficiencia y eficacia en los procesos.
- Mayor rapidez, mejor comprensión y transparencia a la hora de comunicar resultados.
- Cuadros de mando para una visión “panorámica” de toda la información disponible.

2.5 Herramientas tecnológicas

En este apartado, se detallan las herramientas utilizadas para el desarrollo de la solución de Inteligencia de Negocio, las cuales fueron seleccionadas a partir del proceso de análisis realizado en el trabajo previo de seminario.

2.5.1 Herramientas para el almacenamiento de datos

A continuación, se detalla la herramienta tecnológica seleccionada para brindar apoyo al almacenamiento de datos.

2.5.1.1 PostgreSQL versión 9.4

Según PostgreSQL en 2019, es un sistema de base de datos relacional de código abierto, que utiliza e implementa nuevas funcionalidades al lenguaje SQL para almacenar y escalar los datos con seguridad, por más pesados y complicados que estos sean.

También da soporte a todas las características básicas de un SDBD relacional, y agrega otras como las vistas, procedimientos almacenados, índices y disparadores. Además, puede ser ampliado por sus usuarios, pudiendo modificar características ya existentes, agregar nuevas características y distribuirlas libremente, una de las ventajas de ser de código abierto. Puede ser ejecutado en todos los sistemas operativos y permite las conexiones externas a lenguajes de programación como C/C++, Java, Python, entre otros.

Cuenta con una interfaz gráfica conocida como “pgAdmin”, como se ve en la Figura 32, que permite la gestión de las bases de datos de manera intuitiva y sencilla.

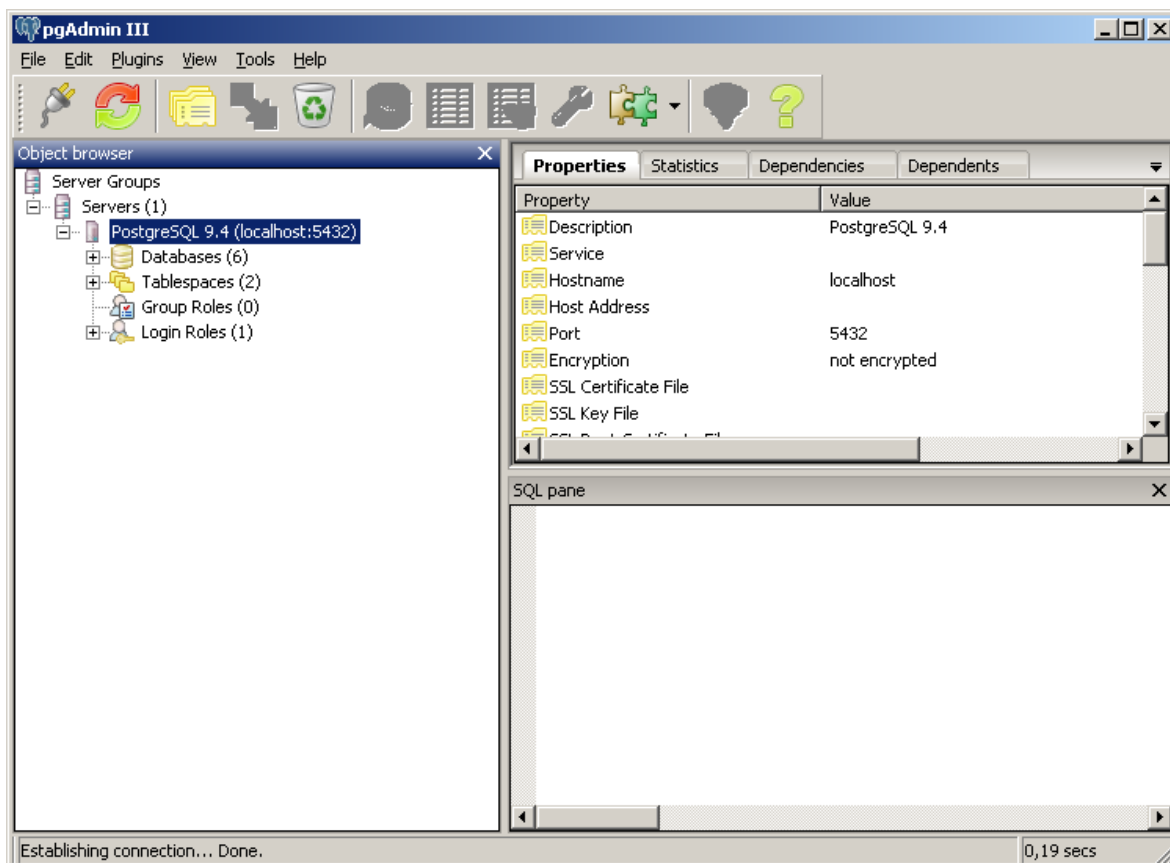


Figura 32. Interfaz gráfica de PostgreSQL.

Al ser una herramienta de código abierto, su sistema de soporte está basado en el apoyo ofrecido por la comunidad, por lo que puede tomar mucho tiempo en dar respuesta al problema que se presente. Además, ofrece actualizaciones de versiones cada 5 años.

En un nivel técnico, proporciona una cantidad limitada de transacciones por segundo en comparación a otros sistemas manejadores de bases de datos, por lo que no es recomendable de utilizar en sistemas donde se necesite de esta característica.

2.5.2 Herramientas para los procesos ETL

A continuación, se detalla la herramienta tecnológica que apoyan a la creación de procesos ETL seleccionada.

2.5.2.1 Pentaho Data Integration (Kettle) versión 8.2

Según Hitachi Ventara en 2019, Pentaho es una suite de Inteligencia de Negocio de código abierto que posee un producto llamado Pentaho Data Integration (o Kettle) utilizado para la integración de datos y procesos ETL. En la Figura 33 se puede ver un ejemplo de su interfaz.

Utiliza un enfoque innovador y una interfaz muy poderosa y fácil de utilizar. Además, está integrado con un motor java que se encarga de procesar las tareas y el movimiento de los datos entre las diferentes bases de datos y archivos. Es capaz de programar tareas y utilizar consultas SQL personalizables, código de Javascript y expresiones regulares.

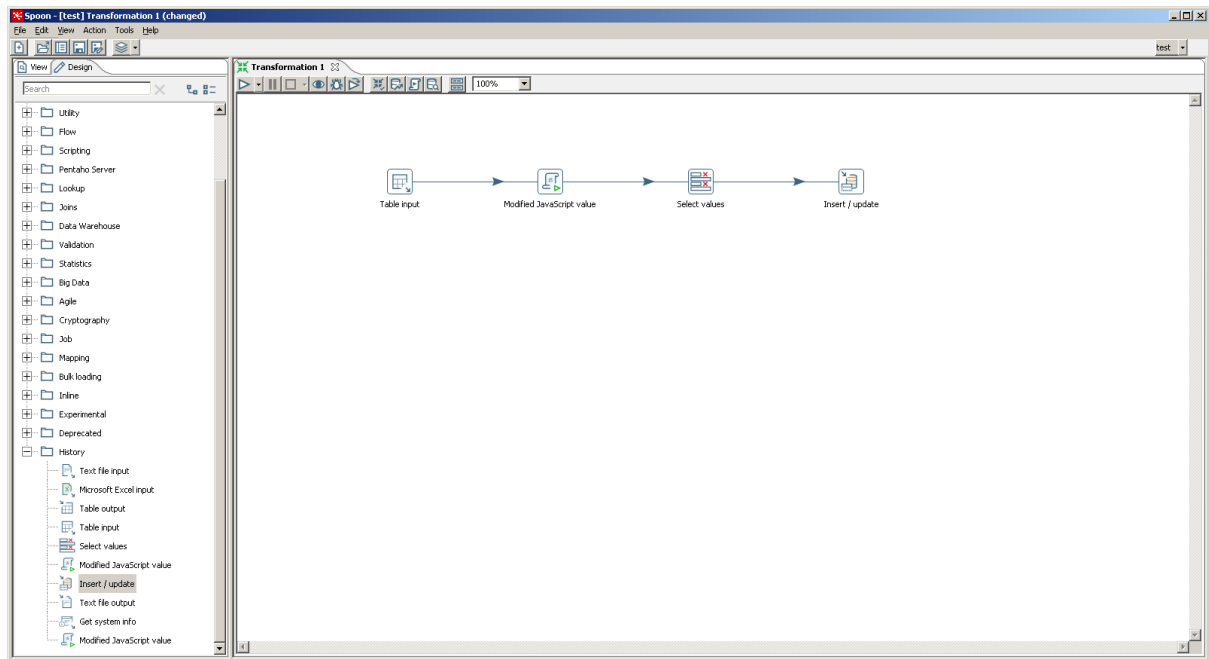


Figura 33. Interfaz de Pentaho Data Integration.

Pentaho Data Integration presenta varios usos entre los cuales encontramos:

- Migración de datos entre diferentes bases de datos y aplicaciones. Incluye soporte para la integración de Big Data, Hadoop, NoSQL.
- Escalabilidad y rendimiento, incluido el almacenamiento en memoria caché. Manejo de errores y cargas masivas que va desde transformaciones simples a otras complejas.
- Carga de almacenes de datos con soporte integrado para las dimensiones lentamente cambiantes y creación de claves subrogadas (conocidas como las claves primarias de las dimensiones).

2.5.3 Herramientas de visualización

A continuación, se detalla la herramienta tecnológica seleccionada para la visualización de datos.

2.5.3.1 Tableau versión 2019.2

Según Tableau en 2019, Tableau es una herramienta de visualización de datos utilizada en la industria de la Inteligencia de Negocio. Ayuda a simplificar los datos crudos y transformarlos a un formato muy fácil de entender. En la Figura 34 se puede ver la interfaz de Tableau.

El análisis de datos es muy rápido con Tableau, y las visualizaciones creadas pueden estar en la forma de hojas de trabajo o cuadros de mando. Los datos que son creados utilizando

Tableau pueden ser entendidos por un profesional en cualquier nivel de una organización. Incluso permite a los usuarios generales la creación de cuadros de mando personalizables, sin necesidad de tener conocimientos tecnológicos. Es por esto que la herramienta ha despertado el interés entre las personas de todos los sectores, desde investigadores, negocios, y cualquier industria que desee poder hacer análisis de sus datos.

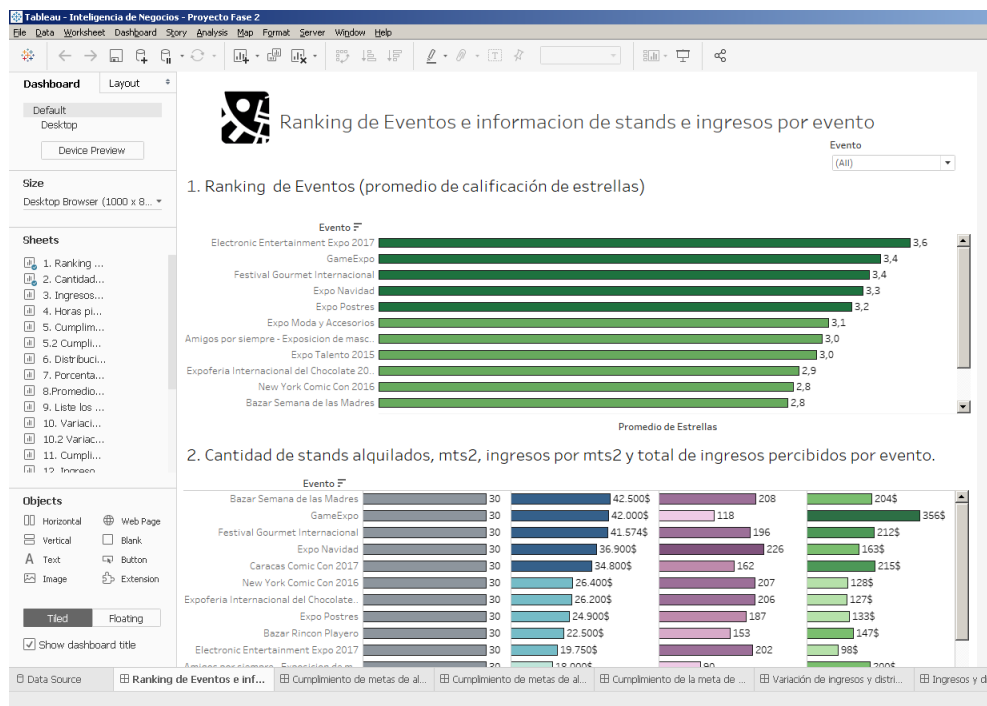


Figura 34. Interfaz de Tableau.

Tableau conecta y extrae los datos almacenados en varios lugares. Puede extraer datos de cualquier plataforma imaginable. Tableau puede extraer una base de datos simple, como Excel, PDF, a una base de datos compleja como Oracle, una base de datos en la nube como los servicios web de Amazon, la base de datos Microsoft Azure SQL, Google Cloud SQL y otras fuentes de datos. Cuando se inicia Tableau, hay disponibles conectores de datos listos que le permiten conectarse a cualquier base de datos.

Los datos extraídos pueden estar conectados en vivo o extraídos al motor de datos de Tableau. Aquí es donde se trabajan con los datos que se extrajeron y se desarrollan las visualizaciones. Los cuadros de mando creados se comparten con los usuarios como un archivo estático.

Los datos pueden publicarse en el servidor de Tableau. Esta es una plataforma empresarial donde se admiten las funciones de colaboración, distribución, gobierno, modelo de seguridad y automatización. Con el servidor de Tableau, los usuarios finales tienen una mejor experiencia en el acceso a los archivos desde todas las ubicaciones, ya sea un escritorio, un dispositivo móvil o un correo electrónico.

CAPÍTULO 3: MARCO METODOLÓGICO

En la actualidad, existen dos metodologías que son consideradas las más adecuadas para llevar a cabo el desarrollo de una solución de Inteligencia de Negocio. Para el desarrollo de este TEG, se decide utilizar el Ciclo de vida dimensional del negocio, propuesto por Ralph Kimball (Kimball y Margy, 2013), en el siguiente capítulo se hace un estudio de dicha metodología, detallando en qué consiste el conjunto de actividades a seguir y cómo están relacionadas para cumplir con los requerimientos y necesidades de la organización.

3.1 Ciclo de vida dimensional del negocio

El ciclo de vida dimensional del negocio proporciona una perspectiva global, incluyendo especificaciones técnicas y de gestión a la hora de desarrollar sistemas complejos de almacenes de datos. Rivadera en 2010 menciona que es una metodología iterativa en donde en cada iteración se genera un conjunto coherente de datos y estructuras. El ciclo de vida dimensional está basado en cuatro principios básicos:

- **Centrarse en el negocio:** hay que concentrarse en la identificación de los requerimientos del negocio y su valor asociado, y usar estos esfuerzos para desarrollar relaciones sólidas con el negocio, agudizando el análisis del mismo y la competencia consultiva de los implementadores.
- **Construir una infraestructura de información adecuada:** diseñar una base de información única, integrada, fácil de usar, de alto rendimiento donde se reflejará la amplia gama de requerimientos de negocio identificados en la empresa.
- **Realizar entregas en incrementos significativos:** crear el almacén de datos en incrementos entregables en plazos de 6 a 12 meses. Hay que usar el valor de negocio de cada elemento identificado para determinar el orden de aplicación de los incrementos.
- **Ofrecer la solución completa:** consiste en proporcionar todos los elementos necesarios para entregar valor a los usuarios de negocios. Para comenzar, esto significa tener un almacén de datos sólido, bien diseñado, con calidad probada, y accesible. También se deberá entregar herramientas de consultas a la medida, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

En la Figura 35 se puede apreciar el ciclo de vida dimensional y cada uno de sus componentes.

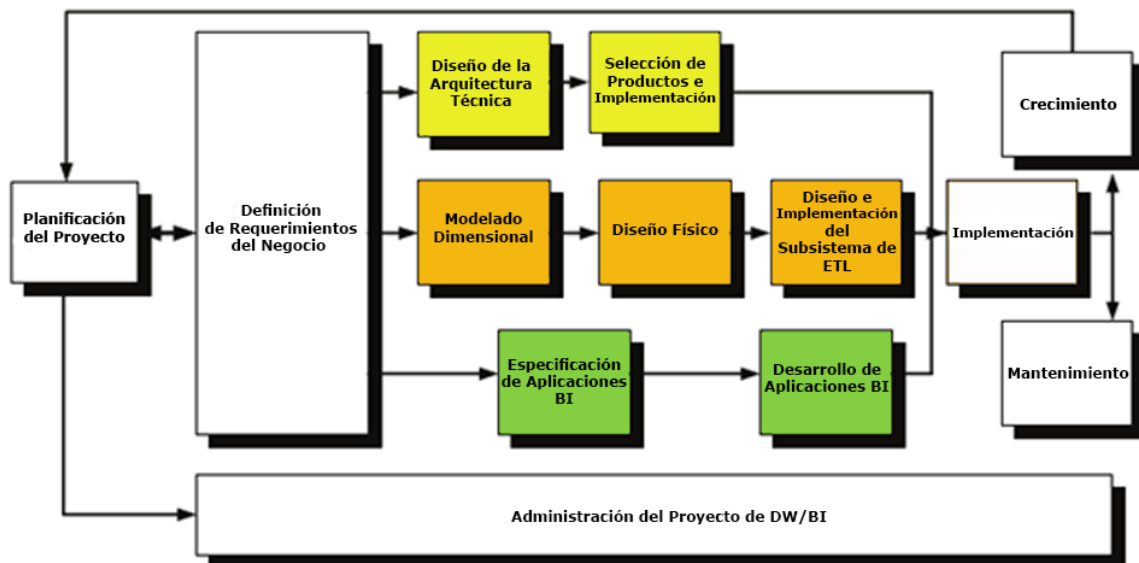


Figura 35. Ciclo de vida dimensional.

En la Figura 35, se pueden observar dos detalles importantes. Primero, hay que resaltar el rol central de la actividad de definición de requerimientos. Los requerimientos del negocio son el soporte inicial de las actividades subsiguientes. También tiene influencia en la planificación del proyecto, notando la doble flecha entre la caja de definición de requerimientos del negocio y la de planificación del proyecto. En segundo lugar, podemos ver tres flujos o caminos que se enfocan en tres diferentes aspectos:

- **Tecnología (flujo superior en color amarillo):** Implica actividades relacionadas con software específico, por ejemplo, PostgreSQL.
- **Datos (flujo del medio en color naranja):** En la que se diseña e implementa el modelo dimensional, y se desarrollan el subsistema de ETL para el llenado del almacén de datos.
- **Aplicaciones de Inteligencia de Negocio (flujo inferior en color verde):** En este flujo se encuentran actividades en las que se diseñan y desarrollan las aplicaciones de negocios para los usuarios finales.

Estos flujos se combinan cuando se instala finalmente el sistema. En la parte de debajo de la figura se muestra la actividad general de administración del proyecto. A continuación, se describirán cada una de las actividades que conforman el ciclo de vida dimensional.

3.1.1 Planificación del proyecto

La planificación busca identificar la definición y el alcance del proyecto de Inteligencia de Negocio, incluyendo justificaciones del negocio y evaluaciones de factibilidad. La planificación del proyecto se focaliza sobre recursos, perfiles, tareas, duraciones y secuencialidad. El plan de proyecto resultante identifica todas las actividades asociadas con el ciclo de vida dimensional e identifica las partes involucradas.

3.1.2 Definición de requerimientos del negocio

Los usuarios finales y sus requerimientos impactan siempre en las implementaciones realizadas de un almacén de datos. Según la perspectiva de Kimball, los requerimientos del negocio se posicionan en el centro del “universo del almacén de datos”. Los requerimientos del negocio deben determinar el alcance del almacén de datos (qué datos debe contener, cómo debe estar organizado, cada cuánto debe actualizarse, quiénes y desde dónde accederán, etc.).

Kimball da consejos y técnicas para descubrir eficazmente los requerimientos del negocio. Estas tácticas y estrategias se focalizan sobre las entrevistas. Parte del proceso de preparación es averiguar a quién se debe realmente entrevistar. Esto normalmente implica examinar cuidadosamente el organigrama de la organización. Hay básicamente cuatro grupos de personas con las que se debe hablar desde el principio: el directivo responsable de tomar las decisiones estratégicas; los administradores intermedios y de negocio responsables de explorar alternativas estratégicas y aplicar decisiones; personal de sistemas, si existen, la gente que realmente sabe qué tipos de problemas informáticos y de datos existen; y por último, la gente que se necesita entrevistar por razones políticas.

3.1.3 Flujo de tecnología

Este flujo o camino consta de las siguientes actividades:

- Diseño de la arquitectura técnica

Según Kimball, los ambientes de almacenes de datos requieren la integración de numerosas tecnologías. Se debe tener en cuenta tres factores: los requerimientos del negocio, los actuales ambientes técnicos y las directrices técnicas estratégicas futuras planificadas. En el apartado 2.3.3 del presente trabajo se indaga más al respecto de la arquitectura de una solución de Inteligencia de Negocio.

- Selección de productos e implementación

Siguiendo con las actividades planteadas por Kimball al utilizar su plan de arquitectura técnica, se deben evaluar y seleccionar componentes tecnológicos específicos como la plataforma de hardware, el sistema de administración de bases de datos, las herramientas de ETL y la herramienta de consulta para el acceso a datos. Una vez que los productos han sido seleccionados, se instalan y prueban para garantizar una integración adecuada de extremo a extremo dentro del almacén de datos.

3.1.4 Flujo de datos

Este flujo o camino consta de las siguientes actividades:

- Modelado dimensional

En esta actividad, se deben realizar las tareas para la creación de un modelo dimensional que sea capaz de dar respuesta a los requerimientos del negocio recopilados en pasos anteriores. Estas actividades consisten en: elección del proceso de negocio, establecer los niveles de granularidad, elección de las dimensiones e identificación de las medidas y las tablas de hecho. En el capítulo 2.3.3.4 del presente trabajo se detalla más a fondo a cerca del modelado dimensional.

- Diseño físico

El diseño físico de la base de datos se focaliza sobre la selección de las estructuras necesarias para soportar el diseño lógico. Algunos de los elementos principales de este proceso son la definición de convenciones estándares de nombres y configuraciones específicos del ambiente de base de datos.

Hay problemas adicionales que abordar, como las preliminares estrategias de ajuste de rendimiento, desde indexación hasta particiones y agregaciones. Si corresponde, las bases de datos OLAP también se diseñan durante este proceso.

- Diseño e implementación del subsistema de ETL

Según Kimball, el sistema de ETL es la base sobre la cual se alimenta el almacén de datos. Si el sistema ETL se diseña adecuadamente, puede extraer los datos de los sistemas de origen de datos, aplicar diferentes reglas para aumentar la calidad y consistencia de los mismos, consolidar la información proveniente de distintos sistemas, y finalmente cargar la información en el almacén de datos en un formato acorde para la utilización por parte de las herramientas de análisis. En el capítulo 2.3.3.2 del presente trabajo se detalla sobre los procesos ETL.

3.1.5 Flujo de aplicaciones de Inteligencia de Negocio

Este flujo o camino consta de las siguientes actividades:

- Especificación de aplicaciones BI

Una parte fundamental de todo proyecto de Inteligencia de Negocio está en proporcionarles a una gran comunidad de usuarios una forma más estructurada y, por lo tanto, más fácil, de acceder al almacén de datos. Este acceso es proporcionado a través de lo que se conoce como aplicaciones de Inteligencia de Negocio.

Las aplicaciones de BI son la cara visible de la Inteligencia de Negocio: los informes y aplicaciones de análisis proporcionan información útil a los usuarios. Las aplicaciones de BI incluyen un amplio espectro de tipos de informes y herramientas de análisis, que van desde informes simples de formato fijo a sofisticadas aplicaciones analíticas que usan complejos algoritmos e información del dominio. Kimball divide a estas aplicaciones en dos categorías basadas en el nivel de sofisticación, y les llama informes estándar y aplicaciones analíticas.

- Desarrollo de aplicaciones BI

Después de la especificación de la aplicación de Inteligencia de Negocio, las actividades de desarrollo de la aplicación incluyen la configuración de los metadatos comerciales y la infraestructura de la herramienta, y luego la construcción y validación de las aplicaciones analíticas y operativas BI especificadas, junto con el portal de navegación.

3.1.6 Implementación

La implementación representa la convergencia de los tres flujos: la tecnología, los datos y las aplicaciones de usuarios finales accesible desde el escritorio del usuario del negocio.

Hay varios factores extras que aseguran el correcto funcionamiento de todas estas piezas, entre ellos se encuentran la capacitación, el soporte técnico, la comunicación, las estrategias de retroalimentación.

3.1.7 Mantenimiento y crecimiento

El desarrollo del ciclo de vida dimensional es un proceso (de etapas bien definidas, con comienzo y fin, pero de naturaleza espiral) pues acompaña a la evolución de la organización durante toda su historia. Se necesita continuar con los relevamientos de forma constante para poder seguir la evolución de las metas por conseguir. Según afirma Kimball, “si se ha utilizado el ciclo de vida dimensional, el almacén de datos está preparado para evolucionar y crecer”. Al contrario de los sistemas tradicionales, los cambios en el desarrollo deben ser vistos como signos de éxito y no de falla.

3.1.8 Administración del proyecto

La administración del proyecto asegura que las actividades del ciclo de vida dimensional se lleven en forma y sincronizadas. Como lo indica el diagrama, la administración o gerencia del proyecto acompaña todo el ciclo de vida. Entre sus actividades principales se encuentra el monitoreo del estado del proyecto y la comunicación entre los requerimientos del negocio y las restricciones de información para poder manejar correctamente las expectativas en ambos sentidos.

CAPÍTULO 4: MARCO APLICATIVO

En este capítulo se detallan cada una de las actividades llevadas a cabo durante la implementación de la solución de Inteligencia de Negocio que satisfacen los requerimientos del caso de estudio, teniendo correspondencia con las actividades que forman parte de la metodología seleccionada.

4.1 Planificación del proyecto

En todo proyecto de software, es necesario realizar una planificación del mismo para garantizar que el desarrollo se realice según las actividades previstas.

En la siguiente tabla se define con detalle cada una de las actividades que se desarrollan durante el proyecto:

Tabla 4. Planificación del proyecto.

Actividad	Duración
Definición de los requerimientos	2 semanas
Diseño de la arquitectura técnica	2 semanas
Selección de productos e implementación	1 semana
Modelado dimensional	2 semanas
Diseño físico	2 semana
Diseño e implementación de subsistema de ETL	3 semanas
Especificaciones y desarrollo de aplicaciones BI	2 semanas
Implementación	2 semanas
Total	16 semanas

La mayor parte del tiempo, como suele ser común en las soluciones de Inteligencia de Negocio, se destina al diseño e implementación del subsistema de ETL, en donde se realiza la transformación e integración adecuada de los datos que son almacenados en el área de preparación y almacén de datos.

La segunda actividad que requiere más tiempo es la de especificación y desarrollo de aplicaciones BI, esto se debe a que se cuentan con un gran número de indicadores que deben ser mostrados en la solución de Inteligencia de Negocio y el desarrollo de cada una de las visualizaciones asociados a ellos necesitara un mayor nivel de atención.

4.2 Definición de requerimientos del negocio

Para el desarrollo de la solución de Inteligencia de Negocio, el INE provee una serie de datos que dan respuesta, tanto de forma directa como indirecta, a los indicadores de la Agenda 2030 correspondientes a los ODS 1, 3, 4 y 8. En la Figura 36 se puede ver una porción de estos datos.

ODS	ESTATUS	INDICADOR ODS	INDICADOR ODS-INE	Periodicidad	UNIDAD DE MEDIDA	1989	1990	1991	1992	1993	
1.2.1		Proporción de la población que vive por debajo del umbral nacional de la pobreza.	Hogares Pobres (NBI)	Anual	Porcentaje						
			Hogares Pobres No Extremos (NBI)	Anual	Porcentaje						
			Hogares Pobres Extremos (NBI)	Anual	Porcentaje						
			Niños 7 a 12 años que no asisten a la escuela	Anual	Porcentaje						
			Hacinamiento crítico	Anual	Porcentaje						
			Viviendas inadecuadas	Anual	Porcentaje						
			Sin servicios básicos	Anual	Porcentaje						
			Alta dependencia económica	Anual	Porcentaje						
			Personas en hogares con pobreza extrema (Por Ingreso)	Anual	Porcentaje			24,0	23,8	19,6	24,0
			Coefficiente Gini	Anual	Valor						
1.3.1		Proporción de la población cubierta por sistemas o niveles mínimos de protección social, desglosada por sexo, distinguiendo entre los niños, los desempleados, los ancianos, las personas con discapacidad, las mujeres embarazadas, los recién nacidos, las víctimas de accidentes de trabajo, los pobres y los vulnerables	Hombres 60 años y más	Anual	Personas		579.525	603.273	627.020	660.767	
			Mujeres 55 años y más	Anual	Personas		904.530	942.636	980.751	1.018.866	
			Total Adultos Mayores	Anual	Personas		1.484.055	1.545.909	1.607.771	1.669.633	
			Beneficiarios (pensionados y pensionadas)	Anual	Personas		234.300	265.891	289.224	305.984	
			Beneficiarios (pensionados y pensionadas)	Anual	Porcentaje		15,8	17,2	18,0	18,3	
			Personas con Discapacidad Certificadas	Anual	Personas						
		Total de Hogares (Equipamiento del Hogar)	2do Semestre	Hogares							

Figura 36. Formato de archivo fuente.

Entre todos los datos suministrados, se puede dar respuesta a 28 indicadores de la Agenda 2030, teniendo 123 indicadores suministrados por el INE, siendo de estos 32 indicadores principales y 91 indicadores complementarios, distribuidos entre los ODS a trabajar. En la siguiente tabla se tiene un resumen de dicha distribución, mientras que en el anexo A se puede ver con detalle, junto con la clasificación de cada indicador por su tipo, ya sea de eficacia o de eficiencia.

Tabla 5. Resumen de contabilización de Indicadores.

ODS	Indicadores Agenda 2030	Cantidad Indicadores Principales	Cantidad Indicadores Complementarios
ODS 1 – Fin de la Pobreza	4	5	38
ODS 3 – Salud y Bienestar	12	16	13
ODS 4 – Educación de Calidad	7	5	20
ODS 8 – Trabajo decente y crecimiento económico	5	6	20
TOTAL	29	32	91

En los archivos que fueron proporcionados por el INE, además de los valores correspondientes a cada medida, se suministran datos como la periodicidad de la medición (anual o semestral), la fuente de donde se originan los datos y la unidad de medida en la que se expresan dichos datos. Se tiene como requisito que cada una de estas características debe ser mostrada en la visualización de los indicadores.

También se tiene como requerimiento que la solución de Inteligencia de Negocio permita la inclusión de nuevos datos, según se vayan obteniendo, para cada indicador ya existente. Es decir, se requiere que sea escalable en el tiempo para cada uno de los indicadores creados. La creación de nuevos indicadores adicionales a los contemplados, se encuentra fuera de los requerimientos y alcance de la solución propuesta.

En cuanto a los usuarios finales de la solución, estos son cualquier persona que desee conocer sobre el estado actual de Venezuela hacia el cumplimiento de la Agenda 2030, ya que los datos mostrados serán publicados por el INE para su libre uso y acceso.

4.3 Flujo de tecnología

El flujo de tecnología hace referencia al grupo de actividades del Ciclo de Vida Dimensional que se encargan de los aspectos tecnológicos de la solución de Inteligencia de Negocio, a continuación, se detallan las tareas realizadas en cada una de estas actividades.

4.3.1 Diseño de la arquitectura técnica

La arquitectura de la solución de Inteligencia de Negocio es aquella detallada en el capítulo 1.3, la cual cuenta con 2 ambientes principales, el ambiente transaccional y el ambiente de Inteligencia de Negocio.

4.3.2 Selección de productos e implementación

Los productos seleccionados para el desarrollo de la solución de Inteligencia de Negocio se detallan en el capítulo 2.5, y la justificación de su elección se puede ver en el capítulo 1.3. En la siguiente tabla se tiene un resumen de los productos seleccionados, junto a las versiones específicas que fueron empleadas.

Tabla 6. Resumen de productos seleccionados.

Tipo	Herramienta	Versión
Sistema manejador de base de datos	PostgreSQL	9.4.22
Procesos ETL	Pentaho Data Integration	8.2
Visualizaciones	Tableau	2019.2 Desktop con licencia de estudiante.

4.4 Flujo de datos

El flujo de datos hace referencia al grupo de actividades del Ciclo de Vida Dimensional que se encargan de los aspectos referentes al manejo de los datos y sus estructuras, a continuación, se detallan las tareas realizadas en cada una de estas actividades.

4.4.1 Modelado de base de datos

Como la solución de Inteligencia de Negocio propuesta cuenta con dos ambientes principales, el ambiente transaccional y el ambiente de Inteligencia de Negocio, se requiere hacer un modelado de las bases de datos donde son almacenados los datos dentro de cada ambiente, a continuación, se detalla cómo es el modelado de ellas.

4.4.1.1 Área de preparación

El área de preparación es aquella donde se pueden almacenar los datos provenientes de las fuentes en Excel, de manera que se puedan integrar y ser accedidos de manera rápida y eficiente. Para el modelado de dicha área, se realizan las siguientes actividades.

- **Preparación de las fuentes**

Antes de poder modelar el área de preparación, se debe hacer un tratamiento previo de los archivos fuente, para tener una estructura uniforme entre ellos y que permita facilitar su carga y garantizar su escalabilidad en el tiempo al momento de nueva carga de datos.

Los archivos fuente tienen la forma que se muestra en la Figura 37, en donde cada columna representa una característica del indicador, como su nombre, periodicidad, unidad de medida, fuente y los años en que se realizaron las mediciones; Mientras que en las filas se encuentra el indicador como tal, obteniendo una fila por cada indicador.

ODS	ESTATUS	INDICADOR ODS	INDICADOR ODS-INE	Periodicidad	UNIDAD DE MEDIDA	FUENTE	1989	1990	1991
		3.3.2 Incidencia de la tuberculosis por cada 100.000 habitantes	Tasa de mortalidad en tuberculosis	Anual	Por 100.000 Habitantes	MPPPS		4,01	4,02
			Tasa de prevalencia de tuberculosis	Anual	Por 100.000 Habitantes	MPPPS		27,7	25,8

Figura 37. Archivo Excel sin tratamiento.

Esta estructura de datos no es conveniente, ya que, si se desean agregar nuevos años de medición a un indicador, es necesario crear una nueva columna de datos correspondiente a ese año en particular, lo que implica una modificación a la estructura de los programas de ETL cada vez que se agregue una nueva columna. Es por esto por lo que se requiere realizar una transposición de las filas y las columnas, de manera que los años de los datos se mantengan todos en una sola columna, mientras que los valores de los indicadores estén distribuidos cada uno en una columna propia, de manera que los datos de nuevas series de tiempo puedan ser agregados sin necesidad de realizar modificaciones técnicas. Para lograr esto se hace lo siguiente:

1. Se inicia con uno de los archivos Excel fuente (Figura 37), para un indicador específico, es necesario hacer una nueva hoja de Excel por cada uno de los indicadores ODS.
2. Transponer las filas y las columnas de manera que exista una única columna para los años, como se observa en la Figura 38. Se agrega una columna correspondiente al código del país y al número del indicador al que dan respuesta los datos.

A	B	C	D	E	F
Código País	Indicador	Año	Tasa de mortalidad materna	FUENTE	UNIDAD DE MEDIDA
01	3.1.1	1990	58,9	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1991	53,1	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1992	53,3	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1993	63,1	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1994	69,9	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1995	66,3	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1996	60,4	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1997	59,6	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1998	51,0	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	1999	59,3	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	2000	60,1	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	2001	67,2	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	2002	68,0	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	2003	57,8	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	2004	59,9	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	2005	59,9	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	2006	60,5	MPPPS	Por 100.000 Nacidos Vivos
01	3.1.1	2007	56,1	MPPPS	Por 100.000 Nacidos Vivos
			63,4	MPPPS	Por 100.000 Nacidos

Figura 40. Archivo Excel transpuesto correspondiente al indicador 3.1.1 con fuente MPPPS.

Ya que se tiene una estructura adecuada de los datos fuentes, se puede realizar el modelado del área de preparación, el cual se detalla a continuación.

- **Modelado del área de preparación**

Como el área de preparación es un repositorio simple en donde son almacenados los datos fuente, no existen relaciones en ella, es por esto por lo que su modelado es sencillo, y se crea una tabla por cada uno de los indicadores de la Agenda 2030 de los que se posee datos, haciendo que se tenga entonces un total de 29 tablas en el área de preparación. En la siguiente Tabla (Tabla 7) se puede ver la correspondencia entre los archivos fuente y el área de preparación de algunos indicadores, mientras que en el anexo B se puede ver la correspondencia de toda el área de preparación.

Tabla 7. Correspondencia de algunos archivos fuente con área de preparación.

Fuentes	Tablas
INDICADORES AGENDA 2030 ODS v7: Esperanza de Vida Mujeres Esperanza de Vida Hombres Tasa de mortalidad materna (PRINCIPAL)	SA_DASH_INDICADOR_311
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad de niños menores de 5 años (PRINCIPAL) Proporción de niños y niñas menores de 5 años con peso inferior al normal Esperanza de Vida Nacional Desnutrición en menores de 5 años según corte internacional	SA_DASH_INDICADOR_321
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad neonatal (PRINCIPAL)	SA_DASH_INDICADOR_322
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad por enfermedad virus de la inmunodeficiencia humana (VIH) (PRINCIPAL)	SA_DASH_INDICADOR_331

4.4.1.2 Modelado dimensional

Una vez identificados los requerimientos principales, establecido los indicadores y haber definido la arquitectura a utilizar, se genera el modelo dimensional conceptual con el que se creara finalmente el almacén de datos. Para generar dicho modelo, se toman en cuenta los siguientes pasos:

- **Determinar el área temática y/o proceso de negocio**

Cada uno de los diferentes ODS a trabajar, puede ser visto como un proceso de negocio y/o área temática. Entonces se tienen las siguientes 4 áreas temáticas:

1. ODS 1: Fin de Pobreza.
2. ODS 3: Salud y Bienestar.
3. ODS 4: Educación de Calidad.
4. ODS 8: Trabajo Decente y Crecimiento Económico.

Cada una de ellas puede tener asociada varias tablas de hechos que están relacionadas a través de las dimensiones.

- **Determinar la Granularidad**

La granularidad es el nivel de detalle asociado a una fila de la tabla de hechos. Para cada tabla de hechos, la granularidad es por año/semestre, y para cada indicador la granularidad está determinada por las dimensiones que se consideran para obtener la medida.

- **Identificación de los Hechos**

Los hechos son aquellas métricas que se desean medir, de tal manera que los valores que son suministrados en las fuentes de datos son todos y cada uno diferentes hechos que deben ser almacenados. Es por esto por lo que se decidió crear una tabla de hechos por cada Indicador de los que se posean datos de la Agenda 2030, llegando entonces a un total de 28 tablas de hechos.

- **Identificación de las Dimensiones**

Una vez identificados los hechos, se debe de determinar las características o perspectivas que se quieren conocer acerca de ellos. Tomando en cuenta la información suministrada en las fuentes de datos y en los requerimientos del negocio se lograron identificar las siguientes dimensiones:

- **Tiempo:** permite conocer el momento de medición de los hechos, puede ser anual o semestral. Contiene la siguiente jerarquía: Año -> Semestre
- **Ubicación:** permite conocer el lugar en donde fue medido el hecho, ya sea a nivel nacional, o más específicamente a nivel parroquial. Contiene una jerarquía especificada de la siguiente manera: País -> Estado -> Municipio -> Parroquia
- **Indicador:** permite conocer a que indicador de la Agenda 2030 está asociado el hecho.
- **Indicador Complementario:** permite conocer a que indicador complementario está asociado el hecho.
- **Unidad de Medida:** los hechos pueden estar representados en diferentes formas y unidades, tales como porcentaje, cantidades, años, entre otros. Esta dimensión permite identificar la unidad de medida asociada al hecho.
- **Fuente:** esta dimensión permite conocer el organismo que se encargó de medir el hecho.

Con los hechos y las dimensiones claramente identificadas, se obtiene el modelo dimensional resultante, que será común entre las 29 tablas de hecho a crear, en la Figura 41 se puede ver dicho modelo.



Figura 41. Modelo Dimensional Conceptual.

4.4.2 Diseño físico

En esta fase se trasladan el área de preparación y los modelos dimensionales conceptuales construidos a un diseño físico, es decir, se implementa físicamente las bodegas de datos que soportan la solución de Inteligencia de Negocio. Se deben tomar en cuenta algunos aspectos importantes como son las convenciones de nombre, tipos de datos y otros aspectos referentes al ambiente de bases de datos.

4.4.2.1 Diseño físico del área de preparación

Para el diseño físico del área de preparación se crea un esquema de base de datos llamado **“ods_intermedia”** en el que son almacenadas todas las tablas correspondientes al área de preparación.

Cada una de las tablas fue nombrada siguiendo una convención de nombre, que consiste en llamarlas **“SA_DASH_INDICADOR_XXX”**, siendo SA el acrónimo de área de preparación en inglés (*staging área*), dash como diminutivo de panel de control, del inglés *dashboard*, y finalmente INDICADOR_XXX, haciendo referencia al número del indicador de la Agenda 2030 cuyos datos están almacenados en esa tabla.

A continuación, en las Tabla 8 y 9, se muestra la estructura física de la tabla SA_DASH_INDICADOR_332 y SA_DASH_INDICADOR_461, correspondientes a los indicadores 3.3.2 y 4.6.1 de la Agenda 2030. En el Anexo C se especifica la estructura física de cada una de las tablas del área de preparación.

Tabla 8. Estructura de la tabla SA_DASH_INDICADOR_332

Nombre de la Tabla: SA_DASH_INDICADOR_332			
Descripción: Tabla que contiene la información del indicador 3.3.2 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	Text	Identificador del país al que están asociados los datos.
cod_indicador	variable	Text	Identificador del indicador ODS.
cod_anio	variable	Text	Identificador del año en que se obtuvieron los datos.
tasa_mortalidad_tuberculosis	18,2	double precision	Valor de la tasa de mortalidad en tuberculosis.
tasa_prevalencia_tuberculosis	18,2	double precision	Valor de la tasa de prevalencia de la tuberculosis.
fuelle	variable	Text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	Text	Identificador de la unidad de medida en que se miden los valores.
Relaciones:		Campos Clave:	

Tabla 9. Estructura de la tabla SA_DASH_INDICADOR_461

Nombre de la Tabla: SA_DASH_INDICADOR_461			
Descripción: Tabla que contiene la información del indicador 4.6.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
tasa_pob_analfabeta	18,2	double precision	Valor de la tasa de población analfabeta.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
Relaciones:		Campos Clave:	

En la Figura 42 y 43 se presentan algunos de los datos almacenados dentro de las SA de los indicadores 3.3.2 y 4.6.1.

	cod_pais text	cod_indicador text	cod_anio double precision	tasa_mortalidad_tuberculosis double precision	tasa_prevalencia_tuberculosis double precision	fuentes text	cod_unidad_medida text
1	01	3.3.2	1990	4.01	27.7	MPPPS	Por 100.000 Habitantes
2	01	3.3.2	1991	4.02	25.8	MPPPS	Por 100.000 Habitantes
3	01	3.3.2	1992	4.09	26.4	MPPPS	Por 100.000 Habitantes
4	01	3.3.2	1993	3.48	24.5	MPPPS	Por 100.000 Habitantes
5	01	3.3.2	1994	3.73	22.6	MPPPS	Por 100.000 Habitantes
6	01	3.3.2	1995	3.5	25.3	MPPPS	Por 100.000 Habitantes
7	01	3.3.2	1996	3.32	25.1	MPPPS	Por 100.000 Habitantes
8	01	3.3.2	1997	3.21	26.1	MPPPS	Por 100.000 Habitantes
9	01	3.3.2	1998	2.95	26.8	MPPPS	Por 100.000 Habitantes
10	01	3.3.2	1999	2.84	27.6	MPPPS	Por 100.000 Habitantes
11	01	3.3.2	2000	2.58	26.6	MPPPS	Por 100.000 Habitantes
12	01	3.3.2	2001	2.96	25.2	MPPPS	Por 100.000 Habitantes
13	01	3.3.2	2002	2.61	24.6	MPPPS	Por 100.000 Habitantes
14	01	3.3.2	2003	2.8	26.2	MPPPS	Por 100.000 Habitantes
15	01	3.3.2	2004	2.64	26.1	MPPPS	Por 100.000 Habitantes
16	01	3.3.2	2005	2.35	25.8	MPPPS	Por 100.000 Habitantes
17	01	3.3.2	2006	2.3	24.8	MPPPS	Por 100.000 Habitantes
18	01	3.3.2	2007	2.43	23.5	MPPPS	Por 100.000 Habitantes
19	01	3.3.2	2008	2.3	23	MPPPS	Por 100.000 Habitantes
20	01	3.3.2	2009	1.94	23.6	MPPPS	Por 100.000 Habitantes
21	01	3.3.2	2010	2.25	22.8	MPPPS	Por 100.000 Habitantes
22	01	3.3.2	2011	1.93	22.6	MPPPS	Por 100.000 Habitantes
23	01	3.3.2	2012		22.84	MPPPS	Por 100.000 Habitantes
24	01	3.3.2	2013		23.08	MPPPS	Por 100.000 Habitantes

Figura 42. Datos de la tabla SA_DASH_INDICADOR_332

	cod_pais text	cod_indicador text	tasa_pob_analfabeta double precision	fuentes text	cod_anio double precision	cod_unidad_medida text
1	01	4.6.1	5.1	INE	1999	Porcentaje
2	01	4.6.1	5	INE	2000	Porcentaje
3	01	4.6.1	4.9	INE	2001	Porcentaje
4	01	4.6.1	5.2	INE	2002	Porcentaje
5	01	4.6.1	4.9	INE	2003	Porcentaje
6	01	4.6.1	4.3	INE	2004	Porcentaje
7	01	4.6.1	4.1	INE	2005	Porcentaje
8	01	4.6.1	3.9	INE	2006	Porcentaje
9	01	4.6.1	3.4	INE	2007	Porcentaje
10	01	4.6.1	3.3	INE	2008	Porcentaje
11	01	4.6.1	3.2	INE	2009	Porcentaje
12	01	4.6.1	3.2	INE	2010	Porcentaje
13	01	4.6.1	2.8	INE	2011	Porcentaje
14	01	4.6.1	2.5	INE	2012	Porcentaje
15	01	4.6.1	2.5	INE	2013	Porcentaje
16	01	4.6.1	2.4	INE	2014	Porcentaje
17	01	4.6.1	2.1	INE	2015	Porcentaje
18	01	4.6.1	1.8	INE	2016	Porcentaje
19	01	4.6.1	1.8	INE	2017	Porcentaje
20	01	4.6.1	1.8	INE	2018	Porcentaje

Figura 43. Datos de la tabla SA_DASH_INDICADOR_461

4.4.2.2 Diseño físico del modelo dimensional

Para el diseño físico del modelo dimensional se crea un esquema de base de datos llamado “ods_dimensional” en el que son almacenadas todas las tablas pertenecientes al almacén de datos.

Tanto las dimensiones como las tablas de hechos se nombran de manera tal de seguir una convención de nombre. En el caso de las dimensiones estas se nombran con el prefijo “DIM” seguido de lo que representa la dimensión, de esta forma, por ejemplo, la dimensión que almacena la información de las unidades de medida se nombra DIM_UNIDAD_MEDIDA y la dimensión fuente se nombra DIM_FUENTE.

En cuanto a las tablas de hechos, estas se nombran utilizando el formato “FACT_DASH_INDXXX”, siendo FACT la forma de llamar a las tablas de hechos en inglés, DASH como diminutivo de panel de control, del inglés *dashboard*, y IND como abreviación de Indicador, seguido del número del indicador de la Agenda 2030 cuyos datos se encuentran representados en esa tabla de hecho. De esta forma la tabla del indicador 1.2.1 tiene como nombre “FACT_DASH_IND121”

A continuación, en las Tabla 10 y 11, se muestra la estructura física de las dimensiones DIM_FUENTE y DIM_TIEMPO_SEMESTRAL. En el Anexo D se especifica la estructura física de cada una de las dimensiones del almacén de datos.

Tabla 10. Estructura de la dimensión fuente

Nombre de la Tabla: DIM_FUENTE			
Descripción: Dimensión que contiene la información de todas las fuentes encargadas de medir los datos de los indicadores de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
sk_fuente	32	bigint	Identificador único de la dimensión fuente, clave subrogada.
cod_fuente	variable	text	Clave del negocio de la fuente que suministra los datos.
desc_fuente	variable	text	Descripción de la fuente que suministra los datos.
Relaciones:		Campos Clave: sk_fuente	

Tabla 11. Estructura de la dimensión tiempo semestral

Nombre de la Tabla: DIM_TIEMPO_SEMESTRAL			
Descripción: Dimensión que contiene la información de todos los años con sus semestres, desde 1980 hasta el 2050.			
Campo	Tamaño	Tipo de Dato	Descripción
sk_tiempo	32	bigint	Identificador único de la dimensión tiempo semestral, clave subrogada.
anio_cod	32	bigint	Identificador del año.
semestre_cod	variable	text	Identificador en números del semestre del año.
semestre_desc	variable	text	Descripción del semestre del año.
Relaciones:		Campos Clave: sk_tiempo	

En las Figuras 44 y 45 se presentan algunos de los datos almacenados dentro de las dimensiones DIM_UNIDAD_MEDIDA y DIM_TIEMPO_SEMESTRAL.

	cod_unidad_medida text	desc_unidad_medida text	sk_unidad_medida [PK] numeric(18,2)
1			0.00
2	porcentaje	Porcentaje	1.00
3	centimetros	Centímetros	2.00
4	Por 1.000 Nacidos Vivos	Por 1.000 Nacidos Vivos	3.00
5	becas	Becas	4.00
6	valor	Valor	5.00
7	libros	Libros	6.00
8	cien_mil_hab	Por 100.000 Habitantes	7.00
9	consultas	Consultas	8.00
10	anio	Año	9.00
11	cien_mil_nac_vivos	Por 100.000 Nacidos Vivos	10.00
12	establecimientos_salud	Establecimientos de salud	11.00
13	cajeros	Cajeros Automáticos por cada 100.000 adultos	12.00
14	planteles_edu	Planteles Educativos	13.00
15	viviendas	Viviendas	14.00
16	canaima	Portátiles Canaima	15.00
17	exámenes	Exámenes	16.00
18	sucursales_banco	Sucursales Bancarias por cada 100.000 adultos	17.00
19	estudiantes	Estudiantes	18.00
20	personas	Personas	19.00
21	millones_usd	Millones de US \$	20.00
22	vidas_salvadas	Vidas Salvadas	21.00
23	hogares	Hogares	22.00

Figura 44. Datos de la dimensión DIM_UNIDAD_MEDIDA

	anio_cod bigint	semestre_cod text	sk_tiempo [PK] bigint	semestre_desc text
1			0	
2	1980	1	1	1er Semestre
3	1981	1	2	1er Semestre
4	1982	1	3	1er Semestre
5	1983	1	4	1er Semestre
6	1984	1	5	1er Semestre
7	1985	1	6	1er Semestre
8	1986	1	7	1er Semestre
9	1987	1	8	1er Semestre
10	1988	1	9	1er Semestre
11	1989	1	10	1er Semestre
12	1990	1	11	1er Semestre
13	1991	1	12	1er Semestre
14	1992	1	13	1er Semestre
15	1993	1	14	1er Semestre
16	1994	1	15	1er Semestre
17	1995	1	16	1er Semestre
18	1996	1	17	1er Semestre
19	1997	1	18	1er Semestre
20	1998	1	19	1er Semestre
21	1999	1	20	1er Semestre
22	2000	1	21	1er Semestre
23	2001	1	22	1er Semestre
24	2002	1	23	1er Semestre
25	2003	1	24	1er Semestre
26	2004	1	25	1er Semestre
27	2005	1	26	1er Semestre
28	2006	1	27	1er Semestre
29	2007	1	28	1er Semestre

Figura 45. Datos de la dimensión DIM_TIEMPO_SEMESTRAL

En cuanto a las tablas de hechos, las 28 que fueron creadas están relacionadas con cada una de las 6 dimensiones almacenadas, en las tablas 12 y 13 se muestra la estructura física de las tablas de hechos FACT_DASH_IND121 y FACT_DASH_IND333. En el Anexo D se especifica la estructura física de cada una de las tablas de hecho que se encuentran en el almacén de datos.

Tabla 12. Estructura física de la tabla de hechos FACT_DASH_IND121

Nombre de la Tabla: FACT_DASH_IND121			
Descripción: Tabla de hecho que contiene toda la información relacionadas a los indicadores principales y secundarios asociados al indicados ODS 1.2.1			
Campo	Tamaño	Tipo de Dato	Descripción
sk_fact_dash_ind121	32	bigint	Identificador propio de la fact, clave subrogada.
sk_tiempo	32	bigint	Clave primaria de la dimensión tiempo semestral.
sk_complementario	32	bigint	Clave primaria de la dimensión indicador complementario
sk_indicador	32	bigint	Clave primaria de la dimensión indicador
sk_ubicacion	32	bigint	Clave primaria de la dimensión ubicación geo
sk_unidad_medida	32	bigint	Clave primaria de la dimensión unidad de medida
sk_fuente	32	bigint	Clave primaria de la dimensión fuente
valor	18,2	double precision	Medida correspondiente al indicador
cantidad	32	int	Medida cuyo valor siempre es 1, usada para contabilizar la cantidad de registros
Relaciones: DIM_TIEMPO_SEMESTRAL (sk_tiempo) DIM_DASH_IND_COMPLEMENTARIO (sk_complementario) DIM_INDICADOR (sk_indicador) DIM_UBICACION (sk_ubicacion) DIM_UNIDAD_MEDIDA (sk_unidad_medida) DIM_FUENTE (sk_fuente)		Campos Clave: sk_tiempo sk_complementario sk_indicador sk_ubicacion sk_unidad_medida sk_fuente sk_fact_dash_ind121	

Tabla 13. Estructura física de la tabla de hechos FACT_DASH_IND333

Nombre de la Tabla: FACT_DASH_IND333			
Descripción: Tabla de hecho que contiene toda la información relacionadas a los indicadores principales y secundarios asociados al indicados ODS 3.3.3			
Campo	Tamaño	Tipo de Dato	Descripción
sk_fact_dash_ind333	32	bigint	Identificador propio de la fact, clave subrogada.
sk_tiempo	32	bigint	Clave primaria de la dimensión tiempo semestral.
sk_complementario	32	bigint	Clave primaria de la dimensión indicador complementario
sk_indicador	32	bigint	Clave primaria de la dimensión indicador
sk_ubicacion	32	bigint	Clave primaria de la dimensión ubicación geo
sk_unidad_medida	32	bigint	Clave primaria de la dimensión unidad de medida
sk_fuente	32	bigint	Clave primaria de la dimensión fuente
valor	18,2	double precision	Medida correspondiente al indicador
cantidad	32	int	Medida cuyo valor siempre es 1, usada para contabilizar la cantidad de registros
Relaciones: DIM_TIEMPO_SEMESTRAL (sk_tiempo) DIM_DASH_IND_COMPLEMENTARIO (sk_complementario) DIM_INDICADOR (sk_indicador) DIM_UBICACION (sk_ubicacion) DIM_UNIDAD_MEDIDA (sk_unidad_medida) DIM_FUENTE (sk_fuente)		Campos Clave: sk_tiempo sk_complementario sk_indicador sk_ubicacion sk_unidad_medida sk_fuente sk_fact_dash_ind333	

En las Figuras 46 y 47 se presentan algunos de los datos almacenados dentro de las tablas de hechos FACT_DASH_IND_121 y FACT_DASH_IND_333.

	sk_fact_dash_ind121 [PK] bigint	sk_tiempo bigint	sk_complementario bigint	sk_indicador bigint	sk_ubicacion bigint	valor double precision	cantidad integer	sk_unidad_medida numeric(18,2)	sk_fuente integer
1	0	0	0	0	0			0.00	0
2	1	105	5	8	1477	9.48	1	1.00	3
3	2	89	9	8	1477	23.4	1	1.00	3
4	3	94	2	8	1477	18.1811973890061	1	1.00	3
5	4	99	6	8	1477	7.21048	1	1.00	3
6	5	107	8	8	1477	3.0954	1	1.00	3
7	6	95	4	8	1477	1.89418	1	1.00	3
8	7	92	5	8	1477	15.3	1	1.00	3
9	8	94	5	8	1477	16.3592	1	1.00	3
10	9	99	2	8	1477	14.8256555574695	1	1.00	3
11	10	106	1	8	1477	20.4193743386175	1	1.00	3
12	11	95	1	8	1477	30.4953096895693	1	1.00	3
13	12	90	10	8	1477	0.486	1	5.00	3
14	13	104	1	8	1477	21.5340159355343	1	1.00	3
15	14	95	5	8	1477	16.24413	1	1.00	3
16	15	94	7	8	1477	17.14745	1	1.00	3
17	16	102	3	8	1477	7.73295910348853	1	1.00	3
18	17	95	6	8	1477	9.24689	1	1.00	3
19	18	102	10	8	1477	0.39	1	5.00	3
20	19	103	6	8	1477	5.81	1	1.00	3
21	20	107	3	8	1477	4.90592656329336	1	1.00	3
22	21	103	1	8	1477	21.2466641579028	1	1.00	3
23	22	95	8	8	1477	5.5097	1	1.00	3
24	23	104	3	8	1477	6.32401593553429	1	1.00	3
25	24	105	1	8	1477	19.5854460324481	1	1.00	3
26	25	92	3	8	1477	10.4761893288244	1	1.00	3
27	26	98	7	8	1477	8.69144	1	1.00	3
28	27	86	9	8	1477	39	1	1.00	3
29	28	105	10	8	1477	0.409	1	5.00	3

Figura 46. Datos de la tabla de hechos FACT_DASH_IND_121

	sk_fact_dash_ind333 [PK] bigint	sk_tiempo bigint	sk_complementario bigint	sk_indicador bigint	sk_ubicacion bigint	sk_fuente bigint	sk_unidad_medida bigint	valor double precision	cantidad integer
1	0								
2	1	95	64	145	1477	7	7	121.475312114814	1
3	2	100	64	145	1477	7	7	114.692332278619	1
4	3	93	65	145	1477	7	7	0.113067515642184	1
5	4	97	65	145	1477	7	7	0.0639681923055446	1
6	5	84	64	145	1477	7	7	104.193376788548	1
7	6	101	64	145	1477	7	7	126.233532175457	1
8	7	94	64	145	1477	7	7	116.332339569567	1
9	8	90	65	145	1477	7	7	0.264830316168527	1
10	9	106	64	145	1477	7	7	293.006336055092	1
11	10	104	65	145	1477	7	7	0.0340561025930555	1
12	11	99	64	145	1477	7	7	151.917059086017	1
13	12	91	64	145	1477	7	7	90.8118236801687	1
14	13	104	64	145	1477	7	7	174.58520433304	1
15	14	96	64	145	1477	7	7	177.006054583884	1
16	15	103	64	145	1477	7	7	158.330705660026	1
17	16	84	65	145	1477	7	7	0.188791195630617	1
18	17	85	64	145	1477	7	7	59.6736627640227	1
19	18	88	64	145	1477	7	7	97.11769938961	1
20	19	91	65	145	1477	7	7	0.100562163444103	1
21	20	86	64	145	1477	7	7	74.8376783221644	1
22	21	85	65	145	1477	7	7	0.137311451254198	1
23	22	106	65	145	1477	7	7	0.0695269271995133	1
24	23	92	65	145	1477	7	7	0.0987276108917781	1
25	24	97	64	145	1477	7	7	170.561777695631	1
26	25	86	65	145	1477	7	7	0.148276001876618	1
27	26	103	65	145	1477	7	7	0.055283067618724	1
28	27	92	64	145	1477	7	7	122.944671113441	1
29	28	87	64	145	1477	7	7	100.064604641372	1

Figura 47. Datos de la tabla de hechos FACT_DASH_IND_333

Una vez que ya se encuentran creadas físicamente las dimensiones y tablas de hechos, se puede crear el modelo dimensional físico, en la Figura 48 y 49 se pueden ver los modelos dimensionales correspondientes al indicador ODS 1.2.1 y 3.3.3

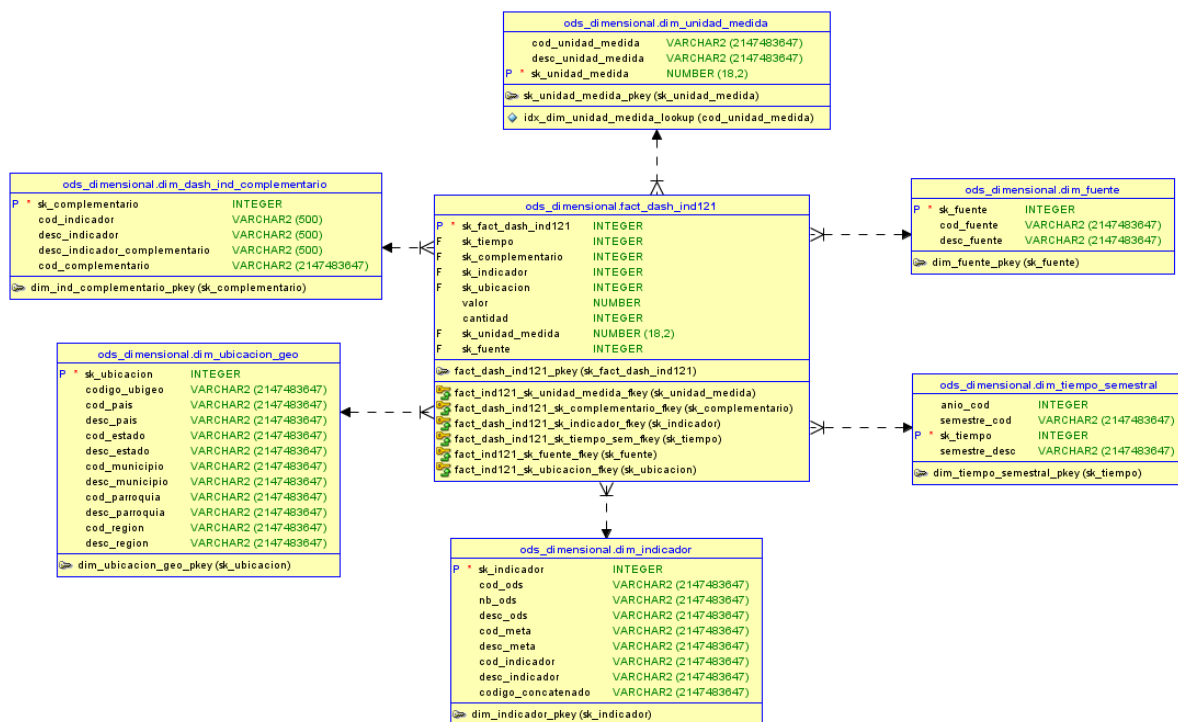


Figura 48. Modelo Dimensional correspondiente al indicador 1.2.1

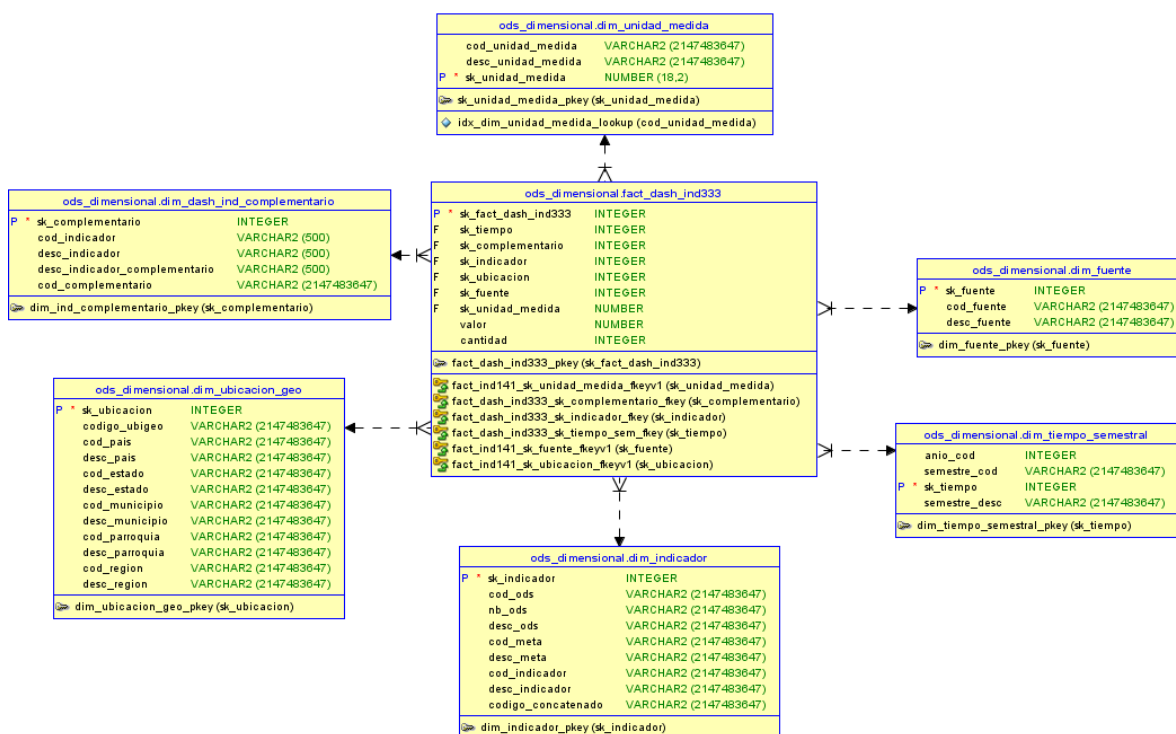


Figura 49. Modelo Dimensional correspondiente al indicador 3.3.3

Finalmente, en la tabla 14, se presenta la correspondencia entre las bodegas de datos creadas y las áreas temáticas.

Tabla 14. Correspondencia de áreas temáticas y bodega de datos creadas

Área Temática / Proceso de Negocio	Bodegas de Datos
ODS 1 - Fin de la Pobreza	FACT_DASHBOARD_IND_121
	FACT_DASHBOARD_IND_122
	FACT_DASHBOARD_IND_131
	FACT_DASHBOARD_IND_141
	FACT_DASHBOARD_IND_1a2
ODS 3 - Salud y Bienestar	FACT_DASHBOARD_IND_311
	FACT_DASHBOARD_IND_321
	FACT_DASHBOARD_IND_322
	FACT_DASHBOARD_IND_331
	FACT_DASHBOARD_IND_332
	FACT_DASHBOARD_IND_333
	FACT_DASHBOARD_IND_341
	FACT_DASHBOARD_IND_342
	FACT_DASHBOARD_IND_361
	FACT_DASHBOARD_IND_371
	FACT_DASHBOARD_IND_381
	FACT_DASHBOARD_IND_3b1
	FACT_DASHBOARD_IND_3b3
ODS 4 - Educación de Calidad	FACT_DASHBOARD_IND_411
	FACT_DASHBOARD_IND_421
	FACT_DASHBOARD_IND_422
	FACT_DASHBOARD_IND_461
	FACT_DASHBOARD_IND_4a1
	FACT_DASHBOARD_IND_4b1
	FACT_DASHBOARD_IND_4c1
ODS 8 - Trabajo Decente y Crecimiento Económico	FACT_DASHBOARD_IND_811
	FACT_DASHBOARD_IND_821
	FACT_DASHBOARD_IND_852
	FACT_DASHBOARD_IND_861
	FACT_DASHBOARD_IND_8101

4.4.3 Diseño e implementación del subsistema de ETL

Una vez que se definen y crean los modelos físicos del área de preparación y del almacén de datos, se procede al diseño e implementación del subsistema de ETL, utilizando la herramienta Pentaho Data Integration. Para ello se construyen transformaciones, las cuales permiten cargar los datos provenientes de las fuentes, a cada una de nuestras tablas de bases de datos, además se crean Jobs que se encargan de unificar dichas transformaciones para facilitar su ejecución. Al igual que las actividades anteriores, se dividen los subsistemas ETL según el esquema de base de datos en que van a actuar.

4.4.3.1 Sistema ETL del área de preparación

Para el desarrollo de los programas ETL del área de preparación se crea un repositorio de Pentaho Data Integration llamado “ODS Intermedia”, en él se encuentran 57 transformaciones y 5 Jobs que se encargan de llenar las tablas pertenecientes al área de preparación. En la Figura 50 y 51 se puede ver parte del contenido de este repositorio.













Name	Type
 ETL_SA_DASH_INDICADOR_1a2	Transformation
 ETL_SA_DASH_INDICADOR_1a2_PARTE_II	Transformation
 ETL_SA_DASH_INDICADOR_121	Transformation
 ETL_SA_DASH_INDICADOR_121_PARTE_II	Transformation
 ETL_SA_DASH_INDICADOR_122	Transformation
 ETL_SA_DASH_INDICADOR_131_F1	Transformation
 ETL_SA_DASH_INDICADOR_131_F2	Transformation
 ETL_SA_DASH_INDICADOR_131_F3	Transformation
 ETL_SA_DASH_INDICADOR_131_F4	Transformation
 ETL_SA_DASH_INDICADOR_141	Transformation
 ETL_SA_DASH_INDICADOR_141_HOGARES	Transformation
 ETL_SA_DASH_INDICADOR_141_PORCENT	Transformation

Figura 50. Estructura del repositorio “ODS Intermedia” para las transformaciones






Name	Type
 Job_final	Job
 job_sa_dash_ods_1	Job
 job_sa_dash_ods_3	Job
 job_sa_dash_ods_4	Job
 job_sa_dash_ods_8	Job

Figura 51. Estructura del repositorio “ODS Intermedia” para los Jobs.

Se realizan Procesos ETL para la carga de cada una de las SA indicadas en el Anexo B en donde se muestra la correspondencia entre los archivos Excel fuente y las SA creadas. A continuación, en la Figura 52, 53 y 54 se muestra el proceso de carga de una SA usando la herramienta Pentaho Data Integration (PDI) Versión 8.2. En este caso se muestra para el Indicador 1.2.1. Este proceso se aplica para cada uno de los indicadores de cada uno de los ODS trabajados.

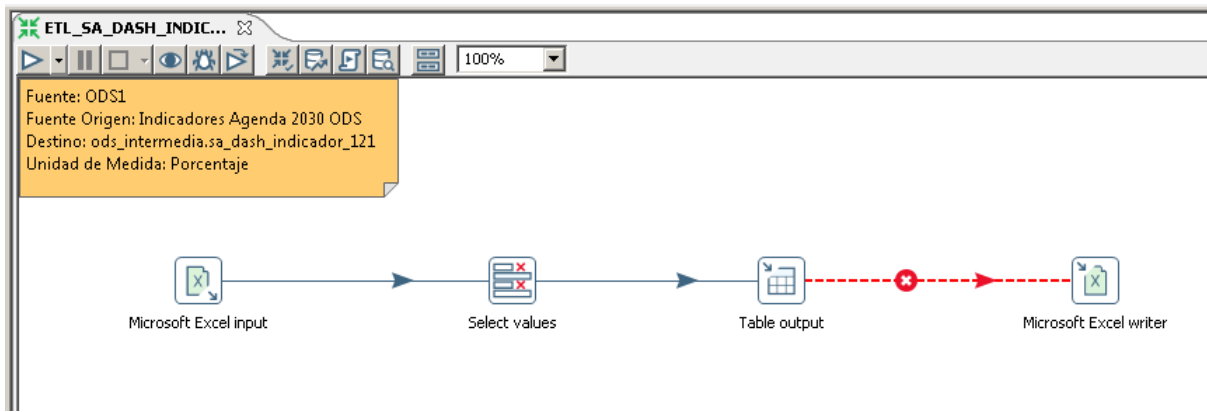


Figura 52. ETL para la carga de la SA del indicador 1.2.1

Name	Type	Length	Precision	Trim type	Repeat	Format	Currency
1 Codigo País	String	-1	-1	none	N		
2 Indicador	String	-1	-1	none	N		
3 Año	Number	-1	-1	none	N		
4 Hogares Pobres (NBI)	Number	-1	-1	none	N		
5 Hogares Pobres No Extremos (NBI)	Number	-1	-1	none	N		
6 Hogares Pobres Extremos (NBI)	Number	-1	-1	none	N		
7 Niños 7 a 12 años que no asisten a la escuela	Number	-1	-1	none	N		
8 Hacinamiento crítico	Number	-1	-1	none	N		
9 Viviendas inadecuadas	Number	-1	-1	none	N		
10 Sin servicios básicos	Number	-1	-1	none	N		
11 Alta dependencia económica	Number	-1	-1	none	N		
12 Personas en hogares con pobreza extrema (Por Ingreso)	Number	-1	-1	none	N		
13 FUENTE	String	-1	-1	none	N		
14 Unidad de Medida	String	-1	-1	none	N		

Figura 53. Campos correspondientes al Excel del indicador 1.2.1

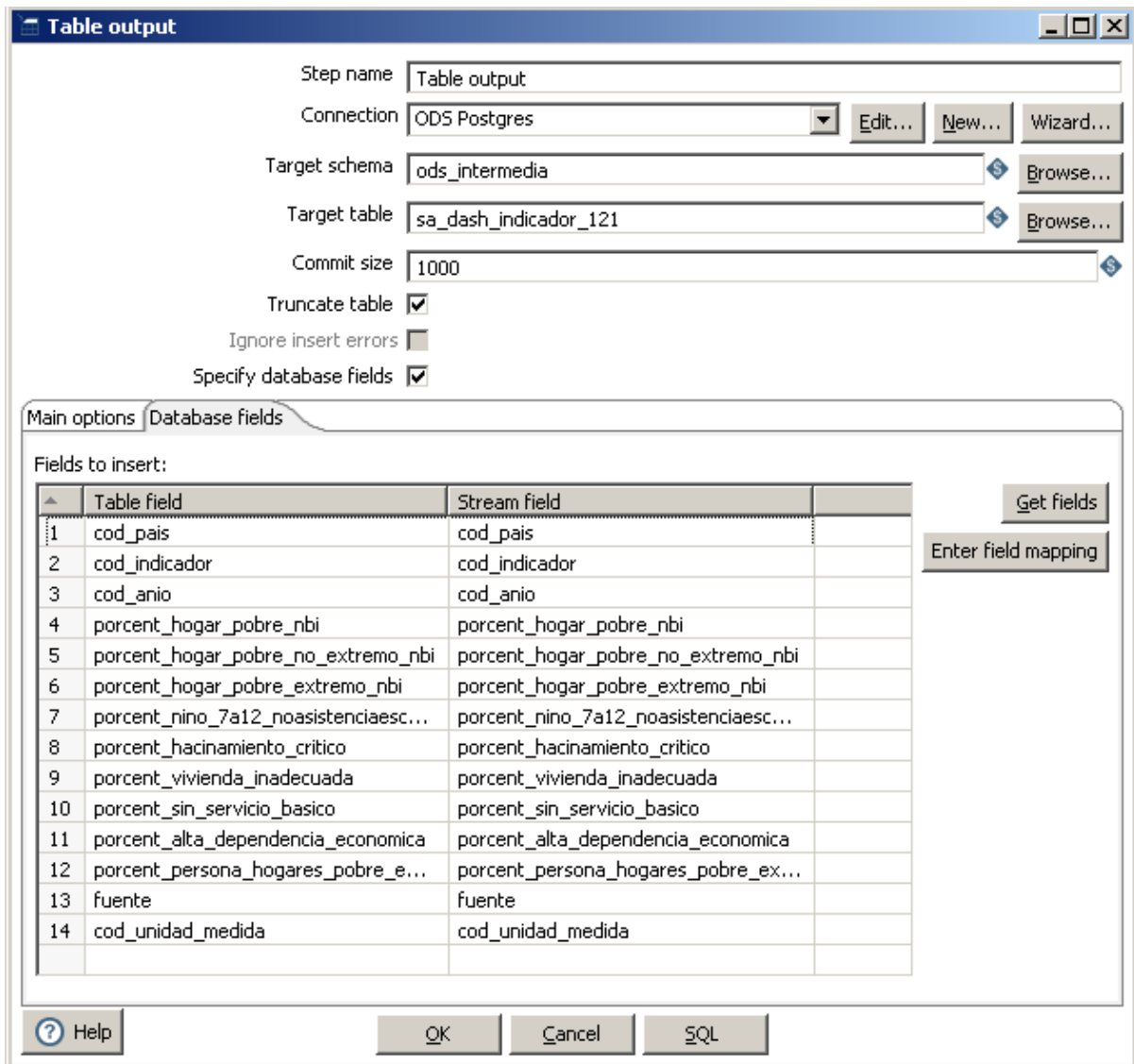


Figura 54. Paso de carga hacia la tabla SA_DASH_INDICADOR_121

Finalmente se crean los Jobs encargados de agrupar las transformaciones creadas, en total se crearon 5 Jobs, siendo 4 que agrupan las transformaciones que cargan las SA por cada ODS, y un último que se encarga de agrupar todos los Jobs creados, implementando un control de alerta de posibles errores. En la Figura 55 se puede ver el Job creado para el ODS 1, mientras que en la Figura 56 se encuentra el Job final encargado de agrupar el resto de los Jobs.

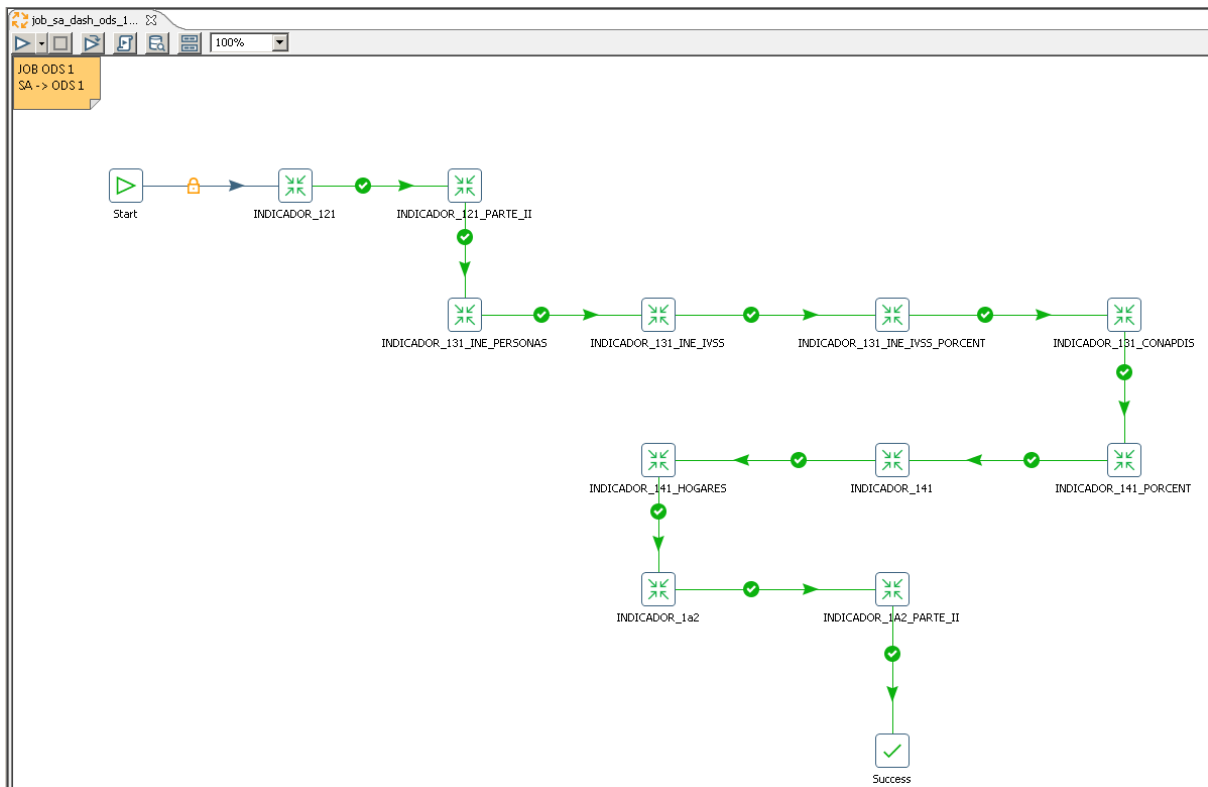


Figura 55. Job para la carga de las SA del ODS 1

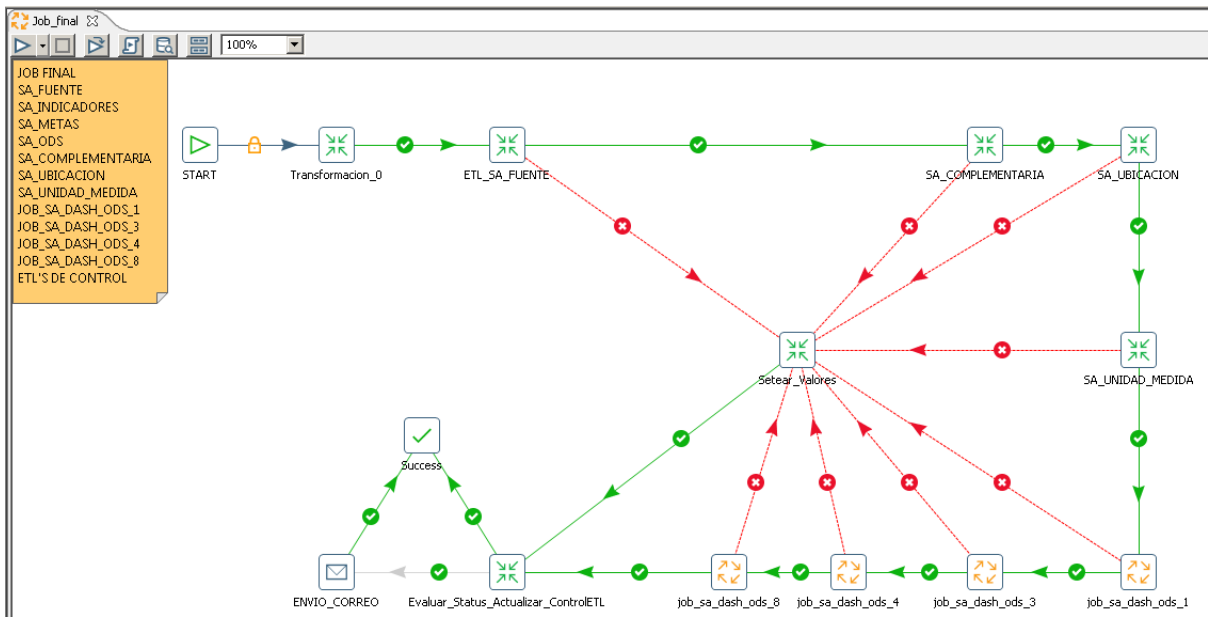


Figura 56. Job final que agrupa el resto de los Jobs y control de errores

4.4.3.2 Sistema ETL del almacén de datos

Para el desarrollo de los programas ETL del almacén de datos, se crea un repositorio de Pentaho Data Integration llamado “ODS Dimensional”, en él se encuentran 39 transformaciones y 5 Jobs que se encargan de llenar las dimensiones y tablas de hechos pertenecientes al almacén de datos. En la Figura 57, 58 y 59 se pueden ver parte del contenido de este repositorio.

Name	Type
✘ DIM_DASH_COMPLEMENTARIO	Transformation
✘ DIM_DASH_FUENTE	Transformation
✘ DIM_INDICADOR	Transformation
✘ DIM_TIEMPO_SEMESTRAL	Transformation
✘ DIM_UBICACION_V3	Transformation
✘ DIM_UNIDAD_MEDIDA	Transformation

Figura 57. Transformaciones para las dimensiones del almacén de datos

Name	Type
✘ FACT_DASH_INDICADOR_1a2	Transformation
✘ FACT_DASH_INDICADOR_3b1	Transformation
✘ FACT_DASH_INDICADOR_3b3	Transformation
✘ FACT_DASH_INDICADOR_4a1	Transformation
✘ FACT_DASH_INDICADOR_4b1	Transformation
✘ FACT_DASH_INDICADOR_4c1	Transformation
✘ FACT_DASH_INDICADOR_121	Transformation
✘ FACT_DASH_INDICADOR_122	Transformation
✘ FACT_DASH_INDICADOR_131	Transformation
✘ FACT_DASH_INDICADOR_141	Transformation
✘ FACT_DASH_INDICADOR_311	Transformation

Figura 58. Transformaciones para las tablas de hechos del almacén de datos

Name	Type
Job_final	Job
Job_ODS1	Job
Job_ODS3	Job
Job_ODS4	Job
Job_ODS8	Job

Figura 59. Jobs que llenan el almacén de datos

Se realizan Procesos ETL para la carga de cada una de las dimensiones y tablas de hechos indicadas en el Anexo C. A continuación, en la Figuras 60, 61, 62 y 63 se muestra el proceso de carga de una dimensión. En este caso se muestra para dimensión DIM_UBICACION_GEO. Este proceso se aplica para cada una de las dimensiones trabajadas.

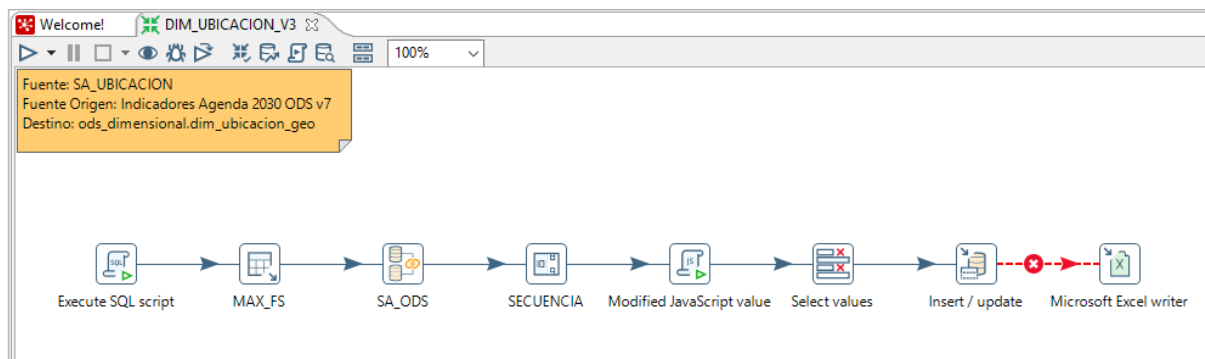


Figura 60. ETL para la carga de la dimensión DIM_UBICACION_GEO

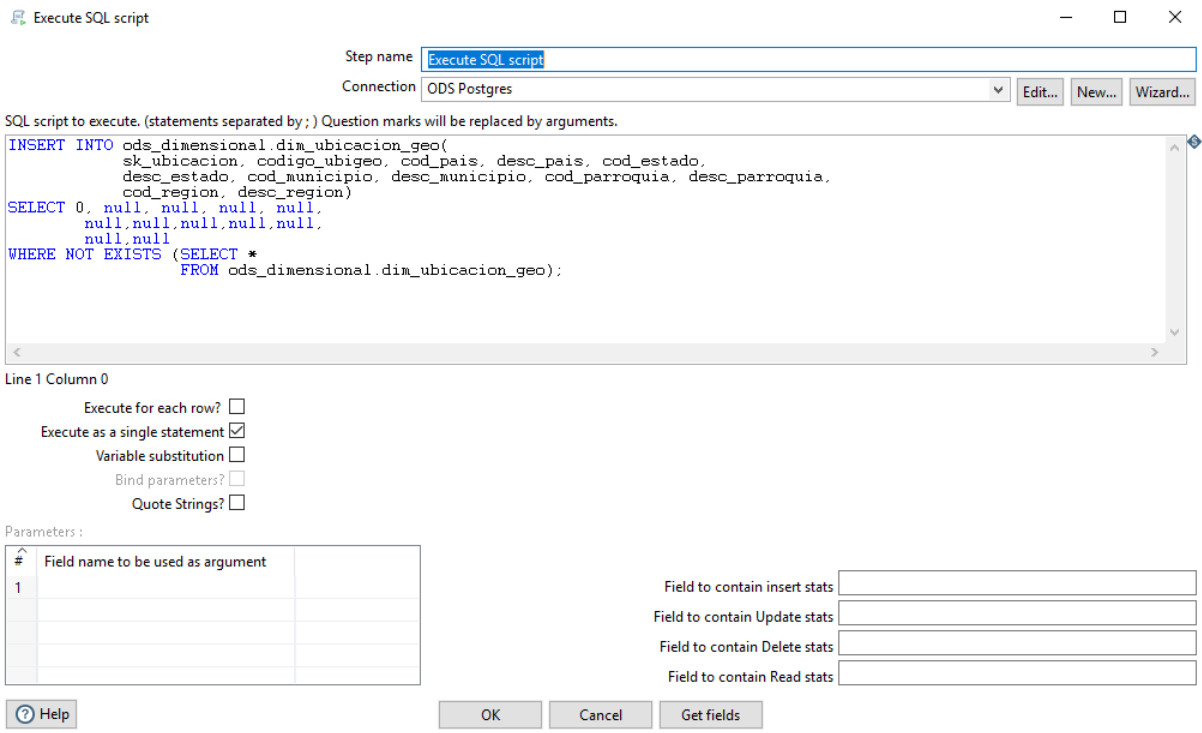


Figura 61. Paso de inserción de SK inicial para la dimensión

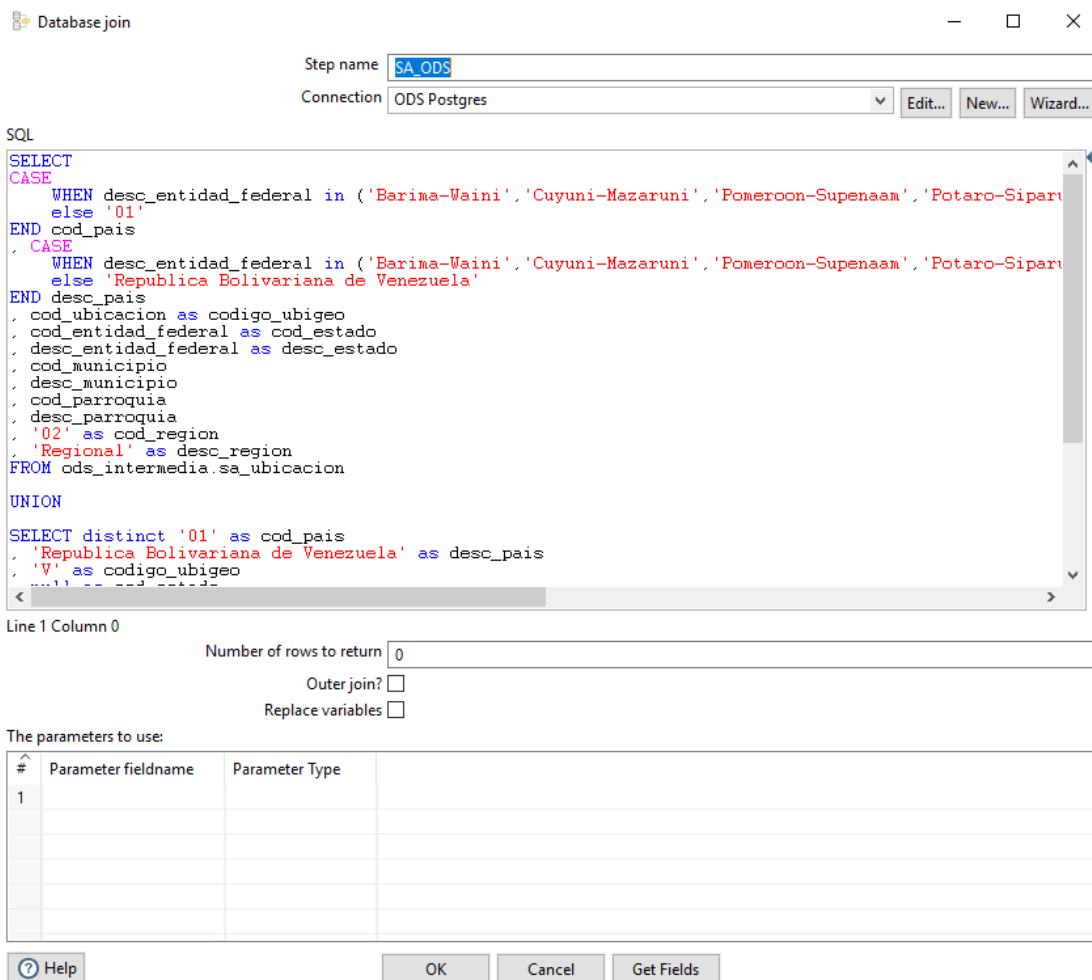


Figura 62. Consulta que obtiene los datos a ser cargados de la SA hacia la dimensión

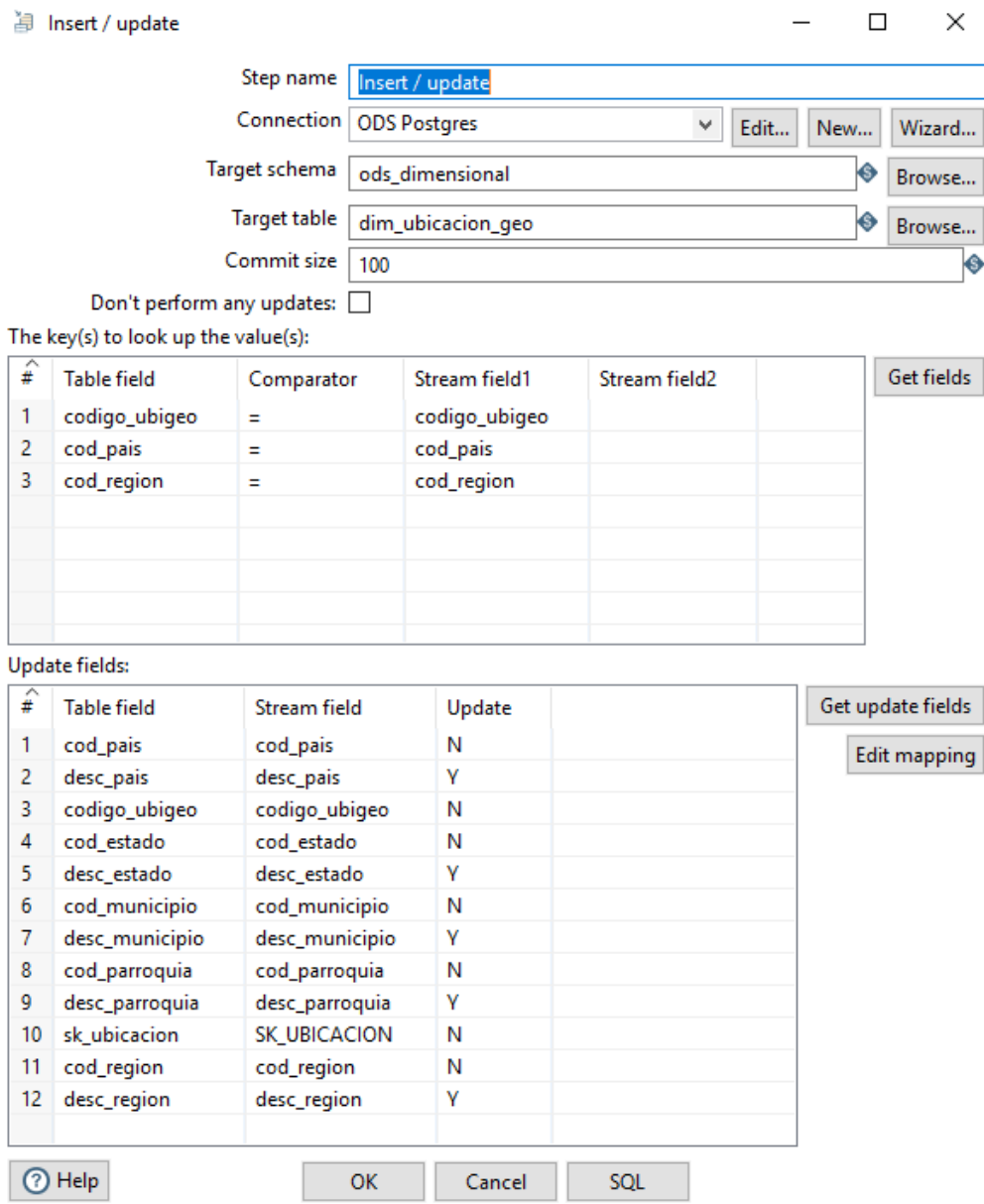


Figura 63. Paso de carga hacia la tabla DIM_UBICACION_GEO

Para la carga de las tablas de hechos, también se sigue un proceso común para cada una de ellas, en las Figuras 64, 65, 66 y 67 se muestra el proceso de carga para la tabla de hecho FACT_DASH_IND_121 correspondiente al indicador 1.2.1

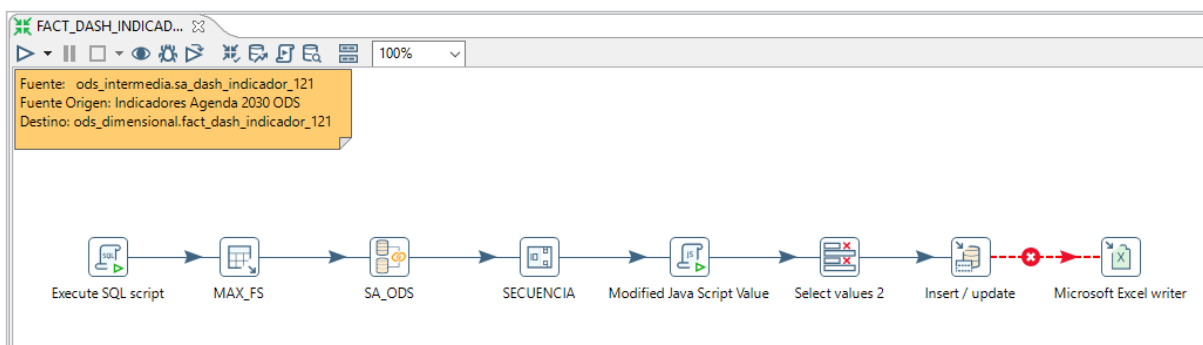


Figura 64. ETL para la carga de la tabla de hechos FACT_DASH_IND_121

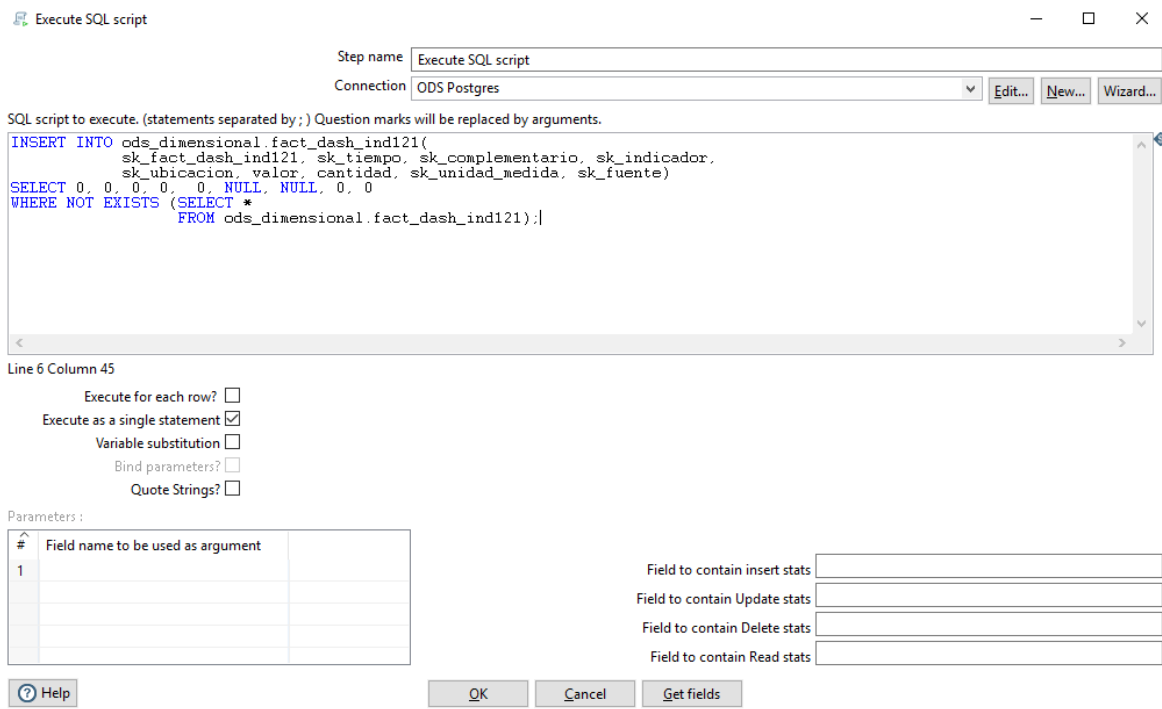


Figura 65. Paso de inserción de SK inicial para la tabla de hechos

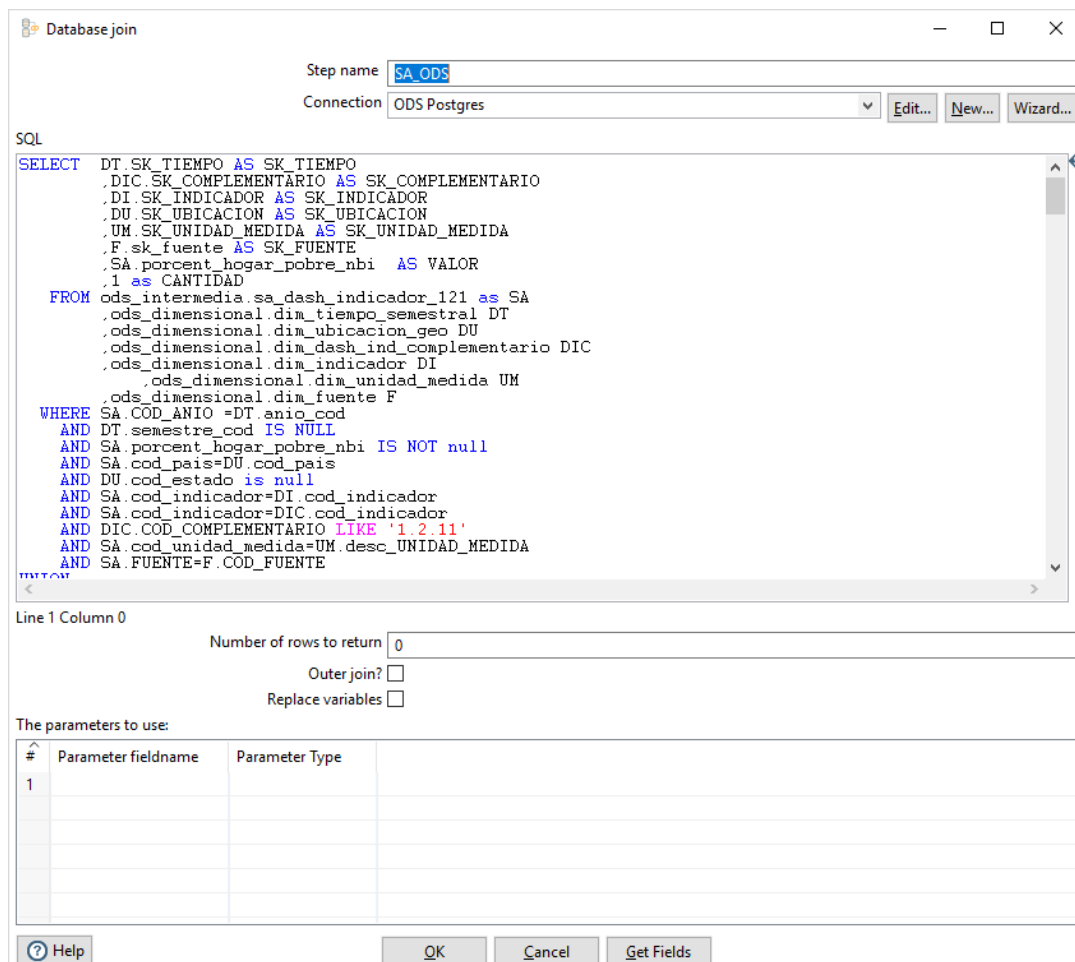


Figura 66. Consulta que obtiene los datos a ser cargados en la tabla de hechos

Insert / update

Step name:

Connection:

Target schema:

Target table:

Commit size:

Don't perform any updates:

The key(s) to look up the value(s):

#	Table field	Comparator	Stream field1	Stream field2
1	sk_tiempo	=	sk_tiempo	
2	sk_complementario	=	sk_complementario	
3	sk_indicador	=	sk_indicador	
4	sk_ubicacion	=	sk_ubicacion	
5	sk_unidad_medida	=	sk_unidad_medida	
6	sk_fuente	=	sk_fuente	

Update fields:

#	Table field	Stream field	Update
1	sk_tiempo	sk_tiempo	N
2	sk_complementario	sk_complementario	N
3	sk_indicador	sk_indicador	N
4	sk_ubicacion	sk_ubicacion	N
5	sk_unidad_medida	sk_unidad_medida	N
6	sk_fuente	sk_fuente	N
7	valor	valor	Y
8	cantidad	cantidad	Y
9	sk_fact_dash_ind121	sk_fact_dash_ind121	N

Figura 67. Paso de carga hacia la tabla FACT_DASH_IND_121

Finalmente se crean los Jobs encargados de agrupar las transformaciones creadas, en total se crearon 5 Jobs, siendo 4 que agrupan las transformaciones que cargan las tablas de hecho por cada ODS, y un último Job que se encarga de agrupar todos los Jobs creados, implementando un control de alerta de posibles en errores. En la Figura 68 se puede ver el Job creado para el ODS 1, mientras que en la Figura 69 se encuentra el Job final encargado de agrupar el resto de los Jobs.

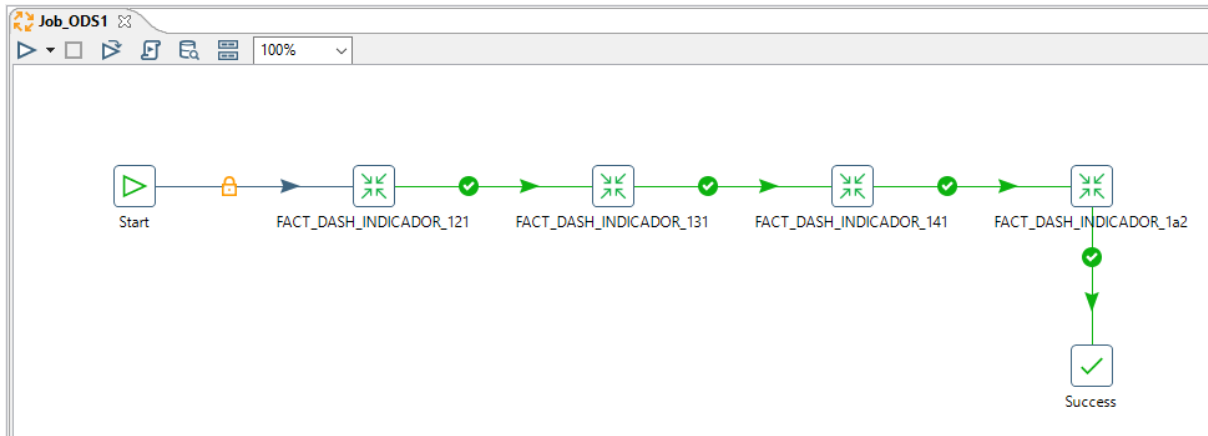


Figura 68. Job para la carga de las tablas de hechos del ODS 1

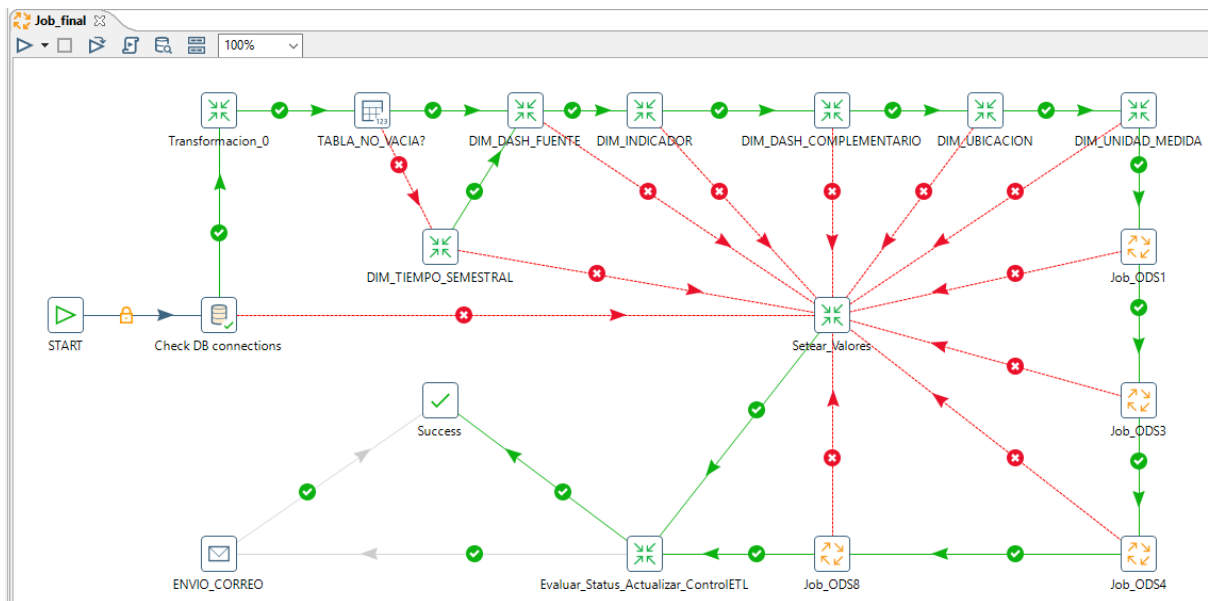


Figura 69. Job final que agrupa el resto de los Jobs y control de errores

4.5 Flujo de aplicaciones de Inteligencia de Negocio

El flujo de aplicaciones de Inteligencia de Negocio, agrupa todas aquellas actividades del Ciclo de Vida Dimensional que se encargan de las visualizaciones que se generan a partir de los datos almacenados en el almacén de datos y que dan respuesta a los indicadores propuestos. A continuación, se detallan las tareas realizadas en cada una de estas actividades.

4.5.1 Especificación de aplicaciones BI

Para esta actividad del Ciclo de Vida Dimensional, se decidió representar cada indicador de tres maneras diferentes, tal como fue indicado en los requerimientos del negocio: en cuadros estadísticos, gráficos de barras y mapas para aquellos indicadores en donde aplique. Se decidió estandarizar dichas visualizaciones, especificando como color principal el rojo para los gráficos, tipo de letra Arial y con un color de fondo idéntico entre todos los indicadores. Además cada indicador debe reflejar su título, la cobertura geográfica que representa, su unidad de medida y la fuente que genera los datos del indicador. En la Figura 70, 71 y 72 se muestran los estándares para cada tipo de representación.

Cuadro 1. Mortalidad neonatal

Año	Tasa
1990	14.04
1991	12.42
1992	13.50
1993	14.79
1994	14.09
1995	13.54
1996	13.82
1997	12.97
1998	12.39
1999	11.88
2000	11.52
2001	11.04
2002	11.92
2003	11.23
2004	11.62
2005	10.79
2006	9.99
2007	9.90
2008	9.81
2009	10.42
2010	10.83
2011	10.81
2012	11.04
2013	10.94
2014	11.38

Cobertura Geográfica: Nacional

Unidad de Medida: Por 1.000 Nacidos Vivos

Fuente: Ministerio del Poder Popular para la Salud

Cuadro 1. Mortalidad neonatal

Año	Tasa
1990	14.04
1991	12.42
1992	13.50
1993	14.79
1994	14.09
1995	13.54
1996	13.82
1997	12.97
1998	12.39
1999	11.88
2000	11.52
2001	11.04
2002	11.92
2003	11.23
2004	11.62
2005	10.79
2006	9.99
2007	9.90
2008	9.81
2009	10.42
2010	10.83
2011	10.81
2012	11.04
2013	10.94
2014	11.38

Cobertura Geográfica: Nacional
 Unidad de Medida: Por 1.000 Nacidos Vivos
 Fuente: Ministerio del Poder Popular para la Salud

Figura 70. Estándar de visualizaciones para cuadros estadísticos

Gráfico 1. Personas en hogares con pobreza extrema (Por Ingreso)

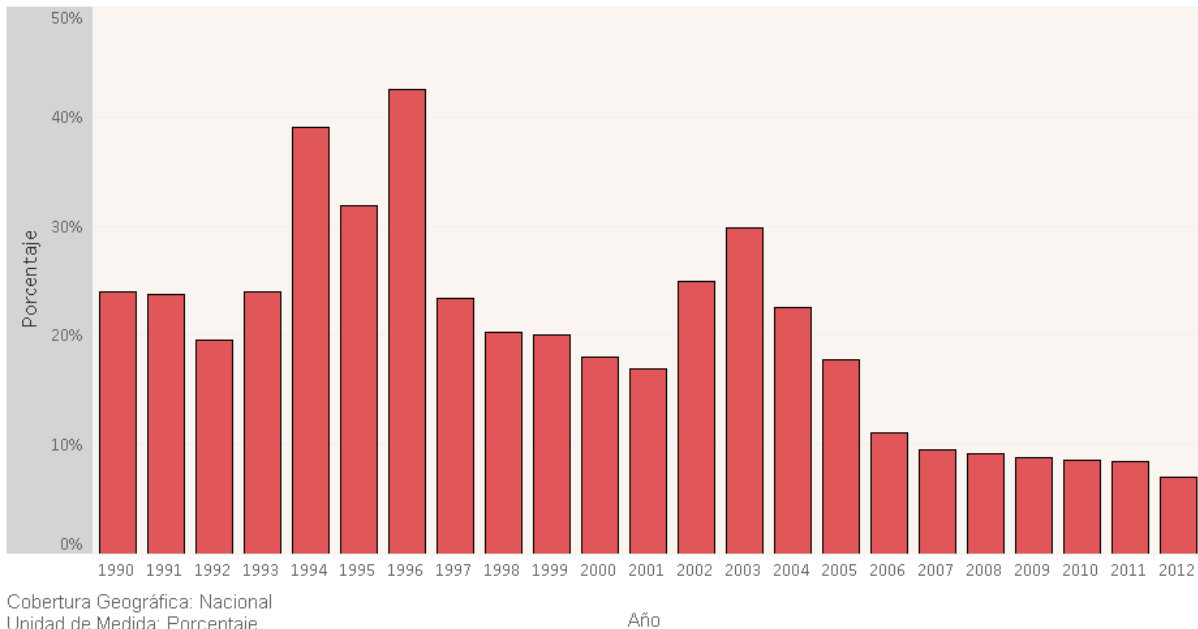
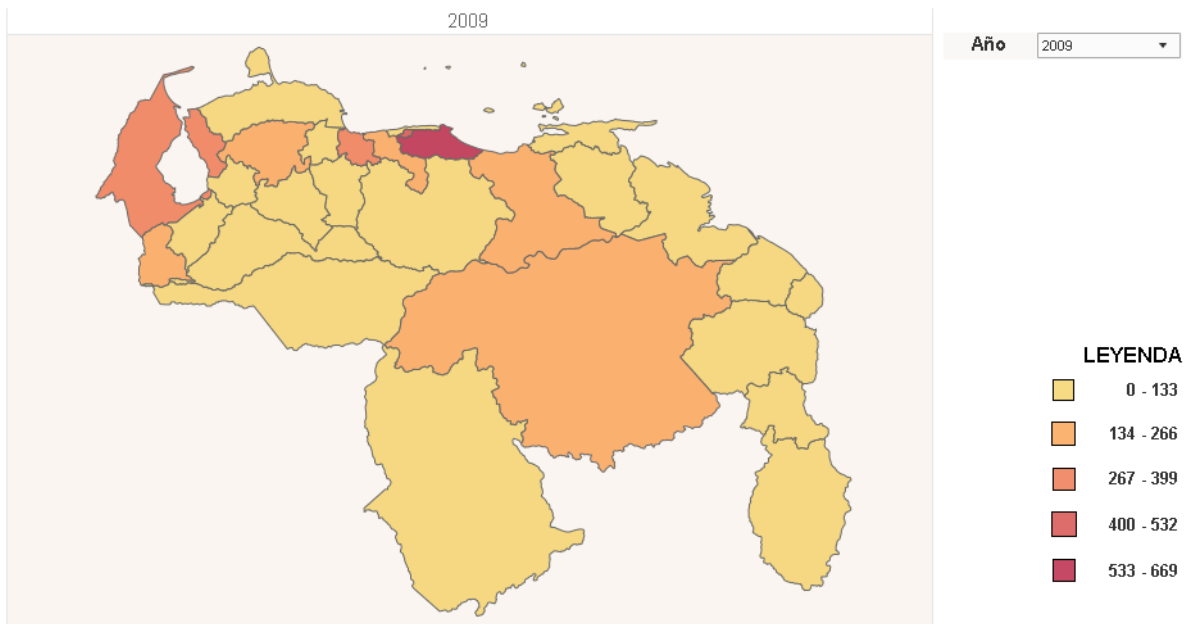


Figura 71. Estándar de visualizaciones para gráficos



Cobertura Geográfica: Regional
 Unidad de Medida: Sucursales Bancarias por cada 100.000 adultos
 Fuente: Superintendencia de las Instituciones del Sector Bancario

Figura 72. Estándar de visualizaciones para mapas

4.5.2 Desarrollo de aplicaciones BI

Para el desarrollo de las aplicaciones BI, se realiza una conexión entre el almacén de datos y la capa lógica de Tableau, como se muestra en la Figura 73.

fact_dash_ind321

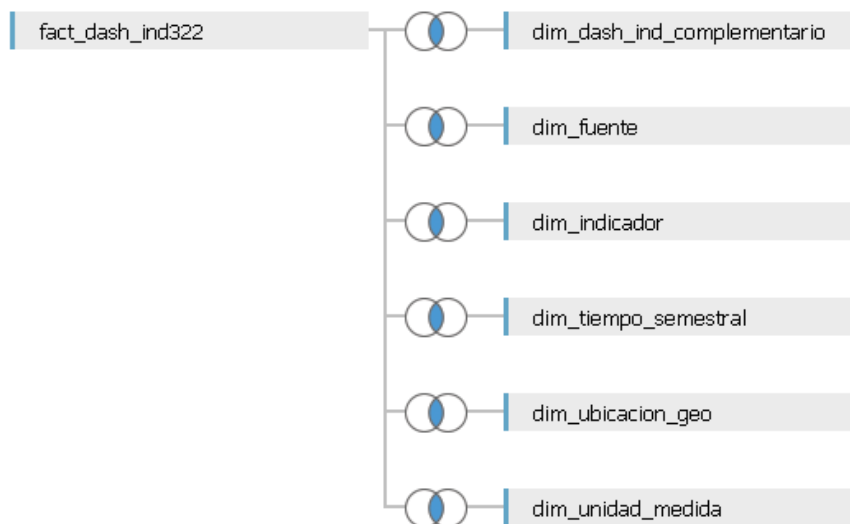


Figura 73. Capa lógica de Tableau

Dicha conexión debe de hacerse por cada uno de las bodegas de datos o tablas de hechos creadas. Una vez que se tiene la conexión establecida, se debe hacer una limpieza de los campos, ocultando aquellos que no tendrán uso para el usuario final (como las claves subrogadas) y renombrando el resto con nombres más representativos para facilitar su uso. En la Figura 74 se puede observar la interfaz para realizar este proceso y en la Figura 75 el resultado una vez culminado.

# dim_dash_ind_complementario Sk Complementa...	Abc dim_dash_ind_complementario Cod Indicador	Abc dim_dash_ind_complementario Desc Indicador	Abc dim_dash_ind_complementario Descripción Indic...	Abc dim_dash_ind_complementario Cod Complemen...
60	3.2.2	Tasa de mortalidad n...	Tasa	3.2.260
60	3.2.2	Tasa de mortalidad n...	Tasa	3.2.260
60	3.2.2	Tasa de mortalidad n...	Tasa	3.2.260
60	3.2.2	Tasa de mortalidad n...	Tasa	3.2.260
60	3.2.2	Tasa de mortalidad n...	Tasa	3.2.260
60	3.2.2	Tasa de mortalidad n...	Tasa	3.2.260

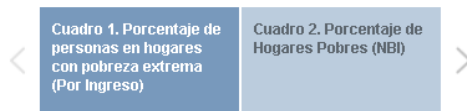
Figura 74. Representación de los campos crudos del almacén de datos en Tableau

Abc dim_dash_ind_complementario Indicador	Abc dim_dash_ind_complementario Forma de Medición	Abc dim_fuente Fuente	# dim_tiempo... Año	Abc dim_ubicacion_geo Cobertura Geogr...	Abc dim_unidad_medida Unidad de Medida
Tasa de mortalidad n...	Tasa	Ministerio del Poder ...	1990	Nacional	Por 1.000 Nacidos Vi...
Tasa de mortalidad n...	Tasa	Ministerio del Poder ...	1991	Nacional	Por 1.000 Nacidos Vi...
Tasa de mortalidad n...	Tasa	Ministerio del Poder ...	1992	Nacional	Por 1.000 Nacidos Vi...
Tasa de mortalidad n...	Tasa	Ministerio del Poder ...	1993	Nacional	Por 1.000 Nacidos Vi...
Tasa de mortalidad n...	Tasa	Ministerio del Poder ...	1994	Nacional	Por 1.000 Nacidos Vi...

Figura 75. Representación de los campos limpios del almacén de datos en Tableau

Una vez preparados los datos en Tableau, se puede proceder a la creación de las visualizaciones pertenecientes a cada indicador, siguiendo el estándar establecido anteriormente. A continuación, se presentan algunas de las visualizaciones creadas para cada indicador.

En la Figura 76 y 77 se muestra el cuadro estadístico creado para el indicador 1.2.1 Proporción de la población que vive por debajo del umbral nacional de pobreza, desglosada por sexo y edad. Como no se poseen datos para representar este indicador de forma directa, se optó por mostrar dos indicadores, los cuales son Porcentaje de Personas en Hogares con Pobreza Extrema (Por Ingreso) y Porcentaje de Hogares Pobres (NBI).

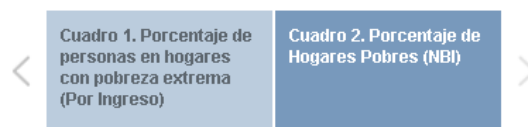


Cuadro 1. Personas en hogares con pobreza extrema (Por Ingreso)

Año	Porcentaje
1990	24.00
1991	23.80
1992	19.60
1993	24.00
1994	39.00
1995	31.90
1996	42.50
1997	23.40
1998	20.30
1999	20.10
2000	18.00
2001	16.90
2002	25.00
2003	29.80
2004	22.50
2005	17.80
2006	11.10
2007	9.60
2008	9.20
2009	8.80
2010	8.60
2011	8.50
2012	7.10

Cobertura Geográfica: Nacional
 Unidad de Medida: Porcentaje
 Fuente: Instituto Nacional de Estadística

Figura 76. Cuadro estadístico del porcentaje de personas en hogares con pobreza extrema



Cuadro 2. Hogares Pobres (NBI)

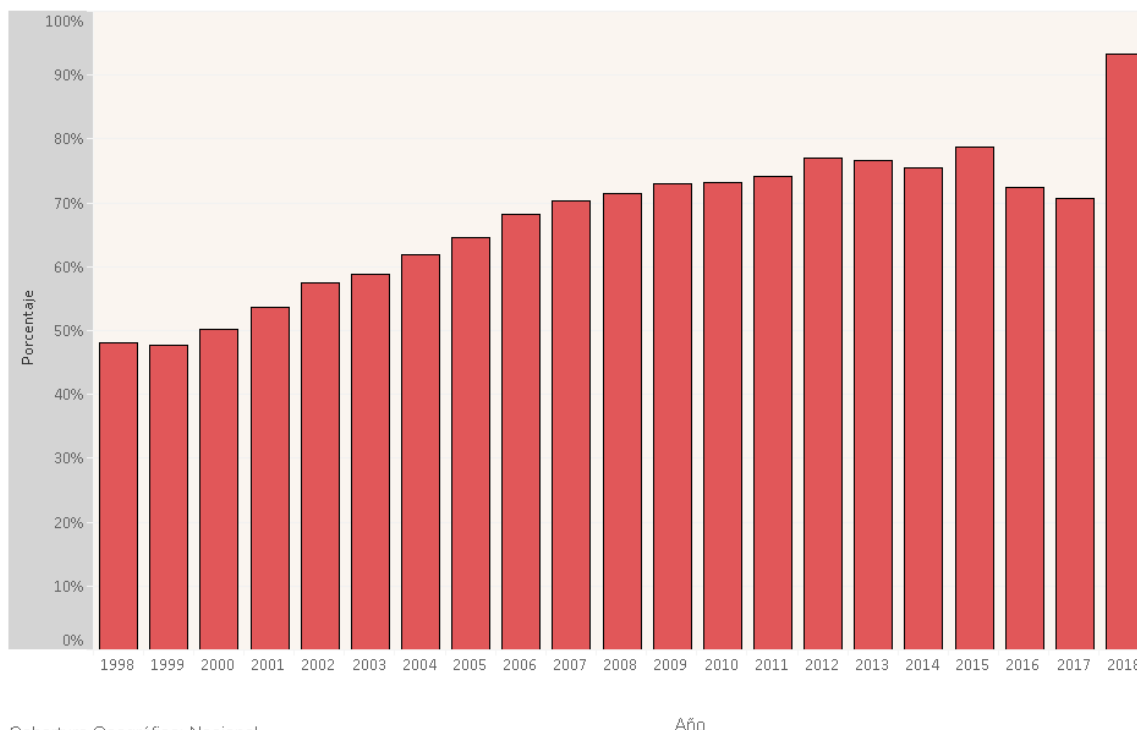
Año	Porcentaje
1999	29.26
2000	30.08
2001	27.83
2002	31.21
2003	30.50
2004	29.64
2005	26.72
2006	23.37
2007	23.27
2008	23.43
2009	23.55
2010	23.46
2011	21.25
2012	21.53
2013	19.59
2014	20.42
2015	19.70

Cobertura Geográfica: Nacional
 Unidad de Medida: Porcentaje
 Fuente: Instituto Nacional de Estadística

Figura 77. Cuadro estadístico del porcentaje de hogares pobres

Para el indicador 4.2.1 Proporción de niños menores de 5 años cuyo desarrollo es adecuado en cuanto a la salud, el aprendizaje y el bienestar psicosocial, desglosada por sexo, se tiene como indicador complementario el de Tasa Neta de Escolaridad Inicial, el cual se muestra su forma de representación en grafico en la Figura 78.

Gráfico 1. Escolaridad Inicial

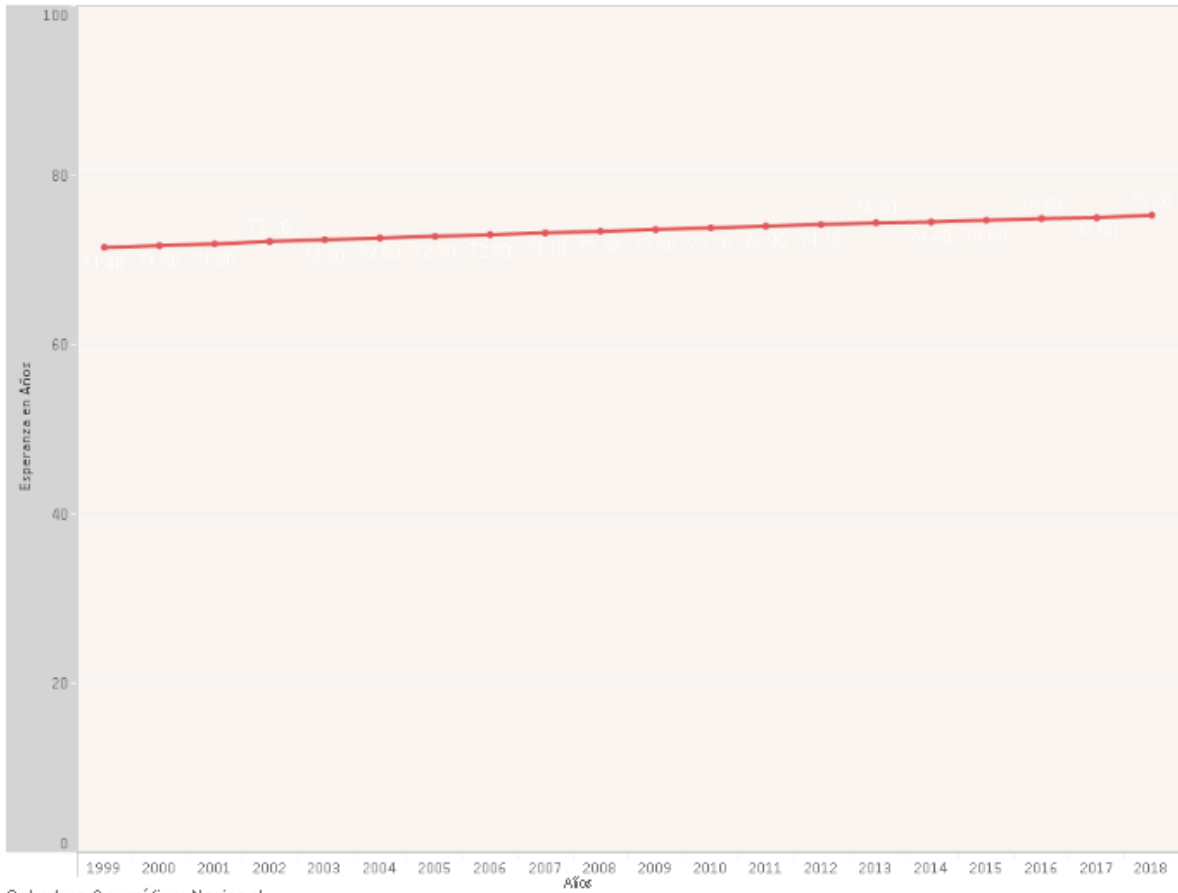


Cobertura Geográfica: Nacional
 Unidad de Medida: Porcentaje
 Fuente: Ministerio del Poder Popular para la Educación

Figura 78. Gráfico del indicador complementario Tasa Neta de Escolaridad Inicial

También se crean gráficos variados para aquellos indicadores que presentan diferentes niveles de desagregación, por ejemplo, para el indicador complementario Esperanza de Vida Nacional, del indicador 3.1.1, existe una desagregación por género, la cual fue representada como se muestra en las Figuras 79, 80 y 81.

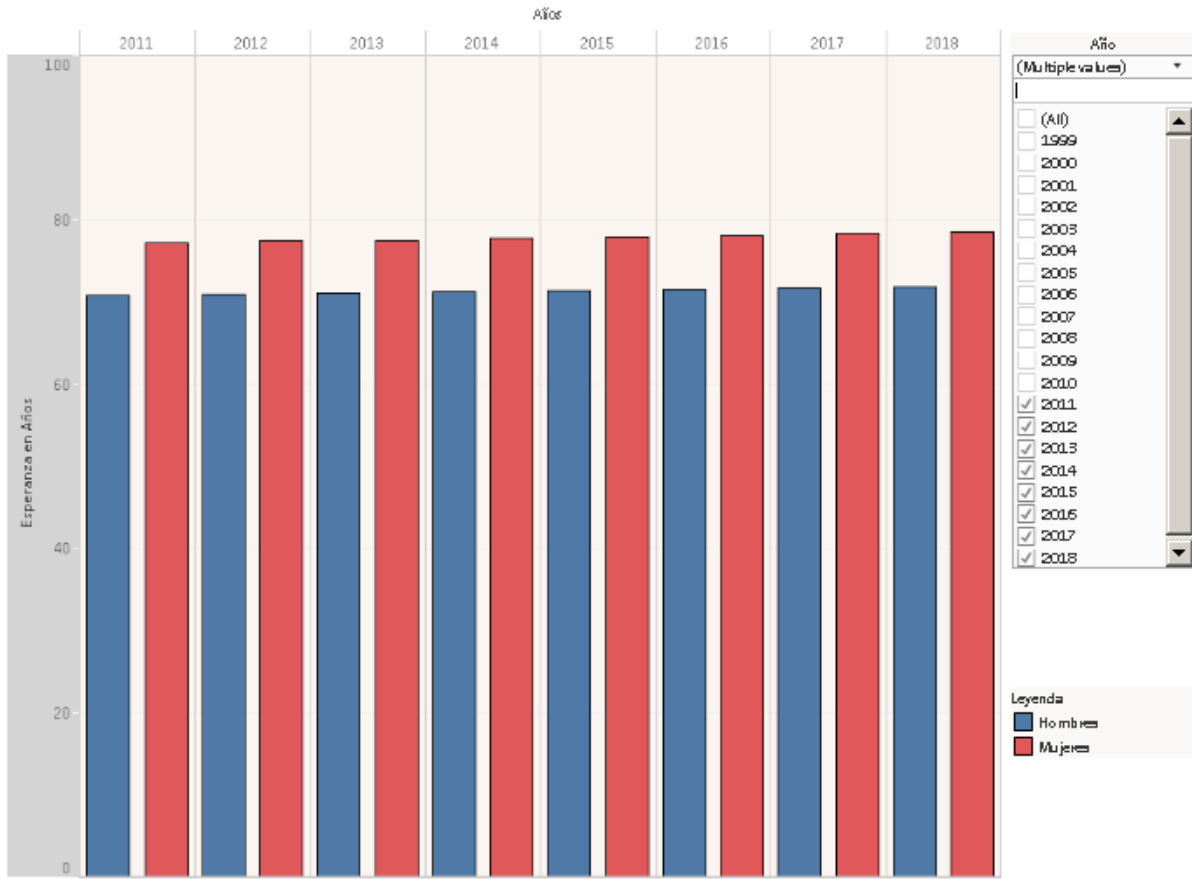
Gráfico 1. Esperanza de Vida Nacional



Cobertura Geográfica: Nacional
 Unidad de Medida: Año
 Fuente: Instituto Nacional de Nutrición

Figura 79. Gráfico de la esperanza de vida nacional

Gráfico 1. Esperanza de Vida de Hombres y Mujeres



Cobertura Geográfica: Nacional
 Unidad de Medida: Año
 Fuente: Instituto Nacional de Nutrición

Figura 80. Gráfico comparativo entre la esperanza de vida de hombres y mujeres

Gráfico 3. Esperanza de Vida por Género

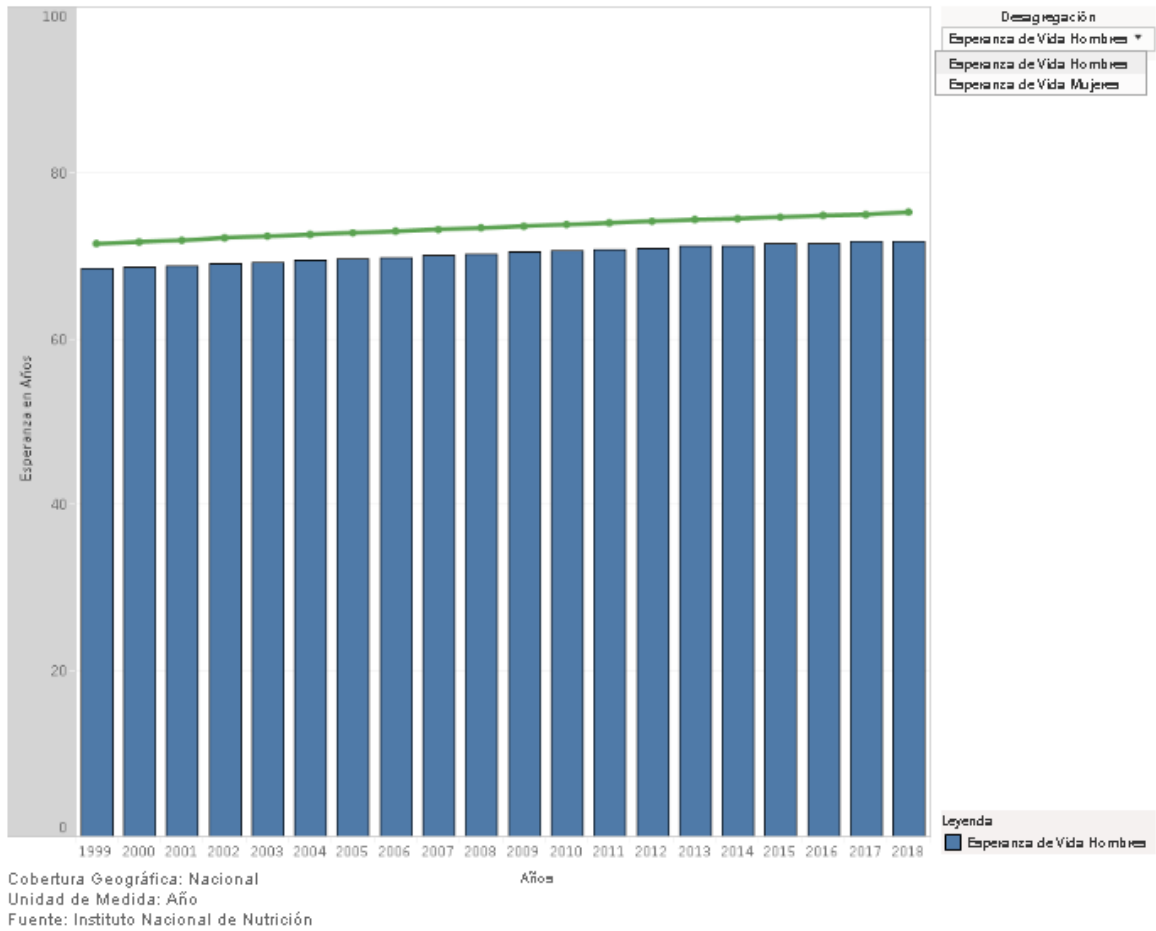


Figura 81. Gráfico comparativo entre la esperanza de vida nacional, de hombres y de mujeres

Cuando los datos se encuentran desagregados por semestre, los gráficos son mostrados como se observa en la Figura 82, la cual pertenece al indicador de Población Desocupada, perteneciente al indicador 8.5.2 Tasa de desempleo, desglosada por sexo, edad y personas con discapacidad.

Gráfico 1. Población Desocupada (M)
Año

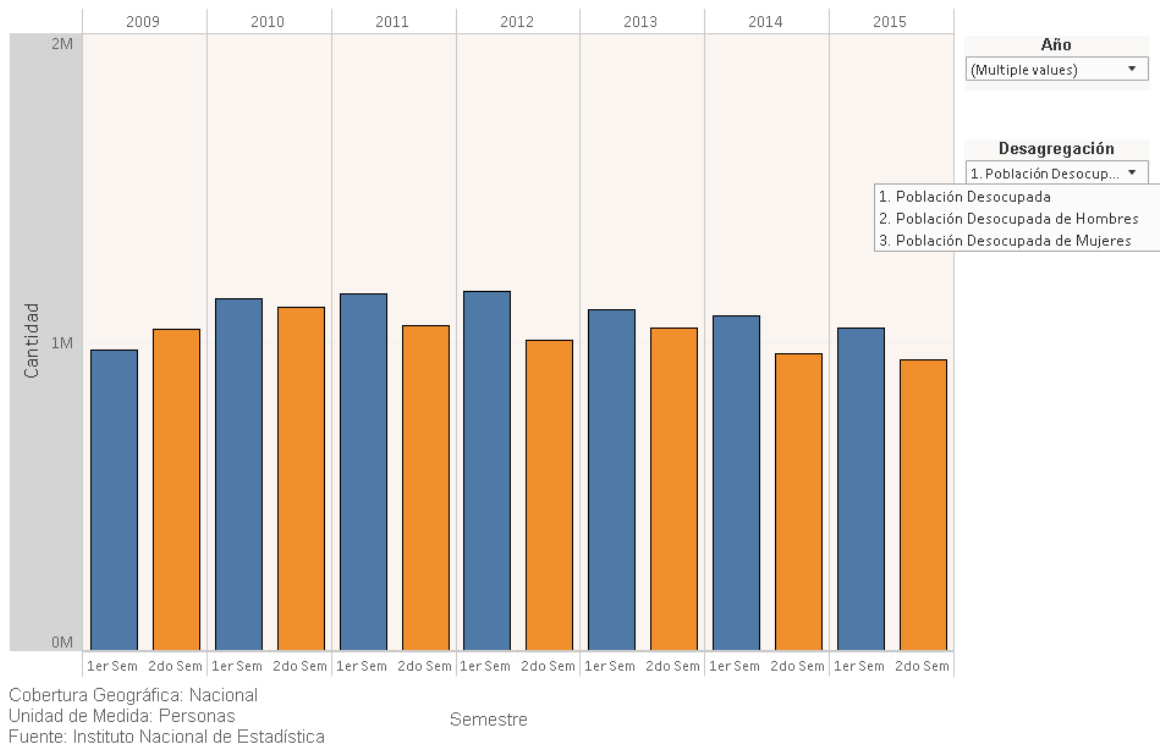


Figura 82. Gráfico de la población desocupada

Por último, para el indicador 8.10.1 se tiene una desagregación por Estados del país, por lo que se pueden representar los datos en un mapa, el cual se muestra en la Figura 83

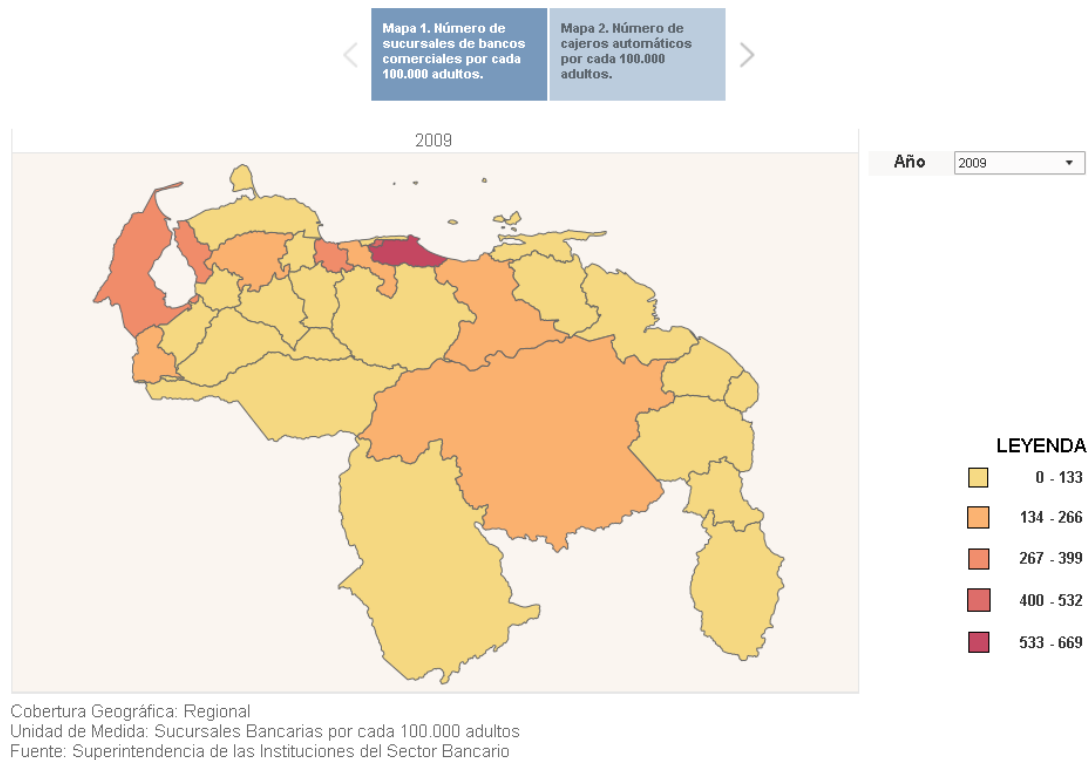


Figura 83. Mapa perteneciente al indicador 8.10.1

Finalmente, luego de haber creado cada una de las visualizaciones pertenecientes a cada indicador de la Agenda 2030, se crea un cuadro de mando que permite la navegación entre cada uno de los indicadores. En las Figura 84 se muestra el inicio de dicho cuadro de mando, el cual contiene una breve descripción de la Agenda 2030 y un botón para acceder al listado de ODS.

INE Instituto Nacional de Estadística
República Bolivariana de Venezuela

Solución de Inteligencia de Negocio para el Monitoreo y Seguimiento de la Agenda 2030 y los ODS

OBJETIVOS DE DESARROLLO SOSTENIBLE

Las Naciones Unidas en el 2015 plantean una serie de objetivos, conocidos como los Objetivos de Desarrollo Sostenible (ODS) que realizan un llamado universal a la adopción de medidas para poner fin a la pobreza, proteger el planeta y garantizar que todas las personas gocen de paz y prosperidad. Son 17 ODS que agrupa un conjunto de metas en los tres ámbitos del desarrollo sostenible (económico, social y ecológico).

Los ODS derivan en 169 metas planteando el horizonte 2015-2030 para ser aplicados por todos los países del mundo sean ricos o pobres y pueden ser representadas mediante indicadores que permiten su evaluación.

[Ingresar](#)

Figura 84. Inicio del cuadro de mando

Una vez seleccionado el botón de ingresar, se presenta el listado de los 17 ODS de la Agenda 2030, de los cuales se pueden seleccionar solo aquellos dentro del alcance de este TEG, es decir, el ODS 1, 3, 4 y 8. En la Figura 85 se puede ver dicho listado.

OBJETIVOS DE DESARROLLO SOSTENIBLE



Figura 85. Listado de los 17 ODS

Al seleccionar uno de los ODS, se presentan las metas asociadas a dicho ODS, como se observa en la Figura 86.

1 FIN DE LA POBREZA



ODS 1. Poner fin a la pobreza en todas sus formas y en todo el mundo

METAS



Figura 86. Listado de metas del ODS 1

Las metas cuyos indicadores no se poseen datos se presentan con un color gris y no son seleccionables, como se observa en la Figura 87.



Figura 87. Icono en gris indicando metas que no poseen datos

Al seleccionar una meta que si posee indicadores, se muestra el listado de dichos indicadores, tal como se muestra en la Figura 88.

1 FIN DE LA POBREZA



Meta 1.2 De aquí a 2030, reducir al menos a la mitad la proporción de hombres, mujeres y niños de todas las edades que viven en la pobreza en todas sus dimensiones con arreglo a las definiciones nacionales

INDICADORES

1.2.1

Proporción de la población que vive por debajo del umbral nacional de pobreza, desglosada por sexo y edad

Figura 88. Listado de indicadores de la meta 1.2

Al seleccionar entonces el indicador 1.2.1, se muestran las visualizaciones asociadas a ese indicador, comenzando por los indicadores principales, como se muestra en la Figura 89.



Figura 89. Indicador principal asociados al indicador 1.2.1

Existen flechas de navegación que permiten el desplazamiento entre los distintos indicadores que se encuentran disponibles, en la Figura 90 se observa el siguiente indicador luego de haber seleccionado la flecha de navegación.



Figura 90. Navegación entre los indicadores

Se obtiene el mismo resultado al navegar entre los indicadores utilizando el menú localizado en el lateral izquierdo, el cual refleja con color azul el indicador que se encuentra seleccionado actualmente, como se refleja en la Figura 91.

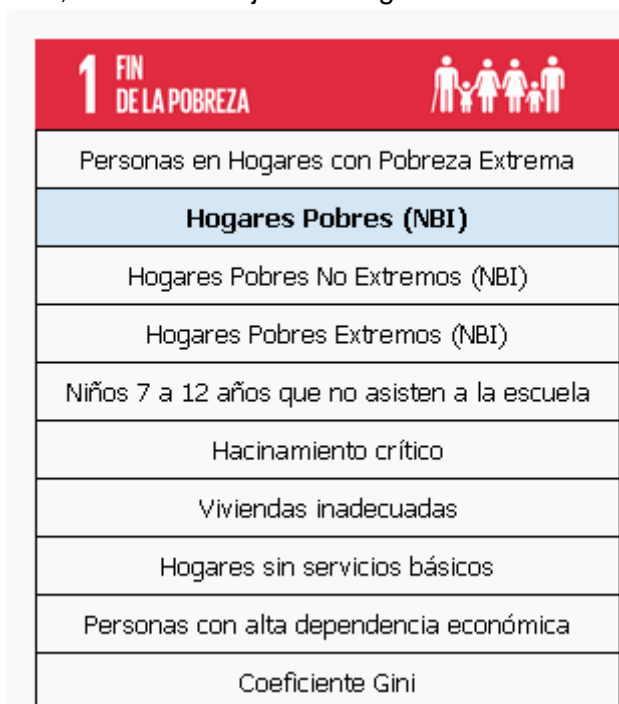


Figura 91. Menú de Navegación de Indicadores

A su vez, también existe una forma de navegación entre las diferentes ventanas existentes, a través de la miga de pan mostrada en la Figura 92, que permite el desplazamiento entre el menú inicial, el listado de ODS, las metas y los indicadores.



Figura 92. Menú de Miga de Pan

4.6 Implementación

En esta actividad del Ciclo de Vida Dimensional, se acopla cada componente de la solución de inteligencia de negocio de tal manera que la ejecución de los procesos de Extracción, Transformación y Carga sea periódica, con el fin de tener acceso a las diversas fuentes de datos con las que se poblará el área de preparación y ésta a su vez sirva como fuente de datos del almacén de datos.

Una vez poblado el almacén de datos con los datos correspondientes, se generan las visualizaciones que pueden mostrar el estado de cada uno de los indicadores de la Agenda 2030.

El usuario final puede consultar estas visualizaciones mediante el cuadro de mando de acceso creado, y crear, modificar o eliminar dichas visualizaciones con el uso de la herramienta Tableau.

Con el fin de garantizar la calidad de los datos al generar las visualizaciones asociadas a los indicadores de la Agenda 2030, los resultados son comparados con los datos crudos provenientes de los archivos de Excel fuente.

4.7 Mantenimiento y crecimiento

Para el desarrollo de la solución de Inteligencia de Negocio, se realizaron diferentes ajustes de mantenimiento y crecimiento, tales como la incorporación continua de indicadores mediante diversas mesas de trabajo con el personal de INE, modificación de algunos indicadores como el nombre que se despliega o la fuente que genero los datos; y el tratamiento de los archivos Excel fuente que permite agregar nuevos datos y series de tiempo de forma sencilla, garantizando el crecimiento en el tiempo.

4.8 Administración del proyecto

Durante el desarrollo de la solución de Inteligencia de Negocio, fue necesaria la gestión y redistribución de los recursos de tiempo disponibles con el fin de poder llevar a cabo las tareas relacionadas con el acceso a las diversas fuentes de datos, el procesamiento de éstos y su carga en el almacén de datos; la creación de las diversas visualizaciones correspondientes a cada indicador, así como el cuadro de mando para la navegación; así como también la adquisición de conocimientos acerca de las herramientas de software utilizadas.

Para garantizar la carga correcta de los datos, se realizan diversas pruebas de carga y calidad de datos para cada uno de los esquemas creados en la base de datos.

En primera instancia, se realizan corridas recurrentes tanto de “Carga inicial” como “Carga Incremental” en ambiente de desarrollo, y se valida que todos los registros que existen en la fuente de datos de cada Indicador (archivo Excel), son cargados correctamente en las estructuras dimensionales correspondientes.

Adicionalmente se realizan actividades de Calidad del Dato, verificando la veracidad y consistencia de los datos dentro de las estructuras dimensionales (bodegas de datos), de acuerdo con los datos suministrados por la INE en las estructuras fuentes en ambiente de desarrollo.

También se realizan ajustes y pruebas a los ETL de las cargas tanto de las SA como de las estructuras dimensionales creadas.

Con el propósito de verificar si fueron satisfechos los requerimientos de los usuarios finales y por ende si se cumplió con sus expectativas, se tomaron en cuenta sus recomendaciones y se realizan pruebas de aceptación a través de la aplicación de una encuesta.

Dicha encuesta fue aplicada a 25 personas, distribuidos entre miembros de la escuela de computación de la UCV y empleados del INE, de manera de simular el uso que se le dará a la solución de Inteligencia de Negocio por el usuario final.

En la encuesta realizada se presenta una serie de preguntas que cuentan con las opciones de respuesta definidas en la escala de Likert:

- Completamente en desacuerdo
- En desacuerdo
- Ni de acuerdo ni en desacuerdo
- De acuerdo
- Completamente de acuerdo

Las preguntas que conforman la encuesta pueden verse en la tabla 15.

Tabla 15. Prueba de Aceptación

1	¿La navegación se realiza de manera intuitiva?
2	¿La distribución de los elementos visuales es adecuada?
3	¿Se distinguen los elementos disponibles de los no disponibles?
4	¿Se presenta con claridad la información de los gráficos de cada indicador?
5	¿Se presenta con claridad la información de las tablas de cada indicador?
6	¿Se presenta con claridad la información de los mapas de cada indicador? (mapas disponibles en el indicador 8.10.1)
7	¿Los colores empleados son adecuados?
8	¿El diseño es consistente?

A continuación se muestran los resultados obtenidos en la aplicación de esta prueba de aceptación:

¿La navegación se realiza de manera intuitiva?

25 respuestas

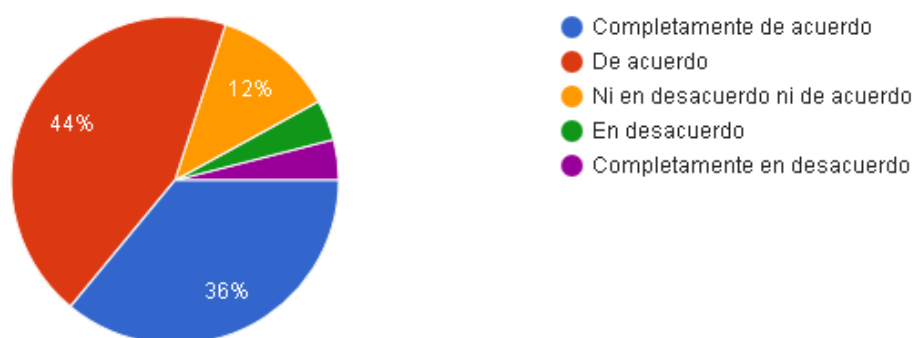


Figura 93. Prueba de aceptación – Resultados Pregunta 1

En la Figura 93 se pueden ver los resultados en cuanto a la navegación de los cuadros de mando, la mayoría de las personas se encuentran satisfechas con la forma en que se maneja, siendo una pequeña minoría las que presentan inconformidad, por esto se puede decir que la navegación entre los diferentes cuadros de mando, ODS, metas e indicadores es adecuada e intuitiva.

¿La distribución de los elementos visuales es adecuada?

25 responses

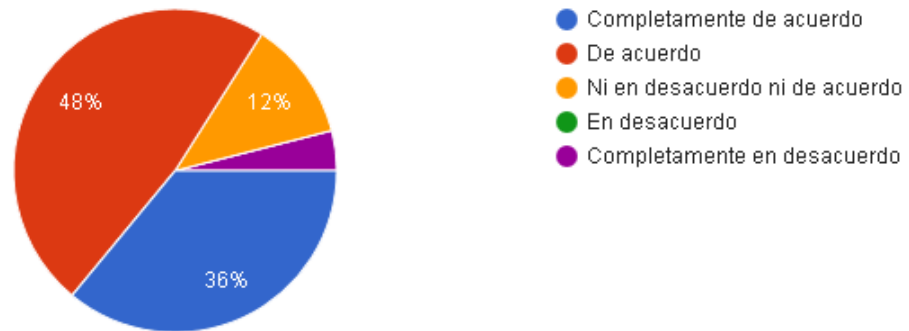


Figura 94. Prueba de aceptación – Resultados Pregunta 2

La distribución de los elementos visuales, desde los botones hasta el posicionamiento de los indicadores presenta un alto porcentaje de aceptación por los encuestados, como se puede ver en la Figura 94, de manera que no es necesario realizar cambios con respecto a este apartado.

¿Se distinguen los elementos disponibles de los no disponibles?

25 responses

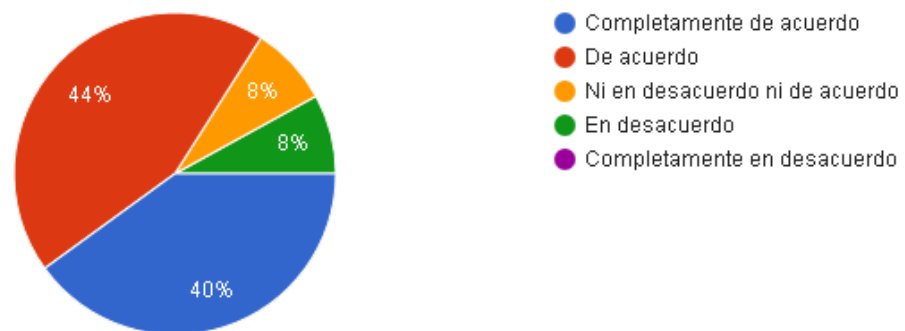


Figura 95. Prueba de aceptación – Resultados Pregunta 3

La distinción de los elementos disponibles de los no disponibles es importante para la navegación, ya que no se cuentan con indicadores para todas las metas existentes, y a la vez tampoco se abarcan todos los 17 ODS de la Agenda 2030. En su mayoría, los encuestados han podido hacer distinción de dichos elementos, con una baja minoría inconforme en cómo se maneja este aspecto, tal como se puede ver en la Figura 95.

¿Se presenta con claridad la información de los gráficos de cada indicador?

25 responses

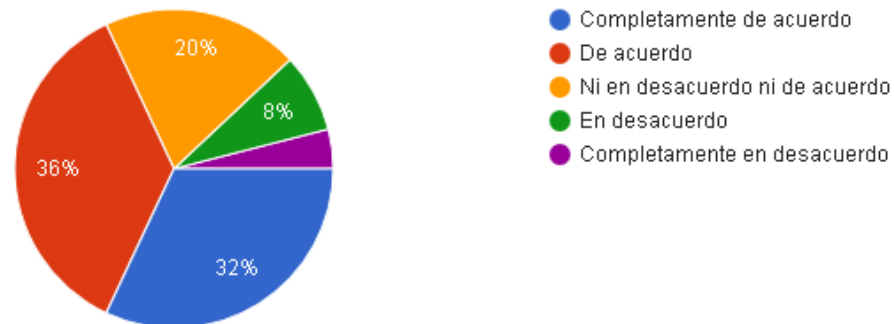


Figura 96. Prueba de aceptación – Resultados Pregunta 4

El diseño de los indicadores de tipo de gráfico puede ser realizado de diversas maneras, siendo complicado poder satisfacer a todos los usuarios con una misma manera de presentarlos. Esto se puede ver en la Figura 96, donde las respuestas son variadas, pero mostrando que la mayoría de los encuestados están conformes con la manera de que se presenta la información en cada uno de los gráficos mostrados.

¿Se presenta con claridad la información de las tablas de cada indicador?

25 responses

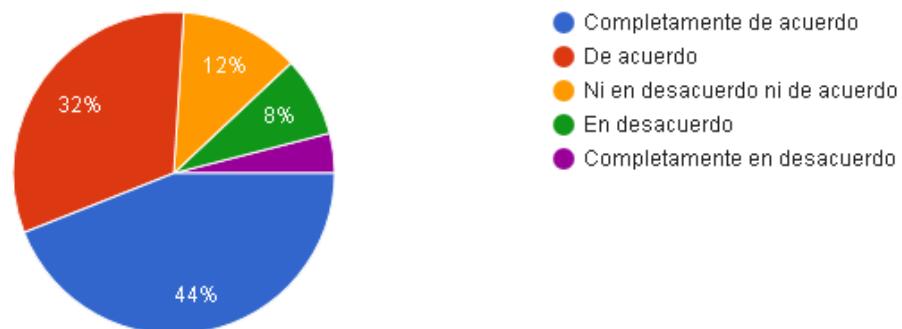


Figura 97. Prueba de aceptación – Resultados Pregunta 5

Las tablas o cuadros estadísticos con las que se presentan la información referente a cada indicador debe de ser entendible y clara para el usuario. En la Figura 97 se puede ver que para la mayoría de los encuestados esto es así, siendo una pequeña minoría los que se encuentran inconformes.

¿Se presenta con claridad la información de los mapas de cada indicador? (mapas disponibles en el indicador 8.10.1)

25 responses

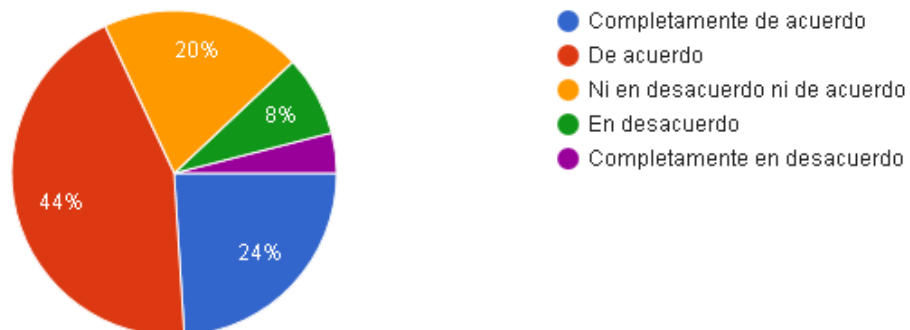


Figura 98. Prueba de aceptación – Resultados Pregunta 6

En la Figura 98 se presentan los resultados con respecto a la forma de presentar la información en los mapas desarrollados en los cuadros de mando. Aun cuando solo se presentan mapas para un indicador en particular, es importante conocer si se muestra de forma clara la información plasmada en ellos. La mayoría de los encuestados se encontraron satisfechos con la manera en que se presenta la información en dichos mapas.

¿Lo colores empleados son adecuados?

25 responses

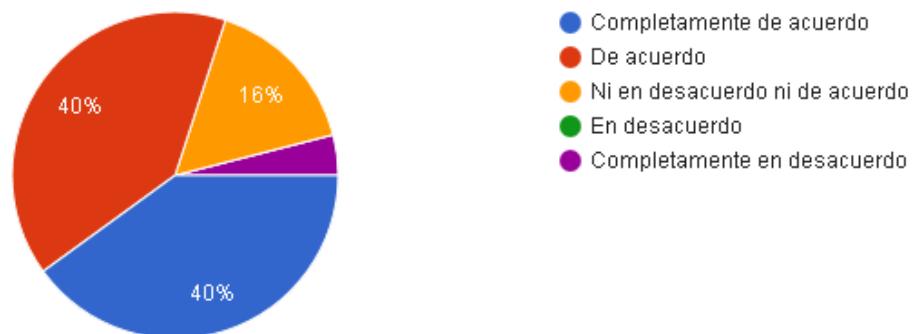


Figura 99. Prueba de aceptación – Resultados Pregunta 7

El uso de los colores en un cuadro de mando es muy importante para garantizar la comodidad de los usuarios finales, en la Figura 99 se puede ver que la mayoría de los encuestados estuvieron conformes con el uso de los colores en cada uno de los diferentes cuadros de mando creados.

¿El diseño es consistente?

25 responses

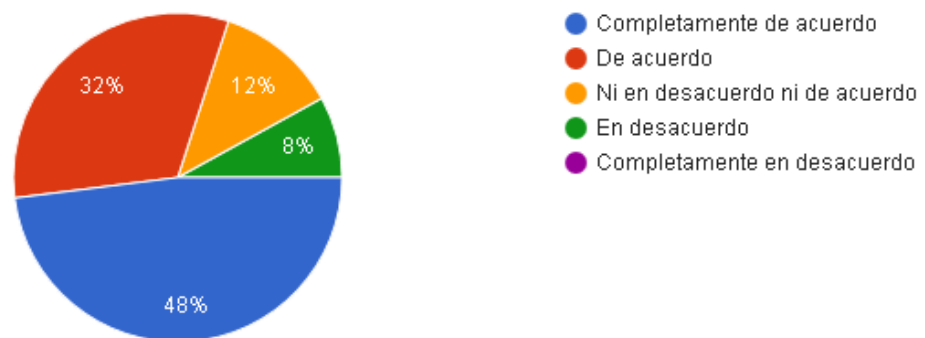


Figura 100. Prueba de aceptación – Resultados Pregunta 8

Finalmente, en la Figura 100 se puede ver los resultados de la pregunta referente a la consistencia del diseño de los cuadros de mando, siendo una gran mayoría de los encuestados los que les parece adecuado y consistente.

CONCLUSIONES

Una vez finalizado el desarrollo de la solución de Inteligencia de Negocio para la obtención de indicadores relacionados con los ODS para el INE, se llegan a las siguientes conclusiones:

Se analizaron satisfactoriamente los requerimientos funcionales y no funcionales del negocio, de manera de desarrollar una solución de Inteligencia de Negocio que fuese satisfactoria tanto para el INE, como para sus usuarios finales.

Se diseñaron los modelos dimensionales que cumpliesen con lo necesario para poder dar respuesta a los indicadores de los ODS trabajados, llegando al diseño e implementación de 29 modelos dimensionales.

En cuanto al diseño e implementación del almacén de datos, encargado de almacenar todos aquellos datos que dan respuesta a los indicadores de los ODS, este se diseñó e implementó de forma correcta, a partir de la creación de las 4 bodegas de datos creadas que conforman el almacén de datos.

Para el llenado de dichas bodegas de datos, se crearon de forma satisfactoria 39 transformaciones encargadas de extraer, transformar y cargar los datos de manera de que fuesen unificados dentro del almacén de datos. De igual forma se crearon otras 57 transformaciones para el llenado del área de preparación.

Se construyeron las diversas visualizaciones para cada uno de los indicadores ODS, siendo de diversos tipos como cuadros estadísticos, gráficos y mapas. También se creó el cuadro de mando que permite la navegación entre los indicadores y sus visualizaciones.

Por último, se realizaron las pruebas de aceptación y calidad de datos para asegurar el desarrollo adecuado de todas las actividades que forman parte del Ciclo de Vida Dimensional utilizado en la Solución de Inteligencia de Negocio, En cuanto a la calidad de los datos, se garantizó la fidelidad de los mismos, de manera que los indicadores reflejan de manera correcta los datos suministrados. Por otro lado con las pruebas de aceptación se pudo ver que los usuarios finales se encuentran satisfechos con el diseño final presentado en los cuadros de mando y las visualizaciones de la Solución de Inteligencia de Negocio.

Finalmente se puede concluir que el desarrollo de una solución de Inteligencia de Negocio que permita la obtención, análisis y visualización de los Indicadores asociados a los ODS para apoyar a la toma de decisiones por parte de la organización fue un éxito, logrando abarcar la creación y despliegue de los datos correspondientes a 29 indicadores de los 4 ODS seleccionados.

El desarrollo de esta solución de Inteligencia de Negocio brinda grandes beneficios para la organización y para el país, ya que permite conocer de forma rápida y oportuna los avances que se han ido realizando en cada uno de los aspectos considerados en los ODS, de manera tal de poder tomar decisiones oportunas para seguir avanzando hacia el logro de estos objetivos. De igual manera, permite la carga de nuevos datos que se puedan ir generando, siendo entonces esta solución resistente al paso del tiempo.

TRABAJOS FUTUROS

Dada la experiencia obtenida a lo largo del desarrollo de esta solución de Inteligencia de Negocio, se proponen las siguientes recomendaciones para trabajos futuros. Estas recomendaciones no pudieron ser llevadas a cabo en esta etapa, debido a que estaban fuera del alcance de este trabajo:

- Continuar con el desarrollo de los modelos dimensionales, almacenes de datos, procesos ETL y transformaciones correspondientes a los ODS que no fueron abarcados en esta solución, de manera de poder dar respuesta de manera robusta, a cada uno de los indicadores propuestos por las Naciones Unidas en la Agenda 2030.
- Desarrollar una aplicación web o móvil que permita desplegar de forma más sencilla las visualizaciones creadas, debido a que los cuadros de mando no permiten tanta flexibilidad como la que proveería una aplicación web o móvil.
- Implementar mayores niveles de desagregación a las visualizaciones, tales como, rangos de edad, género, actividad laboral, regiones geográfica de residencia y niveles de pobreza.

REFERENCIAS BIBLIOGRÁFICAS

Beltran, J. (2005). *Indicadores de Gestión: Herramientas para lograr la competitividad*. Disponible en: www.infoservi.com/infoservi/pdf/Indicadores_De_Gestion.pdf

Cano, J. (2007). *Business Intelligence: competir con información*. Disponible en: http://itemsweb.esade.edu/biblioteca/archivo/Business_Intelligence_competir_con_informacion.pdf

Comisión Mundial de Medio Ambiente y Desarrollo de Naciones Unidas. (1987). *Our Common Future*. Disponible en: http://www.krynica-zdroj.pl/files/files/File/Srodowisko/Brundtland_Report_Extracts.pdf

Davidson, L. Moss, J. (2016). *Pro SQL Server Relational Database Design and Implementation*. Disponible en: <https://books.google.co.ve/books?id=x6PPDQAAQBAJ&pg=PA706&dq=Bill+Inmon+Top-down&hl=es-419&sa=X&ved=0ahUKEwjbgJTLtpzaAhVQs1kKHezaDIEQ6AEIOzAC#v=onepage&q=Bill%20Inmon%20Top-down&f=false>

Dertiano, V. (2015). *Arquitectura BI (Parte III): El enfoque de Ralph Kimball*. Disponible en: <http://blog.mirai-advisory.com/arquitectura-bi-parte-iii-el-enfoque-de-ralph-kimball/>

Devece, R. Guiral, J. y Lapiedra, R. (2011). *Introducción a la gestión de sistemas de información en la empresa*. Disponible en: <https://libros.metabiblioteca.org/bitstream/001/193/8/978-84-693-9894-4.pdf>

Franco, T. (2019). *OLTP vs OLAP: what's the difference between them?* Disponible en: <https://www.imaginarycloud.com/blog/oltp-vs-olap/>

Frost, R. Kenyo, L. Pels, S y Pike, J. (2012). *Designing Business Information Systems: Apps, Websites, and More*. Disponible en: <https://2012books.lardbucket.org/books/designing-business-information-systems-apps-websites-and-more/index.html>

GantaBI. (2018). *5 ventajas de los Indicadores clave de rendimiento (KPI)*. Disponible en: <https://www.gantabi.com/2018/02/20/los-indicadores-kpi/>

Gómez, C. (2006). *El Desarrollo Sostenible: Conceptos básicos, alcance y criterios para su evaluación*. Disponible en: <http://www.unesco.org/new/fileadmin/MULTIMEDIA/FIELD/Havana/pdf/Cap3.pdf>

Guru99. (2018). *What is Dimensional Model in Data Warehouse?* Disponible en: <https://www.guru99.com/dimensional-model-data-warehouse.html#4>

Hitachi Ventara. (2019). *Pentaho Data Integration*. Disponible en: <https://www.hitachivantara.com/go/pentaho.html>

INEI Perú. (2018). *PERÚ: SISTEMA DE MONITOREO Y SEGUIMIENTO DE LOS INDICADORES DE LOS OBJETIVOS DE DESARROLLO SOSTENIBLE*. Disponible en: <http://ods.inei.gob.pe/ods/>

INEGI México. (2018). *Sistema de Información de los Objetivos de Desarrollo Sostenible*. Disponible en: <http://agenda2030.mx>

Inmon, B. (2002). *Building the Data Warehouse, 4th Edition*. Disponible en: <http://download.e-booksshelf.de/download/0000/5851/95/L-G-0000585195-0002361324.pdf>

Inmon, B. (1998). *Data Mart Does Not Equal Data Warehouse*. Disponible en: <https://www.information-management.com/news/data-mart-does-not-equal-data-warehouse>

Kimball, R. (1997). *A Dimensional Modeling Manifesto*. Disponible en: <https://www.kimballgroup.com/1997/08/a-dimensional-modeling-manifesto/>

Kimball, R. Margy, R. (2013). *The Data Warehouse Toolkit Third Edition*. Disponible en: <http://www.essai.rnu.tn/Ebook/Informatique/The%20Data%20Warehouse%20Toolkit,%203rd%20Edition.pdf>

Laudon, J. Laudon K. (2014). *Management Information Systems*. Disponible en: http://dinus.ac.id/repository/docs/ajar/MIS_KC_Laudon.pdf

Loshin, D. (2003). *Business Intelligence: The Savy Manager's Guide, Getting Onboard with Emerging IT*. Disponible en: https://books.google.co.ve/books/about/Business_Intelligence.html?id=VcJNigOP-osC&redir_esc=y

Maniviesa, P. (2015). *Qué son los objetivos SMART?* Disponible en: <http://www.pymerang.com/direccion-de-negocios/funciones-del-director-general/liderazgo/457-el-lider-nace-o-se-hace>

Mos, S. (2015). *¿Qué son los Sistemas de Apoyo a los Ejecutivos (ESS)?* Disponible en: <http://pertutatis.cat/que-son-los-sistemas-de-apoyo-a-los-ejecutivos-ess/>

Oracle. (2019). *Business Intelligence*. Disponible en: <https://www.oracle.com/solutions/business-analytics/business-intelligence/>

Oracle. (2019). *Oracle Database 12c*. Disponible en: <https://www.oracle.com/database/12c-database/>

Ortilles, J. (2012). *¿Por qué un almacén de datos es diferente a una base de datos operacionales?* Disponible en: <http://pentahohispano.blogspot.com/2012/01/por-que-un-almacen-de-datos-es.html>

PostgreSQL. (2019). *What is PostgreSQL?* Disponible en: <https://www.postgresql.org/about/>

Programa de las Naciones Unidas para el Desarrollo. (2019). *Objetivos de Desarrollo Sostenible*. Disponible en: <http://www.ve.undp.org/content/venezuela/es/home/sustainable-development-goals.html>

Rivadera, G. (2010). *La metodología de Kimball para el diseño de almacenes de datos (Data warehouses)*. Disponible en: <https://www.ucasal.edu.ar/hm/ingenieria/cuadernos/archivos/5-p56-rivadera-formateado.pdf>

Sinnexus (2017). *¿Qué es Business Intelligence?* Disponible en: https://www.sinnexus.com/business_intelligence/

Tableau. (2019). *Combine, shape, and clean your data for analysis with Tableau Prep*. Disponible en: <https://www.tableau.com/products/prep>

Tableau. (2019). *What is Tableau?* Disponible en: <https://www.tableau.com/products/what-is-tableau>

Turban, E. Aronson, J. (2005). *Decision Support Systems and Intelligent Systems*. Disponible en: https://www.academia.edu/11191200/TURBAN_EFRAIM_and_JAY_E._ARONSON_DECISION_SUPPORT_SYSTEMS_AND_INTELLIGENT_SYSTEMS_PRENTICE_HALL_UPPER_SADDLE_RIVER_NJ_1998

Wang, J. (2008). *Data Warehousing and Mining: Concepts, Methodologies, Tools, and Applications*. Disponible en: <http://homes.dico.unimi.it/~mdamiani/home/pubblicazioni/libroigi08.pdf>

ANEXOS

A. Contabilización de Indicadores

En la siguiente tabla se muestra detalladamente la distribución según ODS y metas de los indicadores desarrollados.

ODS	Meta	Indicador	Tipo de Indicador	Cantidad Principal	Cantidad Complementario
1 – Fin de la Pobreza	1.2 - De aquí a 2030, reducir al menos a la mitad la proporción de hombres, mujeres y niños de todas las edades que viven en la pobreza en todas sus dimensiones con arreglo a las definiciones nacionales	1.2.1 - Proporción de la población que vive por debajo del umbral nacional de pobreza, desglosada por sexo y edad	Eficacia	2	8
	1.3 - Implementar a nivel nacional sistemas y medidas apropiadas de protección social para todos, incluidos niveles mínimos, y, de aquí a 2030, lograr una amplia cobertura de las personas pobres y vulnerables	1.3.1 - Proporción de la población cubierta por sistemas o niveles mínimos de protección social, desglosada por sexo, distinguiendo entre los niños, los desempleados, los ancianos, las personas con discapacidad, las mujeres embarazadas, los recién nacidos, las víctimas de accidentes de trabajo, los pobres y los vulnerables	Eficacia	1	4
	1.4 - De aquí a 2030, garantizar que todos los hombres y mujeres, en particular los pobres y los vulnerables, tengan los mismos derechos a los recursos económicos y acceso a los servicios básicos, la propiedad y el control de la tierra y otros bienes, la herencia, los recursos naturales, las nuevas tecnologías apropiadas y los servicios financieros, incluida la micro financiación	1.4.1 - Proporción de la población que vive en hogares con acceso a los servicios básicos	Eficacia	1	19
	1.a - Garantizar una movilización significativa de recursos procedentes de diversas fuentes, incluso mediante la mejora de la cooperación para el desarrollo, a fin de proporcionar medios suficientes y previsibles a los países en desarrollo, en particular los países menos	1.a.2 - Proporción del gasto público total que se dedica a servicios esenciales (educación, salud y protección social)	Eficacia	1	7

	adelantados, para que implementen programas y políticas encaminados a poner fin a la pobreza en todas sus dimensiones				
ODS 3 - Salud y Bienestar	3.1 - De aquí a 2030, reducir la tasa mundial de mortalidad materna a menos de 70 por cada 100.000 nacidos vivos	3.1.1 - Tasa de mortalidad materna	Eficacia	1	2
	3.2 - De aquí a 2030, poner fin a las muertes evitables de recién nacidos y de niños menores de 5 años, logrando que todos los países intenten reducir la mortalidad neonatal al menos a 12 por cada 1.000 nacidos vivos y la mortalidad de los niños menores de 5 años al menos a 25 por cada 1.000 nacidos vivos	3.2.1 - Tasa de mortalidad de niños menores de 5 años	Eficacia	1	3
		3.2.2 - Tasa de mortalidad neonatal	Eficacia	1	0
	3.3 - De aquí a 2030, poner fin a las epidemias del SIDA, la tuberculosis, la malaria y las enfermedades tropicales desatendidas y combatir la hepatitis, las enfermedades transmitidas por el agua y otras enfermedades transmisibles	3.3.1 - Número de nuevas infecciones por el VIH por cada 1.000 habitantes no infectados, desglosado por sexo, edad y poblaciones clave.	Eficacia	1	0
		3.3.2 - Incidencia de la tuberculosis por cada 100.000 habitantes	Eficacia	1	1
		3.3.3 - Incidencia de la malaria por cada 1.000 habitantes	Eficacia	1	1
	3.4 - De aquí a 2030, reducir en un tercio la mortalidad prematura por enfermedades no transmisibles mediante su prevención y tratamiento, y promover la salud mental y el bienestar	3.4.1 - Tasa de mortalidad atribuida a las enfermedades cardiovasculares, el cáncer, la diabetes o las enfermedades respiratorias crónicas	Eficacia	5	0
		3.4.2 - Tasa de mortalidad por suicidio	Eficacia	1	0
	3.6 - De aquí a 2020, reducir a la mitad el número de muertes y lesiones causadas por accidentes de tráfico en el mundo	3.6.1 - Tasa de mortalidad por lesiones debidas a accidentes de tráfico	Eficacia	0	1
	3.7 - De aquí a 2030, garantizar el acceso universal a los servicios de salud sexual y reproductiva, incluidos los de planificación familiar, información y educación, y la integración de la salud	3.7.1 - Proporción de mujeres en edad de procrear (entre 15 y 49 años) que cubren sus necesidades de planificación familiar con métodos modernos	Eficacia	0	1

	reproductiva en las estrategias y los programas nacionales				
	3.8 - Lograr la cobertura sanitaria universal, incluida la protección contra los riesgos financieros, el acceso a servicios de salud esenciales de calidad y el acceso a medicamentos y vacunas inocuos, eficaces, asequibles y de calidad para todos	3.8.1 - Cobertura de los servicios de salud esenciales (definida como la cobertura media de los servicios esenciales entre la población general y los más desfavorecidos, calculada a partir de intervenciones trazadoras como las relacionadas con la salud reproductiva, materna, neonatal e infantil, las enfermedades infecciosas, las enfermedades no transmisibles y la capacidad de los servicios y el acceso a ellos)	Eficacia	0	3
	3.b - Apoyar las actividades de investigación y desarrollo de vacunas y medicamentos contra las enfermedades transmisibles y no transmisibles que afectan primordialmente a los países en desarrollo y facilitar el acceso a medicamentos y vacunas esenciales asequibles de conformidad con la Declaración relativa al Acuerdo sobre los Aspectos de los Derechos de Propiedad Intelectual Relacionados con el Comercio y la Salud Pública, en la que se afirma el derecho de los países en desarrollo a utilizar al máximo las disposiciones del Acuerdo sobre los Aspectos de los Derechos de Propiedad Intelectual Relacionados con el Comercio respecto a la flexibilidad para proteger la salud pública y, en particular, proporcionar acceso a los medicamentos para todos	3.b.1 - Proporción de la población inmunizada con todas las vacunas incluidas en cada programa nacional	Eficacia	4	0
		3.b.3 - Proporción de centros de salud que disponen de un conjunto básico de medicamentos esenciales asequibles de manera sostenible	Eficacia	0	1
ODS 4 - Educación de Calidad	4.1 - De aquí a 2030, asegurar que todas las niñas y todos los niños terminen la enseñanza	4.1.1 - Proporción de niños, niñas y adolescentes: a) en los grados 2/3; b) al final	Eficacia	0	11

<p>primaria y secundaria, que ha de ser gratuita, equitativa y de calidad y producir resultados de aprendizaje pertinentes y efectivos</p>	<p>de la enseñanza primaria; y c) al final de la enseñanza secundaria inferior, que han alcanzado al menos un nivel mínimo de competencia en i) lectura y ii) matemáticas, desglosada por sexo</p>			
<p>4.2 - De aquí a 2030, asegurar que todas las niñas y todos los niños tengan acceso a servicios de atención y desarrollo en la primera infancia y educación preescolar de calidad, a fin de que estén preparados para la enseñanza primaria</p>	<p>4.2.1 - Proporción de niños menores de 5 años cuyo desarrollo se encuentra bien encauzado en cuanto a la salud, el aprendizaje y el bienestar psicosocial, desglosado por sexo</p>	Eficacia	0	3
	<p>4.2.2 - Tasa de participación en la enseñanza organizada (un año antes de la edad oficial de ingreso en la enseñanza primaria), desglosada por sexo</p>	Eficacia	2	0
<p>4.6 - De aquí a 2030, asegurar que todos los jóvenes y una proporción considerable de los adultos, tanto hombres como mujeres, estén alfabetizados y tengan nociones elementales de aritmética</p>	<p>4.6.1 - Proporción de la población en un grupo de edad determinado que ha alcanzado al menos un nivel fijo de competencia funcional en a) alfabetización y b) nociones elementales de aritmética, desglosada por sexo</p>	Eficacia	0	1
<p>4.a - Construir y adecuar instalaciones educativas que tengan en cuenta las necesidades de los niños y las personas con discapacidad y las diferencias de género, y que ofrezcan entornos de aprendizaje seguros, no violentos, inclusivos y eficaces para todos</p>	<p>4.a.1 - Proporción de escuelas con acceso a a) electricidad, b) Internet con fines pedagógicos, c) computadoras con fines pedagógicos, d) infraestructura y materiales adaptados a los estudiantes con discapacidad, e) suministro básico de agua potable, f) instalaciones de saneamiento básicas separadas por sexo y g) instalaciones básicas para el lavado de manos (según las definiciones de los indicadores WASH)</p>	Eficacia	0	4
<p>4.b - De aquí a 2020, aumentar considerablemente a nivel mundial el número de becas disponibles para los países en desarrollo, en particular</p>	<p>4.b.1 - Volumen de la asistencia oficial para el desarrollo destinada a becas, desglosado por sector y tipo de estudio</p>	Eficacia	1	0

	los países menos adelantados, los pequeños Estados insulares en desarrollo y los países africanos, a fin de que sus estudiantes puedan matricularse en programas de enseñanza superior, incluidos programas de formación profesional y programas técnicos, científicos, de ingeniería y de tecnología de la información y las comunicaciones, de países desarrollados y otros países en desarrollo				
	4.c - De aquí a 2030, aumentar considerablemente la oferta de docentes calificados, incluso mediante la cooperación internacional para la formación de docentes en los países en desarrollo, especialmente los países menos adelantados y los pequeños Estados insulares en desarrollo	4.c.1 - Proporción del profesorado de educación a) preescolar, b) primaria, c) secundaria inferior y d) secundaria superior que ha recibido al menos la mínima formación docente organizada previa al empleo o en el empleo (por ejemplo, formación pedagógica) exigida para impartir enseñanza a cada nivel en un país determinado	Eficacia	2	1
ODS 8 - Trabajo decente y crecimiento económico	8.1 - Mantener el crecimiento económico per cápita desconformidad con las circunstancias nacionales y, en particular, un crecimiento del producto interno bruto de al menos el 7% anual en los países menos adelantados	8.1.1 - Tasa de crecimiento anual del PIB real per cápita.	Eficacia	1	0
	8.2 - Lograr niveles más elevados de productividad económica mediante la diversificación, la modernización tecnológica y la innovación, entre otras cosas centrándose en los sectores con gran valor añadido y un uso intensivo de la mano de obra	8.2.1 - Tasa de crecimiento anual del PIB real por persona empleada	Eficacia	1	0
	8.5 - De aquí a 2030, lograr el empleo pleno y productivo y el trabajo decente para todas las mujeres y los hombres,	8.5.2 - Tasa de desempleo, desglosada por sexo, edad y personas con discapacidad	Eficacia	1	12

	incluidos los jóvenes y las personas con discapacidad, así como la igualdad de remuneración por trabajo de igual valor				
	8.6 - De aquí a 2020, reducir considerablemente la proporción de jóvenes que no están empleados y no cursan estudios ni reciben capacitación	8.6.1 - Proporción de jóvenes (entre 15 y 24 años) que no cursan estudios, no están empleados ni reciben capacitación.	Eficacia	1	8
	8.10 - Fortalecer la capacidad de las instituciones financieras nacionales para fomentar y ampliar el acceso a los servicios bancarios, financieros y de seguros para todos	8.10.1 - a) Número de sucursales de bancos comerciales por cada 100.000 adultos y b) número de cajeros automáticos por cada 100.000 adultos	Eficacia	2	0
Total				32	91

B. Correspondencia entre archivos fuentes y área de preparación

En la siguiente tabla se muestra detalladamente la correspondencia entre los archivos fuentes de datos con el área de preparación.

FUENTES	TABLAS
INDICADORES AGENDA 2030 ODS v7: Hogares Pobres (NBI) (PRINCIPAL) Hogares Pobres No Extremos (NBI) Hogares Pobres Extremos (NBI) Niños 7 a 12 años que no asisten a la escuela Hacinamiento crítico Viviendas inadecuadas Sin servicios básicos Alta dependencia económica Personas en hogares con pobreza extrema (Por Ingreso) (PRINCIPAL) Coeficiente Gini	SA_DASH_INDICADOR_121
INDICADORES AGENDA 2030 ODS v7: Hombres 60 años y más Mujeres 55 años y más Total Adultos Mayores Beneficiarios (pensionados y pensionadas) (PRINCIPAL) Personas con Discapacidad Certificadas	SA_DASH_INDICADOR_131
INDICADORES AGENDA 2030 ODS v7: Total de Hogares (Equipamiento del Hogar) Sin servicios básicos (PRINCIPAL) Con servicios básicos Que poseen nevera Que poseen lavadora Que poseen televisor Que poseen cocinas a gas o eléctrica Que poseen secadora de ropa Que poseen calentador Que poseen aire acondicionado Que poseen filtro de agua Que poseen radio Que poseen horno microondas Que poseen teléfono móvil celular Que poseen televisión por cable Que poseen computadora Que poseen acceso a internet Total de Viviendas (Servicios)	SA_DASH_INDICADOR_141

Abastecimiento de Agua por Acueducto Alumbrado Público Recolección Directa de Basura	
INDICADORES AGENDA 2030 ODS v7: Total Inversión Social (PRINCIPAL) Inversión Social en Educación Inversión Social en Salud Inversión Social en Vivienda Inversión Social en Seguridad Social Inversión Social en Desarrollo Social Inversión Social en Cultura y Comunicación Inversión Social en Ciencia y Tecnología	SA_DASH_INDICADOR_1a2
INDICADORES AGENDA 2030 ODS v7: Esperanza de Vida Mujeres Esperanza de Vida Hombres Tasa de mortalidad materna (PRINCIPAL)	SA_DASH_INDICADOR_311
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad de niños menores de 5 años (PRINCIPAL) Proporción de niños y niñas menores de 5 años con peso inferior al normal Esperanza de Vida Nacional Desnutrición en menores de 5 años según corte internacional	SA_DASH_INDICADOR_321
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad neonatal (PRINCIPAL)	SA_DASH_INDICADOR_322
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad por enfermedad virus de la inmunodeficiencia humana (VIH) (PRINCIPAL)	SA_DASH_INDICADOR_331
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad en tuberculosis Tasa de prevalencia de tuberculosis (PRINCIPAL)	SA_DASH_INDICADOR_332
INDICADORES AGENDA 2030 ODS v7: Tasa de incidencia en la malaria (PRINCIPAL) Tasa de mortalidad por malaria	SA_DASH_INDICADOR_333
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad por diabetes (PRINCIPAL) Tasa de mortalidad por cáncer (PRINCIPAL) Tasa de mortalidad por enfermedades del corazón (PRINCIPAL) Tasa de mortalidad por enfermedades crónicas de las vías respiratorias inferiores (PRINCIPAL) Tasa de mortalidad por enfermedades cerebrovasculares (PRINCIPAL)	SA_DASH_INDICADOR_341
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad por suicidios y homicidios (PRINCIPAL)	SA_DASH_INDICADOR_342
INDICADORES AGENDA 2030 ODS v7: Tasa de mortalidad por accidentes de todo tipo (PRINCIPAL)	SA_DASH_INDICADOR_361
INDICADORES AGENDA 2030 ODS v7: Nacimientos en madres menores de 20 años	SA_DASH_INDICADOR_371
INDICADORES AGENDA 2030 ODS v7: Misión Barrio Adentro I y II: Consultas Misión Barrio Adentro I y II: Exámenes Misión Barrio Adentro I y II: Vidas Salvadas	SA_DASH_INDICADOR_381
INDICADORES AGENDA 2030 ODS v7: BCG: Vacuna contra tuberculosis (PRINCIPAL) Penta: Vacuna contra difteria, tétanos, tosferina, hepatitis tipo B (PRINCIPAL) SRP: Vacuna contra sarampión, rubéola y parotiditis (PRINCIPAL) TD: Vacuna contra tétanos y difteria (PRINCIPAL)	SA_DASH_INDICADOR_3b1
INDICADORES AGENDA 2030 ODS v7: Establecimientos del sistema Público Nacional de salud	SA_DASH_INDICADOR_3b3
INDICADORES AGENDA 2030 ODS v7: Matrícula Total Matrícula Educación Inicial Matrícula Educación Primaria Matrícula Educación Media Tasa Neta de Escolaridad Inicial Tasa Neta de Escolaridad Primaria Tasa Neta de Escolaridad Media	SA_DASH_INDICADOR_411

Tasa Bruta de Escolaridad Inicial Tasa Bruta de Escolaridad Primaria Textos Entregados Tasa Bruta de Escolaridad Media	
INDICADORES AGENDA 2030 ODS v7: Canaima Entregados Tasa Bruta de Escolaridad Inicial Tasa Neta de Escolaridad Inicial	SA_DASH_INDICADOR_421
INDICADORES AGENDA 2030 ODS v7: Tasa Bruta de Escolaridad Inicial (PRINCIPAL) Tasa Neta de Escolaridad Inicial (PRINCIPAL)	SA_DASH_INDICADOR_422
INDICADORES AGENDA 2030 ODS v7: Población Ocupada Analfabeta	SA_DASH_INDICADOR_461
INDICADORES AGENDA 2030 ODS v7: Total Planteles Canaima Entregados Planteles Dependencia Oficial Planteles Dependencia Privada	SA_DASH_INDICADOR_4a1
INDICADORES AGENDA 2030 ODS v7: Becas Estudiantiles (PRINCIPAL)	SA_DASH_INDICADOR_4b1
INDICADORES AGENDA 2030 ODS v7: Docentes en el Sistema Educativo Nacional Docentes Educación Inicial (PRINCIPAL) Docentes Educación Primaria (PRINCIPAL)	SA_DASH_INDICADOR_4c1
INDICADORES AGENDA 2030 ODS v7: Tasa de crecimiento anual del PIB real per cápita. (PRINCIPAL)	SA_DASH_INDICADOR_811
INDICADORES AGENDA 2030 ODS v7: Tasa de crecimiento anual del PIB real por persona empleada (PRINCIPAL)	SA_DASH_INDICADOR_821
INDICADORES AGENDA 2030 ODS v7: Población económicamente Activa (PEA) Población menor de 15 años Población mayor de 64 años Población entre 15 y 64 años Razón de dependencia Población Desocupada Tasa de Desocupación (PRINCIPAL) Población económicamente Activa Hombres (PEA) Población Desocupada de Hombres Tasa de Desocupación Hombres Población económicamente Activa Mujeres (PEA) Población Desocupada de Mujeres Tasa de Desocupación Mujeres	SA_DASH_INDICADOR_852
INDICADORES AGENDA 2030 ODS v7: Población Económicamente Activa de 15 a 24 años Desocupados de 15 a 24 años Tasa de desocupación de 15 a 24 años (PRINCIPAL) Población Económicamente Activa Hombres de 15 a 24 años Desocupados de 15 a 24 años Hombres Tasa de desocupación de 15 a 24 años Hombres Población Económicamente Activa Mujeres de 15 a 24 años Desocupados de 15 a 24 años Mujeres Tasa de desocupación de 15 a 24 años Mujeres	SA_DASH_INDICADOR_861
INDICADORES AGENDA 2030 ODS v7: 8.10.1.a Número de sucursales de bancos comerciales por cada 100.000 adultos. (PRINCIPAL) 8.10.1.b Número de cajeros automáticos por cada 100.000 adultos. (PRINCIPAL)	SA_DASH_INDICADOR_8101

C. Estructura física de las tablas pertenecientes al área de preparación

A continuación se detalla la estructura física de cada una de las tablas pertenecientes al área de preparación.

Correspondientes al ODS 1 – Fin de la Pobreza

Nombre de la Tabla: SA_DASH_INDICADOR_121			
Descripción: Tabla que contiene la información del indicador 1.2.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
porcent_hogar_pobre_nbi	18,2	double precision	Valor del porcentaje de hogares pobres por NBI.
porcent_hogar_pobre_no_extremo_nbi	18,2	double precision	Valor del porcentaje de hogares pobres no extremos por NBI.
porcent_hogar_pobre_extremo_nbi	18,2	double precision	Valor del porcentaje de hogares pobres extremos por NBI.
porcent_nino_7a12_noasistenciaescuela	18,2	double precision	Valor del porcentaje de niños de 7 a 12 años que no asisten a la escuela.
porcent_hacinamiento_critico	18,2	double precision	Valor del porcentaje de hacinamiento crítico.
porcent_vivienda_inadecuada	18,2	double precision	Valor del porcentaje de viviendas inadecuadas.
porcent_sin_servicio_basico	18,2	double precision	Valor del porcentaje de hogares sin servicios básicos.
porcent_alta_dependencia_economica	18,2	double precision	Valor del porcentaje de hogares con personas con alta dependencia económica.
porcent_persona_hogares_pobre_extrema_ingreso	18,2	double precision	Valor del porcentaje de personas en hogares con pobreza extrema por ingreso.
porcent_coeficiente_gini	18,2	double precision	Valor del coeficiente Gini.
fuentes	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_131			
Descripción: Tabla que contiene la información del indicador 1.3.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
cant_hombres_60omas_anio	18,2	double precision	Valor de la cantidad de hombres de 60 o más años.
cant_mujeres_55omas_anio	18,2	double precision	Valor de la cantidad de mujeres de 55 o más años.
cant_total_adultos_mayores	18,2	double precision	Valor de la cantidad total de adultos mayores
cant_poblacion_menor_15anio	18,2	double precision	Valor de la cantidad de la población menor de 15 años.
cant_poblacion_mayor_64anio	18,2	double precision	Valor de la cantidad de la

			población mayor de 64 años.
cant_poblacion_entre_15y64anio	18,2	double precision	Valor de la cantidad de la población entre 15 y 64 años.
porcent_razon_dependencia	18,2	double precision	Valor del porcentaje de razón de dependencia.
cant_beneficiarios_pensionados	18,2	double precision	Valor de la cantidad de beneficiarios pensionados.
cant_persona_discapacidad_certificada	18,2	double precision	Valor de la cantidad de personas con discapacidad.
porcent_beneficiarios_pensionados	18,2	double precision	Valor del porcentaje de beneficiarios pensionados.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_141			
Descripción: Tabla que contiene la información del indicador 1.4.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
cant_total_equipamiento_hogat	18,2	double precision	Valor de la cantidad total de hogares por su equipamiento.
cant_poseen_nevera	18,2	double precision	Valor de la cantidad de hogares que poseen nevera.
cant_poseen_lavadora	18,2	double precision	Valor de la cantidad de hogares que poseen lavadora.
cant_poseen_tv	18,2	double precision	Valor de la cantidad de hogares que poseen televisión.
cant_poseen_cocina_gasoelectrica	18,2	double precision	Valor de la cantidad de hogares que poseen cocina a gas o eléctrica.
cant_poseen_secadora_ropa	18,2	double precision	Valor de la cantidad de hogares que poseen secadora de ropa.
cant_poseen_calentador	18,2	double precision	Valor de la cantidad de hogares que poseen calentador.
cant_poseen_aire_acondicionado	18,2	double precision	Valor de la cantidad de hogares que poseen aire acondicionado.
cant_poseen_filtro_agua	18,2	double precision	Valor de la cantidad de hogares que poseen filtro de agua.
cant_poseen_radio	18,2	double precision	Valor de la cantidad de hogares que poseen radio.
cant_poseen_horno_microondas	18,2	double precision	Valor de la cantidad de hogares que poseen horno microondas.
cant_poseen_tlf_movil	18,2	double precision	Valor de la cantidad de hogares que poseen teléfono móvil.
cant_poseen_tv_cable	18,2	double precision	Valor de la cantidad de hogares que poseen televisión por cable.
cant_poseen_computadora	18,2	double precision	Valor de la cantidad de hogares que poseen computadora.
cant_poseen_acceso_internet	18,2	double precision	Valor de la cantidad de hogares que poseen acceso a internet.
cant_total_servicios	18,2	double precision	Valor de la cantidad total de hogares que poseen servicios.
cant_abastecimienti_agua_acueducto	18,2	double precision	Valor de la cantidad de hogares que poseen abastecimiento de agua por acueducto.
cant_alumbrado_publico	18,2	double precision	Valor de la cantidad de hogares que poseen acceso a alumbrado

			público.
cant_recoleccion_directa_basura	18,2	double precision	Valor de la cantidad de hogares que poseen recolección directa de basura.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_1a2			
Descripción: Tabla que contiene la información del indicador 1.a.2 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_total_inversion_social	18,2	double precision	Valor de la cantidad total de inversión social.
cant_inversion_social_educacion	18,2	double precision	Valor de la cantidad total de inversión social en educación.
cant_inversion_social_salud	18,2	double precision	Valor de la cantidad total de inversión social en salud.
cant_inversion_social_vivienda	18,2	double precision	Valor de la cantidad total de inversión social en vivienda.
cant_inversion_social_seguridad_social	18,2	double precision	Valor de la cantidad total de inversión social en seguridad social.
cant_inversion_social_desarrollo_social	18,2	double precision	Valor de la cantidad total de inversión social en desarrollo social.
cant_inversion_social_cultura_comunicacion	18,2	double precision	Valor de la cantidad total de inversión social en cultura.
cant_inversion_social_ciencia_tecnologia	18,2	double precision	Valor de la cantidad total de inversión social en ciencia y tecnología.
Relaciones:		Campos Clave:	

Correspondientes al ODS 3 – Salud y Bienestar

Nombre de la Tabla: SA_DASH_INDICADOR_311			
Descripción: Tabla que contiene la información del indicador 3.1.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_tasa_mortalidad_materna	18,2	double precision	Valor de la tasa de mortalidad materna.

cant_esperanza_vida_mujer	18,2	double precision	Valor de la cantidad de esperanza de vida de la mujer.
cant_esperanza_vida_hombre	18,2	double precision	Valor de la cantidad de esperanza de vida del hombre.
cant_esperanza_vida_nacional	18,2	double precision	Valor de la cantidad de esperanza de vida nacional.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_321			
Descripción: Tabla que contiene la información del indicador 3.2.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_desnutricion_menor_5anio_corte_inicial	18,2	double precision	Valor de la cantidad de niños menores de 5 años con desnutrición.
cant_proporcion_nino_nina_menor_5_peso_inferior_normal	18,2	double precision	Valor de la cantidad de niños menores de 5 años con peso inferior al normal.
cant_esperanza_vida_nacional	18,2	double precision	Valor de la cantidad de esperanza de vida nacional.
tasa_mortalidad_nino_menor_5anio	18,2	double precision	Valor de la tasa de mortalidad de niños menores de 5 años.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_322			
Descripción: Tabla que contiene la información del indicador 3.2.2 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
tasa_mortalidad_neonatal	18,2	double precision	Valor de la tasa de mortalidad neonatal.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_331			
Descripción: Tabla que contiene la información del indicador 3.3.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de

			medida en que se miden los valores.
tasa_mortalidad_enfermedad_vih	18,2	double precision	Valor de la tasa de mortalidad por VIH.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_332			
Descripción: Tabla que contiene la información del indicador 3.3.2 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
tasa_mortalidad_tuberculosis	18,2	double precision	Valor de la tasa de mortalidad en tuberculosis.
tasa_prevalencia_tuberculosis	18,2	double precision	Valor de la tasa de prevalencia de la tuberculosis.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_333			
Descripción: Tabla que contiene la información del indicador 3.3.3 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
tasa_incidencia_malaria	18,2	double precision	Valor de la tasa de incidencia de la malaria.
tasa_mortalidad_malaria	18,2	double precision	Valor de la tasa de mortalidad de la malaria.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_341			
Descripción: Tabla que contiene la información del indicador 3.4.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
tasa_mortalidad_diabetes	18,2	double precision	Valor de la tasa de mortalidad por diabetes.
tasa_mortalidad_cancer	18,2	double precision	Valor de la tasa de mortalidad por

			cáncer.
tasa_mortalidad_enfermedad_corazon	18,2	double precision	Valor de la tasa de mortalidad por enfermedades del corazón.
tasa_mortalidad_enfermedad_cronica_via_respiratoria,	18,2	double precision	Valor de la tasa de mortalidad por enfermedades crónicas respiratorias.
tasa_mortalidad_enfermedad_cerebrovascular	18,2	double precision	Valor de la tasa de mortalidad por enfermedades cerebrovasculares.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_342			
Descripción: Tabla que contiene la información del indicador 3.4.2 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
tasa_mortalidad_suicidio_homicidio	18,2	double precision	Valor de la tasa de mortalidad por suicidios y homicidios.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_361			
Descripción: Tabla que contiene la información del indicador 3.6.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
tasa_mortalidad_accidente_todo_tipo	18,2	double precision	Valor de la tasa de mortalidad por accidentes de todo tipo.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_371			
Descripción: Tabla que contiene la información del indicador 3.6.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_nacimiento_madre_menor_20año	18,2	double precision	Valor de la cantidad de nacimientos en madres menores de 20 años.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_381			
Descripción: Tabla que contiene la información del indicador 3.8.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_mision_barrio_adentro_i_ii_consulta	18,2	double precision	Valor de la cantidad de consultas realizadas en misión barrio adentro 1 y 2.
cant_mision_barrio_adentro_i_ii_examen	18,2	double precision	Valor de la cantidad de exámenes realizadas en misión barrio adentro 1 y 2.
cant_mision_barrio_adentro_i_ii_vivasalva	18,2	double precision	Valor de la cantidad de vidas salvadas en misión barrio adentro 1 y 2.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_3b1			
Descripción: Tabla que contiene la información del indicador 3.b.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_bcg_vacuna_contra_tuberculosis	18,2	double precision	Valor de la cantidad de vacunas bcg.
cant_penta_vacuna_difteria_tetanos_tosferina	18,2	double precision	Valor de la cantidad de vacunas penta.
cant_srp_vacuna_contra_sarampion	18,2	double precision	Valor de la cantidad de vacunas srp
cant_td_vacuna_contra_tetanos	18,2	double precision	Valor de la cantidad de vacunas td.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_3b3			
Descripción: Tabla que contiene la información del indicador 3.b.3 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_establecimiento_sistema_publico_nacional_salud	18,2	double precision	Valor de la cantidad de establecimientos del sistema público nacional de salud.
Relaciones:		Campos Clave:	

Correspondientes al ODS 4 – Educación de Calidad

Nombre de la Tabla: SA_DASH_INDICADOR_411			
Descripción: Tabla que contiene la información del indicador 4.1.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_matricula_inicial	18,2	double precision	Valor de la cantidad de matrículas de educación inicial.
cant_matricula_primaria	18,2	double precision	Valor de la cantidad de matrículas de educación primaria.
cant_matricula_secundaria	18,2	double precision	Valor de la cantidad de matrículas de educación secundaria.
tasa_escolaridad_inicial_neta	18,2	double precision	Valor de la tasa de escolaridad inicial neta.

tasa_escolaridad_primaria_neta	18,2	double precision	Valor de la tasa de escolaridad primaria neta.
tasa_escolaridad_secundaria_neta	18,2	double precision	Valor de la tasa de escolaridad secundaria neta.
tasa_escolaridad_inicial_bruta	18,2	double precision	Valor de la tasa de escolaridad inicial bruta.
tasa_escolaridad_primaria_bruta	18,2	double precision	Valor de la tasa de escolaridad primaria bruta.
tasa_escolaridad_secundaria_bruta	18,2	double precision	Valor de la tasa de escolaridad secundaria bruta.
cant_matriculas_total	18,2	double precision	Valor de la cantidad total de matrículas.
cant_textos_entre	18,2	double precision	Valor de la cantidad de textos entregados.
	18,2	double precision	Valor de la tasa de mortalidad por enfermedades cerebrovasculares.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_421			
Descripción: Tabla que contiene la información del indicador 4.2.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_canaimas_entr	18,2	double precision	Valor de la cantidad de canaimas entregados
tasa_escolaridad_inicial_bruta	18,2	double precision	Valor de la tasa de escolaridad inicial bruta.
tasa_escolaridad_inicial_neta	18,2	double precision	Valor de la tasa de escolaridad inicial neta.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_422			
Descripción: Tabla que contiene la información del indicador 4.2.2 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
tasa_escolaridad_inicial_bruta	18,2	double precision	Valor de la tasa de escolaridad inicial bruta.
tasa_escolaridad_inicial_neta	18,2	double precision	Valor de la tasa de escolaridad inicial neta.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_461			
Descripción: Tabla que contiene la información del indicador 4.6.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
tasa_pob_analfabeta	18,2	double precision	Valor de la tasa de población analfabeta.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_4a1			
Descripción: Tabla que contiene la información del indicador 4.a.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_total_planteles	18,2	double precision	Valor de la cantidad total de planteles.
cant_canaimas_entr	18,2	double precision	Valor de la cantidad de canaimas entregados.
cant_planteles_oficial	18,2	double precision	Valor de la cantidad de planteles oficiales.
cant_planteles_privado	18,2	double precision	Valor de la cantidad de planteles privados.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_4b1			
Descripción: Tabla que contiene la información del indicador 4.b.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_becas_estudiantiles	18,2	double precision	Valor de la cantidad de becas estudiantiles.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_4c1			
Descripción: Tabla que contiene la información del indicador 4.c.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cant_docentes_educacion_inicial	18,2	double precision	Valor de la cantidad de docentes en educación inicial.
cant_docentes_educacion_primaria	18,2	double precision	Valor de la cantidad de docentes en educación primaria.
cant_docentes_educacion_secundaria	18,2	double precision	Valor de la cantidad de docentes en educación secundaria.
cant_docentes_sistema_nacional	18,2	double precision	Valor de la cantidad de docentes en el sistema nacional.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_811			
Descripción: Tabla que contiene la información del indicador 8.1.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cod_semestre	variable	text	Identificador del semestre en que se obtuvieron los datos.
porcent_crecimiento_pib_capita	18,2	double precision	Valor del porcentaje del crecimiento del PIB per cápita.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_821			
Descripción: Tabla que contiene la información del indicador 8.2.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cod_semestre	variable	text	Identificador del semestre en que se obtuvieron los datos.
porcent_crecimiento_pib_empleada	18,2	double precision	Valor del porcentaje del crecimiento del PIB por persona empleada.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_852			
Descripción: Tabla que contiene la información del indicador 8.5.2 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cod_semestre	variable	text	Identificador del semestre en que se obtuvieron los datos.
cant_pob_menor_15_anios double precision,	18,2	double precision	Valor de la cantidad de la población menor de 15 años.
cant_pob_mayor_64_anios double precision,	18,2	double precision	Valor de la cantidad de la población mayor de 64 años.
cant_pob_entre_15_64_anios double precision,	18,2	double precision	Valor de la cantidad de la población entre 15 y 64 años.
porcent_razon_dependencia double precision,	18,2	double precision	Valor del porcentaje de la razón de dependencia.
cant_poblacion_economicamente_activa double precision,	18,2	double precision	Valor de la cantidad de la población económicamente activa.
cant_poblacion_desocupada double precision,	18,2	double precision	Valor de la cantidad de la población desocupada.
porcent_tasa_desocupacion double precision,	18,2	double precision	Valor de la tasa de la población desocupada.
cant_pea_hombre double precision,	18,2	double precision	Valor de la cantidad de la población económicamente activa hombres.
cant_poblacion_desocupada_hombre double precision,	18,2	double precision	Valor de la cantidad de la población desocupada hombres
porcent_tasa_desocupacion_hombre double precision,	18,2	double precision	Valor de la tasa de desocupación de hombres.
cant_pea_mujer double precision,	18,2	double precision	Valor de la cantidad de la población económicamente activa mujeres.
cant_poblacion_desocupada_mujer double precision,	18,2	double precision	Valor de la cantidad de la población desocupada mujeres.
porcent_tasa_desocupada_mujer	18,2	double precision	Valor de la tasa de desocupación de mujeres.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_861			
Descripción: Tabla que contiene la información del indicador 8.6.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_indicador	variable	text	Identificador del indicador ODS.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
cod_unidad_medida	variable	text	Identificador de la unidad de medida en que se miden los valores.
cod_semestre	variable	text	Identificador del semestre en que se obtuvieron los datos.
cant_activa_15_a_24	18,2	double precision	Valor de la cantidad de la población económicamente activa entre 15 y 24 años.

cant_desocu_15_a_24	18,2	double precision	Valor de la cantidad de la población económicamente desocupada entre 15 y 24 años.
percent_desocu_15_a_24	18,2	double precision	Valor del porcentaje de la población económicamente desocupada entre 15 y 24 años.
cant_activa_15_a_24_hombre	18,2	double precision	Valor de la cantidad de la población económicamente activa de hombres entre 15 y 24 años.
cant_desocu_15_a_24_hombre	18,2	double precision	Valor del porcentaje de la población económicamente desocupada entre 15 y 24 años de hombres.
percent_desocu_15_a_24_hombre	18,2	double precision	Valor del porcentaje de la población económicamente desocupada entre 15 y 24 años de hombres.
cant_activa_15_a_24_mujer	18,2	double precision	Valor de la cantidad de la población económicamente activa de mujeres entre 15 y 24 años.
cant_desocu_15_a_24_mujer	18,2	double precision	Valor del porcentaje de la población económicamente desocupada entre 15 y 24 años de mujeres.
percent_desocu_15_a_24_mujer	18,2	double precision	Valor del porcentaje de la población económicamente desocupada entre 15 y 24 años de mujeres.
Relaciones:		Campos Clave:	

Nombre de la Tabla: SA_DASH_INDICADOR_8101			
Descripción: Tabla que contiene la información del indicador 8.10.1 de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
cod_pais	variable	text	Identificador del país al que están asociados los datos.
cod_anio	variable	text	Identificador del año en que se obtuvieron los datos.
fuelle	variable	text	Identificador de la fuente que origina los datos.
desc_estado	variable	text	Descripción del estado al que están asociados los datos.
desc_municipio	variable	text	Descripción del municipio al que están asociados los datos.
desc_parroquia	variable	text	Descripción de la parroquia al que están asociados los datos.
desc_institucion	variable	text	Descripción de la institución bancaria al que están asociados los datos.
desc_oficina	variable	text	Descripción de la oficina bancaria al que están asociados los datos.
cant_cajeros_atm	18,2	double precision	Valor de la cantidad de cajeros automáticos.
cant_taquillas_externas	18,2	double precision	Valor de la cantidad de taquillas externas.
Relaciones:		Campos Clave:	

D. Estructura física de las tablas pertenecientes al almacén de datos

A continuación se detalla la estructura física de cada una de las tablas pertenecientes al almacén de datos.

Nombre de la Tabla: DIM_FUENTE			
Descripción: Dimensión que contiene la información de todas las fuentes encargadas de medir los datos de los indicadores de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
sk_fuente	32	bigint	Identificador único de la dimensión fuente, clave subrogada.
cod_fuente	variable	text	Clave del negocio de la fuente que suministra los datos.
desc_fuente	variable	text	Descripción de la fuente que suministra los datos.
Relaciones:		Campos Clave: sk_fuente	

Nombre de la Tabla: DIM_TIEMPO_SEMESTRAL			
Descripción: Dimensión que contiene la información de todos los años con sus semestres, desde 1980 hasta el 2050, con jerarquía de Año>Semestre.			
Campo	Tamaño	Tipo de Dato	Descripción
sk_tiempo	32	bigint	Identificador único de la dimensión tiempo semestral, clave subrogada.
anio_cod	32	bigint	Identificador del año.
semestre_cod	variable	text	Identificador en números del semestre del año.
semestre_desc	variable	text	Descripción del semestre del año.
Relaciones:		Campos Clave: sk_tiempo	

Nombre de la Tabla: DIM_INDICADOR			
Descripción: Dimensión que contiene la información de cada uno de los ODS, metas e indicadores de la Agenda 2030.			
Campo	Tamaño	Tipo de Dato	Descripción
sk_indicador	32	bigint	Identificador único de la dimensión indicador, clave subrogada.
cod_ods	variable	text	Identificador del ODS.
nb_ods	variable	text	Nombre del ODS.
desc_ods	variable	text	Descripción del ODS
cod_meta	variable	text	Identificador de la meta.
desc_meta	variable	text	Descripción de la meta.
cod_indicador	variable	text	Identificador del indicador ODS.
desc_indicador	variable	text	Descripción del indicador ODS.
codigo_concatenado	variable	text	Concatenación de los códigos de ODS, meta e indicador.
Relaciones:		Campos Clave: sk_indicador	

Nombre de la Tabla: DIM_UBICACION_GEO			
Descripción: Dimensión que contiene la información de las regiones de Venezuela, con jerarquía de País>Estado>Parroquia>Municipio.			
Campo	Tamaño	Tipo de Dato	Descripción
sk_ubicacion	32	bigint	Identificador único de la dimensión ubicación geo, clave subrogada.
codigo_ubigeo	variable	text	Identificador concatenado de los códigos de país, estado, municipio y parroquia.
cod_pais	variable	text	Identificador del país.
desc_pais	variable	text	Descripción del país.
cod_estado	variable	text	Identificador del estado.
desc_estado	variable	text	Descripción del estado.

cod_municipio	variable	text	Identificador del municipio.
desc_municipio	variable	text	Descripción del municipio.
cod_parroquia	variable	text	Identificador de la parroquia.
desc_parroquia	variable	text	Descripción de la parroquia.
cod_region	variable	text	Identificador de la región.
desc_region	variable	text	Descripción de la región.
Relaciones:		Campos Clave: sk_ubicacion	

Nombre de la Tabla: DIM_DASH_IND_COMPLEMENTARIO			
Descripción: Dimensión que contiene la información de los indicadores complementarios.			
Campo	Tamaño	Tipo de Dato	Descripción
sk_complementario	32	bigint	Identificador único de la dimensión indicador complementario, clave subrogada.
cod_indicador	variable	text	Identificador del indicador ODS.
desc_indicador	variable	text	Descripción del indicador ODS.
desc_indicador_complementario	variable	text	Descripción del indicador complementario.
cod_complementario	variable	text	Identificador del indicador complementario.
Relaciones:		Campos Clave: sk_complementario	

Nombre de la Tabla: DIM_UNIDAD_MEDIDA			
Descripción: Dimensión que contiene la información de las unidades de medida.			
Campo	Tamaño	Tipo de Dato	Descripción
sk_unidad_medida	32	bigint	Identificador único de la dimensión unidad de medida, clave subrogada.
cod_unidad_medida	variable	text	Identificador de la unidad de medida.
desc_unidad_medida	variable	text	Descripción de la unidad de medida.
Relaciones:		Campos Clave: sk_complementario	

Las tablas de hecho presentan una estructura similar, sola cambiando su nombre y clave subrogada por el identificador del indicador, es por eso que la siguiente especificación física aplica para todas las tablas de hecho, cambiando las variables X.X.X por el número del indicador correspondiente.

Nombre de la Tabla: FACT_DASH_INDXXX			
Descripción: Tabla de hecho que contiene toda la información relacionadas a los indicadores principales y secundarios asociados al indicadores ODS X.X.X			
Campo	Tamaño	Tipo de Dato	Descripción
sk_fact_dash_indXXX	32	bigint	Identificador propio de la fact, clave subrogada.
sk_tiempo	32	bigint	Clave primaria de la dimensión tiempo semestral.
sk_complementario	32	bigint	Clave primaria de la dimensión indicador complementario
sk_indicador	32	bigint	Clave primaria de la dimensión indicador
sk_ubicacion	32	bigint	Clave primaria de la dimensión ubicación geo
sk_unidad_medida	32	bigint	Clave primaria de la dimensión unidad de medida
sk_fuente	32	bigint	Clave primaria de la dimensión fuente
valor	18,2	double precision	Medida correspondiente al indicador
cantidad	32	int	Medida cuyo valor siempre es 1, usada para contabilizar la cantidad de registros
Relaciones: DIM_TIEMPO_SEMESTRAL (sk_tiempo) DIM_DASH_IND_COMPLEMENTARIO (sk_complementario) DIM_INDICADOR (sk_indicador) DIM_UBICACION (sk_ubicacion) DIM_UNIDAD_MEDIDA (sk_unidad_medida) DIM_FUENTE (sk_fuente)			Campos Clave: sk_tiempo sk_complementario sk_indicador sk_ubicacion sk_unidad_medida sk_fuente sk_fact_dash_indXXX