

Modelo Dimensional para procesos académicos, investigativos y de seguimiento a  
graduados en Instituciones de Educación Superior públicas

David Antonio Fuentes Vargas

Trabajo de Grado para Optar al Título de Magíster en Ingeniería de Sistemas e Informática

Director

Luis Carlos Gómez Flórez

MSc. en Ingeniería de Sistemas e Informática

Codirector

Martha Eliana Mendoza Becerra

PhD. en Ingeniería de Sistemas y Computación

Universidad Industrial de Santander

Facultad de Ingenierías Fisiomecánicas

Escuela de Ingeniería de Sistemas e Informática

Maestría en Ingeniería de Sistemas e Informática

Bucaramanga

2023

**Dedicatoria**

A mis padres, amigos y familia por todo su apoyo incondicional en este proceso, a pesar de que muchas veces pensé que no podría lograrlo estas personas lograron mantenerme enfocado y me regalaron la confianza necesaria para lograr esta meta.

**Agradecimientos**

Quiero agradecer a mis directores de tesis por todo su apoyo y asesoría, conté con mucha suerte, siento que sin ellos no hubiera podido lograrlo. Agradecer a todas las personas que me regalaron su tiempo y aceptaron participar en algunas de las fases del proyecto: la profesora Piedad Arenas, la profesora Martha Vitalia, la ingeniera Maritza, el profesor Alexander Bustamante, el Ingeniero Diego Quiñonez, la profesora Sonia Gamboa y la profesora Sandra Villamizar.

**Tabla de Contenido**

Resumen..... 8

Abstract..... 9

Capítulo 1 ..... 10

    Introducción ..... 10

    1.1. Planteamiento del problema ..... 11

    1.2. Objetivos ..... 13

        1.2.1. Objetivo General ..... 13

        1.2.2. Objetivos Específicos ..... 13

Capítulo 2 ..... 14

    2.1. Contexto Teórico y Estado del Arte ..... 14

        2.1.1. Acreditación Institucional ..... 14

        2.1.2. Bodegas de Datos (DW) ..... 15

        2.1.3. Revisión del Estado del Arte ..... 20

Capítulo 3 ..... 31

    3.1 Metodología ..... 31

        3.1.1. Etapa 1: Observación ..... 31

        3.1.2. Etapa 2: Identificación del problema ..... 31

        3.1.3. Etapa 3: Desarrollo de la solución ..... 32

        3.1.4. Etapa 4: Prueba de la solución ..... 46

Capítulo 4 ..... 46

    4.1. Modelado dimensional ..... 46

4.1.1.	Matriz bus .....	47
4.1.2.	Dimensiones .....	49
4.1.3.	Modelos dimensionales .....	51
4.1.4.	Casos de diseño utilizados en los modelos dimensionales .....	83
4.2.	Validación del modelo dimensional .....	87
4.2.3.	Planificación del grupo focal .....	88
4.2.4.	Diseño del grupo focal .....	89
4.2.5.	Construcción de la sesión de grupo focal .....	90
4.2.6.	Análisis de los datos y reporte de resultado .....	91
Capítulo 5	.....	97
5.1.	Prototipo para procesos académicos .....	97
5.1.1.	Desarrollo del Back Room .....	97
5.1.2.	Desarrollo del Front Room .....	103
5.1.3.	Integración .....	111
5.1.4.	Despliegue .....	112
5.2.	Evaluación de la utilidad del prototipo .....	112
5.2.1.	Evaluación del prototipo .....	112
Conclusiones	.....	117
Referencias	.....	119

## Lista de Tablas

Tabla 1. ....	15
Tabla 2. ....	17
Tabla 3. ....	21
Tabla 4. ....	23
Tabla 5. ....	24
Tabla 6. ....	26
Tabla 7. ....	29
Tabla 8. ....	37
Tabla 9. ....	40
Tabla 10. ....	41
Tabla 11. ....	41
Tabla 12. ....	42
Tabla 13. ....	48
Tabla 14. ....	52
Tabla 15. ....	53
Tabla 16. ....	53
Tabla 17. ....	53
Tabla 18. ....	54
Tabla 19. ....	54
Tabla 20. ....	54
Tabla 21. ....	54
Tabla 22. ....	55
Tabla 23. ....	55
Tabla 24. ....	55
Tabla 25. ....	55
Tabla 26. ....	56

Tabla 27. ....	56
Tabla 28. ....	56
Tabla 29. ....	56
Tabla 30. ....	57
Tabla 31. ....	57
Tabla 32. ....	60
Tabla 33. ....	61
Tabla 34. ....	62
Tabla 35. ....	62
Tabla 36. ....	62
Tabla 37. ....	63
Tabla 38. ....	67
Tabla 39. ....	68
Tabla 40. ....	68
Tabla 41. ....	68
Tabla 42. ....	69
Tabla 43. ....	71
Tabla 44. ....	72
Tabla 45. ....	73
Tabla 46. ....	73
Tabla 47. ....	75
Tabla 48. ....	76
Tabla 49. ....	76
Tabla 50. ....	77
Tabla 51. ....	79
Tabla 52. ....	79
Tabla 53. ....	80

Tabla 54. ....	80
Tabla 55. ....	82
Tabla 56. ....	82
Tabla 57. ....	82
Tabla 58. ....	83
Tabla 59. ....	88
Tabla 60. ....	88
Tabla 61. ....	90
Tabla 62. ....	91
Tabla 63. ....	92
Tabla 64. ....	96
Tabla 65. ....	104
Tabla 66. ....	114
Tabla 67. ....	115
Tabla 68. ....	115
Tabla 69. ....	116
Tabla 70. ....	116
Tabla 71. ....	117

## Lista de Figuras

Figura 1.....	17
Figura 2.....	18
Figura 3.....	19
<b>Figura 4.....</b>	<b>27</b>
Figura 5.....	33
Figura 6.....	34
Figura 7.....	35
Figura 8.....	37
Figura 9.....	38
Figura 10.....	39
Figura 11.....	42
Figura 12.....	43
Figura 13.....	44
Figura 14.....	44
Figura 15.....	45
Figura 16.....	46
Figura 17.....	59
Figura 18.....	65
Figura 19.....	71
Figura 20.....	75
Figura 21.....	78
Figura 22.....	81
Figura 23.....	84
Figura 24.....	85
Figura 25.....	86
Figura 26.....	87

Figura 27. ....	91
Figura 28. ....	93
Figura 29. ....	94
Figura 30. ....	94
Figura 31. ....	95
Figura 32. ....	95
Figura 33. ....	98
Figura 34. ....	101
Figura 35. ....	102
Figura 36. ....	103
Figura 37. ....	104
Figura 38. ....	106
Figura 39. ....	108
Figura 40. ....	109
Figura 41. ....	110
Figura 42. ....	110

## Resumen

**Título:** modelo dimensional para procesos académicos, investigativos y de seguimiento a graduados en instituciones de educación superior públicas\*

**Autor:** David Antonio Fuentes Vargas\*\*

**Palabras claves:** modelado dimensional, procesos académicos, procesos investigativos, procesos de seguimiento a graduados, Instituciones de Educación Superior, bodegas de datos, acreditación institucional

**Descripción:** Uno de los más grandes problemas que enfrentan las Instituciones de Educación Superior consiste en lidiar con los grandes volúmenes de datos que generan cada semestre. En el caso particular de Colombia, esta situación se agudiza cuando el Consejo Nacional de Acreditación (CNA) solicita que las instituciones que deseen acreditarse muestren la calidad de su servicio por medio de datos. El presente proyecto propone suplir esta necesidad con un conjunto de modelos dimensionales basados en los factores de acreditación establecidos por el CNA que se relacionan con procesos académicos, investigación y seguimiento a graduados. Para el diseño de los modelos dimensionales de los factores antes mencionados se extraen requerimientos basados en informes de autoevaluación presentados por Instituciones de Educación Superior públicas al CNA, el Acuerdo 03 de 2014 y la opinión de tres expertas en acreditación. Con los requerimientos se realiza el conjunto de modelos dimensionales, que son validados por tres expertos en modelado dimensional por medio de un grupo focal. Finalmente, se implementa el modelo dimensional de matrícula estudiantil correspondiente al área de procesos académicos con base en los datos proporcionados por la Universidad Industrial de Santander, el cual también es validado por el área usuaria por medio de un test de satisfacción.

\* Trabajo de Grado

\*\* Facultad de Ingenierías Fisicomecánicas. Maestría en Ingeniería de Sistemas e Informática. Director: Luis Carlos Gómez Flórez. MSc en Ingeniería de Sistemas e Informática. Codirectora: Martha Eliana Mendoza Becerra. PhD en Ingeniería de Sistemas y Computación

## Abstract

**Title:** dimensional model for academic, research and graduate follow-up process in public higher education institutions\*

**Author:** David Antonio Fuentes Vargas\*\*

**Keywords:** dimensional modeling, academic processes, investigative processes, graduate follow-up processes, Higher Education Institutions, data warehouses, institutional accreditation

**Description:** One of the biggest problems faced by Higher Education Institutions is dealing with the large volumes of data they generate each semester. In the particular case of Colombia, this situation worsens when the National Accreditation Council (CNA) requests that the institutions that wish to be accredited show the quality of their service through data. This project proposes to meet this need with a set of dimensional models based on the accreditation factors established by the CNA that are related to academic processes, research and graduate follow-up. For the design of the dimensional models of the factors before removing, requirements are extracted based on self-assessment reports submitted by public Higher Education Institutions to the CNA, the Accreditation Factors Agreement arranged by the CNA and the opinion of three experts in accreditation. With the requirements, the set of dimensional models is made, which are validated by three experts in dimensional modeling through a focus group. Finally, the student dimensional model corresponding to the area of academic enrollment processes is implemented based on the data provided by the Universidad Industrial de Santander, which is also validated by the user area through a satisfaction test.

\* Degree work

\*\* Faculty of Physical-Mechanical Engineering. Master's Degree in Systems Engineering and Informatics. Director: Luis Carlos Gómez Flórez. MSc in Systems Engineering and Computer Science. Co-director: Martha Eliana Mendoza Becerra. PhD in Systems Engineering and Computing

## Capítulo 1

### Introducción

Las Instituciones de Educación Superior (IES) tienen el deber de mantener un nivel de exigencia que se encuentre acorde al desarrollo del país, y la calidad con la que ofrecen su servicio educativo es el activo con el cual han podido reclamar presencia en las instancias de diálogo con los actores del desarrollo social (ACOFI & Confedi, 2018). Tanta es la importancia del aseguramiento de la calidad educativa que hace parte de los diecisiete objetivos de desarrollo sostenible establecidos por la Organización de las Naciones Unidas (United Nations, 2015). En este sentido, un sector de las IES ha optado por medir la calidad con que ofrecen su servicio participando en el proceso de acreditación en alta calidad establecido por el Consejo Nacional de Acreditación (CNA), entidad creada y reglamentada por la Ley 30 de 1992 y ampliamente reconocida por la comunidad académica y científica como el referente nacional en evaluación de la calidad de la Educación Superior. El proceso de acreditación institucional requiere la presentación de informes de soporte basados en datos, generando un gran desafío para las instituciones ya que deben involucrarse activamente en la gestión de sus datos institucionales. Por otro lado, las Bodegas de Datos (DW) son tecnologías enfocadas en dar soporte a las decisiones estratégicas por medio del uso de datos, este proyecto propone el planteamiento del modelo de una DW que ofrezca soporte a los procesos académicos, investigativos y de seguimiento a graduados de una IES desde el punto de vista de los aspectos a evaluar que utiliza el CNA para otorgar la acreditación institucional en alta calidad.

El modelo propuesto se basa en requerimientos manejados en los informes de autoevaluación institucional presentados por IES públicas ante el CNA para participar en procesos de acreditación, también utilizan el contenido del acuerdo 03 de 2014, donde

figuran los aspectos a evaluar establecidos por el CNA para ponderar la calidad de la institución. Además de las fuentes expuestas, el conjunto de requerimientos también cuenta con la revisión de tres expertas en acreditación que dieron sus observaciones y propusieron requerimientos adicionales.

El conjunto de modelos dimensionales cuenta con la validación de tres expertos en modelado de DW, los cuales formularon sus recomendaciones por medio de un grupo focal. Dentro del proyecto también se realiza la implementación de un prototipo del modelo de matrículas, siendo el más complejo de los relacionados con el área de procesos académicos, la forma de validar este resultado fue por medio de la aplicación de un test de satisfacción de usuario.

Este documento cuenta con cinco capítulos, en el primero se describe la introducción, el planteamiento del problema y los objetivos; el segundo capítulo expone tanto el contexto teórico de la acreditación institucional y las bodegas de datos, como la revisión del estado del arte de las DW para Educación Superior; en el tercer capítulo se expone la metodología de desarrollo de la DW del proyecto; en el cuarto se presenta el modelado dimensional y la validación del mismo con un grupo focal; finalmente, en el último capítulo se implementa un prototipo de matrícula estudiantil con los datos dispuestos por la Universidad Industrial de Santander y su respectiva validación por medio de un test de nivel de satisfacción.

### **1.1. Planteamiento del problema**

Un objetivo típico de las IES consiste en mejorar la calidad con que imparten su servicio. La acumulación permanente de datos en las IES es un recurso que se utiliza para justificar ante el CNA la calidad con que imparten educación. Sin embargo, la realidad actual del manejo de los datos en las IES colombianas presenta diversas

dificultades(Camargo, 2022). La información debe extraerse de bases de datos operacionales de distintos tipos que van dirigidas a usuarios capacitados (Mao, 2010), otra dificultad es que el manejo de los datos se limita a la publicación de informes de gestión ante la comunidad universitaria con herramientas de visualización, cuyo diseño debe volverse a hacer cada año por no usarse servicios de análisis enfocados en almacenamiento. En el caso particular de la Universidad Industrial de Santander se publican los informes anuales a través de la iniciativa denominada UIS en Cifras (2022), donde se muestran a través de Power BI (una plataforma de visualización de datos desarrollada por Microsoft) datos de programas, estudiantes, graduados, profesores y grupos de investigación (UIAES-UIS, 2022). El aspecto para mejorar de esta iniciativa no solo sería incorporar herramientas de análisis y modelado dimensional, también se sugiere tener en cuenta la planificación estratégica en función de mejorar la calidad de los procesos académicos, investigativos y de seguimiento a graduados; la información que muestran puede basarse en algún modelo lo suficientemente sólido como para garantizar una alta calidad, el más adecuado es el definido por el CNA, una entidad que lleva trabajando más de veinte años en ese tema (Ley 30, 1992).

Es momento de exponer la solución propuesta por el presente trabajo. Es claro que surge la necesidad de tecnologías que además de soportar los aspectos de evaluación que exige el CNA para acreditar la alta calidad en procesos académicos, investigativos y de seguimiento a graduados ayuden a las IES públicas a conocer el estado en que se encuentran en términos de calidad. Las DW son sistemas de almacenamiento de información que permitirían a partir de los datos dar un soporte a las decisiones de mejoramiento estratégico tomadas por la IES (FRANKENFIELD, 2022). Según Khan (2012) entre los beneficios que se pueden asociar a una DW se encuentra el soporte estratégico y operacional en la toma de decisiones, proporcionar información oportuna,

facilitar el análisis de datos actuales e históricos, revelar tendencias estratégicas y oportunidades, entre otros. Este sistema está desarrollado específicamente para facilitar las consultas, en contraposición de los sistemas operacionales. En cuanto a la creación anual de dashboards en Power BI puede establecerse un flujo de datos de Extracción, Transformación y Carga (ETL) que permita ir almacenando los datos a través del tiempo utilizando una herramienta más enfocada en la ingeniería de datos, y no tanto en la analítica. Este trabajo propone un modelo dimensional de DW para las IES públicas basado en los factores del CNA relacionados con procesos académicos, investigativos y de seguimiento a egresados; el cual permitirá dar un apoyo a la planificación estratégica de la institución.

## **1.2. Objetivos**

### **1.2.1. Objetivo General**

Proponer un modelo dimensional para Instituciones de Educación Superior utilizando la Metodología de desarrollo de bodega de datos para empresas MiPymes, para soportar la toma de decisiones en procesos académicos, investigativos y de seguimiento a egresados basado en los lineamientos de acreditación institucional del CNA y utilizando bases de datos de la Universidad Industrial de Santander como caso de estudio.

### **1.2.2. Objetivos Específicos**

1. Realizar una revisión Sistemática de la literatura para comprender el estado actual del modelado dimensional aplicado a Instituciones de Educación Superior utilizando la guía de desarrollo de revisiones sistemáticas de la literatura para ingeniería.

2. Definir un modelo dimensional para procesos académicos, investigativos y de seguimiento a egresados que soporte las decisiones estratégicas institucionales con base en los lineamientos de acreditación de calidad del CNA; y validarlo en un grupo focal conformado por expertos en bodegas de datos y administradores de bases de datos siguiendo la norma ISO-25010: Medidas de calidad del producto, definida en la sub-característica adaptabilidad.
3. Construir un prototipo para las facultades de Ingenierías basado en el modelo dimensional propuesto para procesos académicos que permita su visualización por medio de consultas Ad-hoc y de reportes estándar; evaluando su utilidad con los funcionarios UIS responsables de la toma de decisiones en las áreas de interés, por medio de un test de nivel de satisfacción con base en la métrica ISO/IEC 25022: Medidas de calidad de uso, definida para la sub-característica utilidad.

## **Capítulo 2**

### **2.1. Contexto Teórico y Estado del Arte**

#### **2.1.1. Acreditación Institucional**

El Consejo Nacional de Acreditación (CNA) tiene el objetivo fundamental de garantizar a la sociedad que las Instituciones de Educación Superior (IES) cumplen los más altos requisitos de calidad y cumplen con sus propósitos y objetivos (Ley 30, 1992). Las IES que optan por acreditarse en alta calidad deben someterse a un proceso que incluye la presentación de un informe de autoevaluación, que es redactado con base en los aspectos a evaluar definidos por el CNA y publicados en el Acuerdo 03 de 2014 (CESU, 2014).

Los aspectos a evaluar del CNA se agrupan en características, que a su vez están englobadas en factores de evaluación. Los doce factores con que el CNA acredita a las IES se presentan en la Tabla 1. Este trabajo define un modelo dimensional de DW que brinde un soporte a los aspectos a evaluar del CNA que se relacionen con procesos académicos, investigativos y de seguimiento a graduados.

**Tabla 1.**

*Factores del CNA*

<b>Factores de acreditación del CNA (CESU, 2014)</b>	
1. Misión y Proyecto Institucional	7. Pertinencia e impacto social
2. Estudiantes	8. Procesos de autoevaluación y autorregulación
3. Profesores	9. Organización, administración y gestión
4. Procesos Académicos	10. Planta física y recursos de apoyo académico
5. Visibilidad nacional e internacional	11. Bienestar Institucional
6. Investigación y creación artística y cultural	12. Recursos Financieros

De los doce factores, en este estudio solo se tendrán en cuenta los cinco que se relacionan con procesos académicos, investigativos y de seguimiento a egresados, que son: 2. Estudiantes, 3. Profesores, 4. Procesos Académicos, 6. Investigación y creación Artística y Cultural y 7. Pertinencia e impacto social.

Este trabajo se enfoca en estos temas de evaluación de calidad porque son las tres áreas en que se agrupa el uso de DW aplicadas a IES según un estudio publicado en el año 2019 por el autor del proyecto (Fuentes-Vargas et al., 2019).

Para construir el modelo se hará uso de los informes de autoevaluación presentados por IES públicas de Colombia, el Acuerdo 03 de 2014 y la opinión de tres expertas en acreditación institucional.

### **2.1.2. Bodegas de Datos (DW)**

Las Bodegas de Datos (DW) son definidas por Kimball & Ross (2013) como una copia de datos transaccionales estructurados para consultas y análisis; Inmon (2005) es presenta una definición más detallada, las define como una colección de datos, temáticos,

integrados, no volátiles y variantes en el tiempo, que dan soporte a las decisiones de gestión. La característica temática de los datos de una DW se debe a que provienen de sistemas operacionales categorizados en temáticas por la misma organización. Otra característica, y la más importante según (Inmon, 2005), es la integración, debido a que los datos que vienen de múltiples fuentes deben ser convertidos a un único formato. La no volatilidad de los datos en una DW se debe a que los datos se cargan (generalmente, pero no siempre) en masa, y se accede a ellos, pero no se actualizan; la carga se realiza en una instantánea (un conjunto de registros que refleja una vista estática), de modo que cuando se producen cambios posteriores se debe escribir un nuevo registro de instantáneas, guardando un registro histórico de datos que no se puede cambiar, una DW se puede consultar, pero no modificar. La última característica es la varianza en el tiempo, esta implica que cada unidad de datos en la DW es precisa en algún momento de tiempo, en general todas las consultas ejecutadas en una DW tienen algún elemento de tiempo asociado a ellas (Inmon, 2005).

### **2.1.3. Data Mart**

Un data mart es una DW útil para las necesidades de un equipo o unidad de negocios específico. Es de menor tamaño, más especializado y puede incluir resúmenes de datos más útiles para la comunidad de usuarios (Amazon Web Services, 2022). Con este enfoque, los datos analíticos se implementan de forma aislada a nivel de departamento sin preocuparse por compartir e integrar información en toda la organización (Kimball & Ross, 2013).

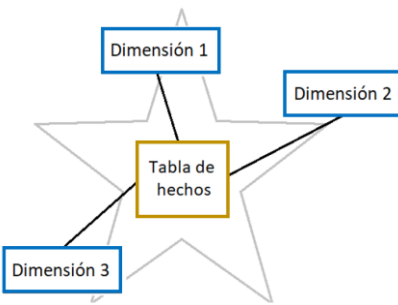
### **2.1.4. Modelado dimensional**

El enfoque dimensional es también algunas veces llamado enfoque de unión de estrellas, Kimball lo categoriza como una técnica para simplificar las bases de datos

(Kimball & Ross, 2013). Aunque las estructuras de bases de datos en tercera forma normal sirven para remover redundancias y son inmensamente útiles en el procesamiento operacional, al utilizar un solo lugar de la base de datos en inserciones y actualizaciones son demasiado complicadas para el análisis estratégico. Los usuarios no pueden entender, navegar y recordar modelos normalizados que parecen inmensas autopistas (Kimball & Ross, 2013). Afortunadamente, el modelado dimensional resuelve el problema de los esquemas demasiado complejos durante la presentación de datos. Kimball destaca dos requisitos por los cuales esta técnica es ampliamente preferida para presentar datos analíticos, estos son, la entrega de datos comprensibles para usuarios comerciales y un rendimiento de consulta rápido. En el proyecto se utilizaron dos de los esquemas de representación dimensional, que son el esquema estrella y el esquema copo de nieve. En la Tabla 2 se explican de forma breve.

**Tabla 2.**

*Esquemas de representación dimensional*

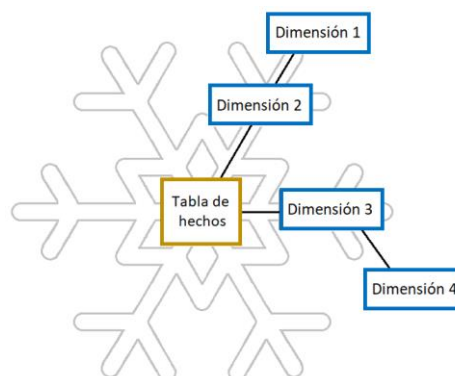
Esquema	Descripción	Figura
Estrella	El esquema en estrella es un enfoque de modelado maduro ampliamente adoptado por las DW relacionales (Microsoft, 2022), reciben este nombre por su semejanza con la estructura de una estrella, como se aprecia en la Figura 1.. A las medidas del entorno de negocio se les denomina hechos, alrededor de los cuales hay un contexto que define las condiciones y en qué momento se registraron, a tal conjunto de características se les denomina dimensiones (Kimball & Ross, 2013).	<p><b>Figura 1.</b></p> <p><i>Esquema Estrella</i></p> 

Copo de  
nieve

Cada tabla de dimensiones puede contener múltiples jerarquías independientes. Mientras en los esquemas en estrella las jerarquías del modelo de datos original se contraen o no se normalizan, en el esquema copo de nieve estas son mostradas explícitamente. Se puede formar a partir de un esquema en estrella expandiendo (normalizando) las jerarquías en cada dimensión, formando el esquema que se aprecia en la Figura 2 (Kimball & Ross, 2013).

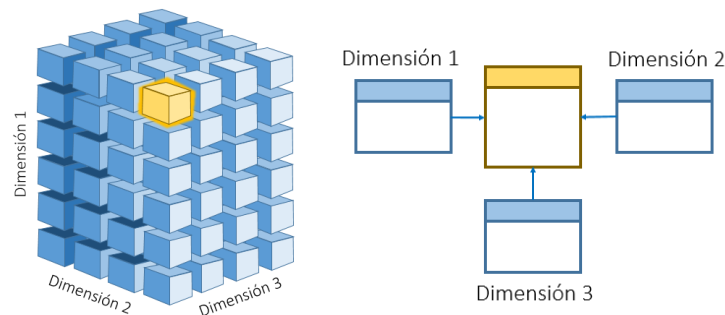
**Figura 2.**

*Esquema copo de nieve*



### 2.1.5. Cubo dimensional

Los modelos dimensionales implementados en entornos de bases de datos multidimensionales se denominan cubos de procesamiento analítico en línea (OLAP), como se ilustra en la Figura 3, donde se ejemplifica una medida (en color amarillo) vista desde la perspectiva de sus dimensiones. Este concepto puede abstraerse para los hipercubos, los cuales no pueden representarse visualmente con facilidad por manejar más de tres dimensiones (Microsoft, 2016). Cuando los datos se cargan en un cubo OLAP, se almacenan e indexan utilizando formatos y técnicas diseñadas para datos dimensionales. Las agregaciones de rendimiento o las tablas de resumen pre-calculadas suelen ser creadas y administradas por el motor de cubo OLAP. En consecuencia, los cubos ofrecen un rendimiento de consulta superior debido a los pre-cálculos, las estrategias de indexación y otras optimizaciones. Los usuarios pueden profundizar agregando o eliminando atributos de sus análisis con un excelente rendimiento sin emitir nuevas consultas. La desventaja es que se pagan estas capacidades con el rendimiento de carga, especialmente para grandes conjuntos de datos (Kimball & Ross, 2013).

**Figura 3.***Cubo OLAP***2.1.6. Proceso de Extracción, Transformación y Carga (ETL)**

El ETL describe el proceso integral por el cual las empresas recogen todos sus datos (estructurados, sin estructurar y gestionados por equipos de cualquier parte del mundo) para ponerlos en un estado que tenga utilidad empresarial (Google, 2022). Este proceso se agrupa en tres componentes principales: la extracción, que consiste en captar los datos del sistema fuente; la transformación, donde se preparan los datos para ser importados en la bodega de datos sometiéndolos a operaciones como reformateo, recálculo, modificación de su estructura, adición de elementos de tiempo, entre otras; y la carga, donde se mueven físicamente los datos extraídos y transformados a la base de datos destino (Khan, 2012).

Luego de presentar el contexto teórico de la acreditación en alta calidad y el modelado dimensional de DW, se hace una revisión sistemática de la literatura que permita comprender el estado actual del modelado dimensional aplicado a procesos académicos, investigativos y de seguimiento a graduados en IES.

### **2.1.3. Revisión del Estado del Arte**

#### **2.1.3.1. Planeación de la Revisión Sistemática de la Literatura (RSL)**

##### **2.1.3.1.1. RSL existentes sobre DW aplicadas a IES**

El proceso inicia seleccionando revisiones publicadas del 2018 al 2022. Para ello, se utilizan las bases de datos: ACM Digital Library, IEEE Xplore, Science Direct, Scopus, Springer Link y Web of Knowledge.

Esta revisión busca tener en cuenta estudios con un impacto investigativo comprobable, para ello se hace la selección de cada revisión examinando si la revista donde se publica esta consignada en el Journal Citations Report (JCR), una base de datos realizada por el Institute for Scientific Information (ISI), la principal fuente de información del mundo en servicios de indexación, revistas y conferencias de investigación (International Scientific Institute, 2018). También son seleccionadas aquellas revisiones pertenecientes a la medida de influencia científica Scimago Journal & Country Rank (SJR), cuyo objetivo es medir el prestigio actual de los estudios en procesos de evaluación investigativa (González-Pereira et al., 2010). Solo tres revisiones sobre DW en IES cumplieron estas condiciones.

Luego del proceso de selección, se asigna un porcentaje de calidad basado en los criterios (1. Reportar criterios de inclusión/exclusión, 2. Búsqueda adecuada, 3. Síntesis de los estudios incluidos, 4. Evaluación de la calidad de los estudios incluidos, 5. Presentación de detalles suficientes de los estudios incluidos) establecidos en la Base de Datos de Resúmenes de Revisiones de Efectos (DARE) (University of York., 2014), otorgando un peso de 1 a cada criterio que se cumple de forma satisfactoria, 0.5 de forma parcial y 0 si no se cumple. Antes de fijar un porcentaje de calidad, se verifica si el estudio

fue publicado en una revista perteneciente al JCR/SJR, si la revista está en el JCR se presenta su Factor de Impacto (FI), en caso de encontrarse en el SJR se muestra tanto su Indicador de Prestigio (IP) como el cuartil (C) más alto alcanzado en el año de publicación, que puede ser Q1 (verde), Q2 (amarillo), Q3 (naranja) o Q4 (rojo). Las métricas FI, IP y C de la Tabla 3 corresponden al año en el cual se publicó el estudio.

Cabe recalcar que el cuarto criterio de la DARE no es tenido en cuenta para el cálculo de calidad de los mapeos (Moscoso-Zea et al., 2018) y (Alkhalil et al., 2021), esto se debe a que la evaluación de la calidad es un paso opcional para los mapeos sistemáticos (Petersen et al., 2008).

**Tabla 3.**

*Revisiones existentes sobre DW en IES*

Año	Título	JCR FI (Clarivate, 2022)	SJR (Scimago Lab, 2022) FI	C	Calidad (%)
2018	A Holistic View of Data Warehousing in Education (Moscoso-Zea et al., 2018)	4,1	0,61	Q1	75
2019	Una Revisión de Bodegas de Datos para Educación Superior (Fuentes-Vargas et al., 2019)	-	0,15	Q4	60
2021	Applying Big Data Analytics in Higher Education: A Systematic Mapping Study (Alkhalil et al., 2021)	-	0,34	Q3	75

La calidad de las revisiones de la Tabla 3 tiene una media de 70%, lo cual indica que se está realizando una búsqueda adecuada, sin embargo, está haciendo falta sintetizar los estudios abordados. Estos hallazgos son tenidos en cuenta en la presente revisión sistemática, buscando no repetir esos mismos errores.

**2.1.3.1.2. Preguntas de investigación**

Para presentar el enfoque de la literatura actual sobre DW en IES y relacionarlo con los factores del CNA, se deberán plantear unas preguntas de investigación que

permitan verificar si los temas abordados en cada estudio se relacionan con uno o más de los tópicos del factor 2, 3, 4, 6 o 7 de la Tabla 1.

Teniendo en cuenta este contexto, se definen las preguntas de investigación con base en la guía propuesta por Kitchenham et al. (2007), que son:

- Pregunta 1: ¿Qué particularidades se están presentando en los estudios sobre el modelado de una bodega de datos para educación superior?
- Pregunta 2: ¿Qué tópicos del CNA se están abordando en investigaciones sobre bodegas de datos para soportar procesos académicos, investigativos y de seguimiento a graduados?

La primera pregunta busca identificar experiencias pasadas que permitan enriquecer el proceso general de realización de una DW para IES. La segunda pregunta es más específica, se enfoca en determinar los tópicos del CNA en los que se están realizando DW con más y menos frecuencia.

Los resultados obtenidos de estas preguntas de investigación permitirán detectar información de valor para el modelado general de una DW para IES y encontrar una relación entre el estado del arte sobre DW para IES y los factores de calidad del CNA relacionados con procesos académicos, investigativos y de seguimiento a graduados.

#### ***2.1.3.1.3. Desarrollo del protocolo de revisión***

Para reducir la posibilidad de sesgo del investigador se debe definir un protocolo de revisión sistemática, el cual consiste en: fijar una cadena de búsqueda, un proceso de búsqueda, los criterios de inclusión/exclusión y evaluar la calidad de los estudios primarios.

Teniendo en cuenta la guía de Kitchenham et al. (2007) se define la cadena de búsqueda presentada en la Tabla 4.

**Tabla 4.***Cadena de búsqueda***Cadena de búsqueda**


---

(dimension\* model\* OR design) AND (business intelligence OR data warehouse\*) AND (educat\* OR high\* educat\* OR academic\*).

---

Se realiza la búsqueda en las mismas seis bases de datos usadas para los estudios de revisión de la Tabla 3. Luego de seleccionar los estudios se aplica la técnica bola de nieve, que consiste en revisar las referencias de cada uno e identificar aún más estudios candidatos. Después de recopilar el conjunto de estudios candidatos, se aplican los criterios de inclusión/exclusión.

El criterio para incluir un estudio candidato fueron los siguientes:

1. El estudio debe haber sido publicado después del 2018.

Los criterios para excluir un estudio candidato fueron los siguientes:

2. El estudio NO es relevante para el desarrollo de Bodegas de Datos (DW) para educación superior.
3. Artículos publicados en revista que NO están indexadas en el Journal Citations Report (JCR) y/o en Scimago Journal & Country Rank (SJR). Este criterio no se aplica para literatura gris.

**2.1.3.2. Resultados de la Revisión Sistemática**

Partiendo de la cadena de búsqueda y aplicando la técnica de bola de nieve se encontraron ciento diecisiete (117) estudios candidatos, de los cuales trece (13) fueron publicados en revistas pertenecientes al JCR/SJR y veinticinco (25) fueron de literatura gris: veintitrés (23) son artículos de conferencias y dos (2) trabajos de grado.

Para medir la calidad de los trece artículos publicados en revistas se tienen en cuenta los siete (7) aspectos presentados en las columnas de la Tabla 5. Las casillas de la segunda, tercera y cuarta columna, se llenan con el mismo procedimiento usado en la

Tabla 3. Los dos aspectos subsecuentes se relacionan con el grado de comprensibilidad del modelo, la columna *Explic. Dimen.* muestra el número de dimensiones del esquema que se explicaron en el estudio y se entendieron (E), por ejemplo, si en los modelos presentados en un estudio se explican cuatro (4) de seis (6) dimensiones en la parte escrita y en los esquemas solo se entienden tres (3), se consigna en la tabla un 4(E=3/6); el caso de las medidas (*Explic. Medidas*) es análogo y se consigna de la misma manera. Otro aspecto a tener en cuenta es el de implementación, la columna *Imp.* muestra si el estudio fue puesto en marcha en una IES; el último factor se relaciona con el porcentaje de calidad (*Calidad (%)*), el cual se calcula a partir de diez de los doce criterios proporcionados por (Kitchenham & Brereton, 2013) y (Dingsøyr & Dybå, 2008), se exceptúa el criterio 2 (identificar el método de investigación) por ser de carácter cualitativo y el criterio 7 (evaluación de la técnica de revisión sistemática) por ser solo relevante para trabajos de revisión; A cada criterio se da un peso de 1 si se cumple de forma satisfactoria, 0.5 de forma parcial y 0 si no se cumple.

**Tabla 5.**

*Calidad de los estudios primarios publicados en revistas presentes en el JCR/SJR*

Año (Referencia)	FI JCR (Clarivate, 2022)	SJR (Scimago Lab, 2022)		Explic. Dimen.	Explic. Medidas	Imp.	Calidad (%)
		P	C				
(Boulila et al., 2018)	2,41	0.40	Q2	0 (E=2/2)	0 (E=5/5)	Si	90
(Scholtz et al., 2018)	1,44	0.54	Q2	Ninguna	Ninguna	Si	95
(Boulila et al., 2018)	2,41	0.40	Q2	0 (E=2/2)	0 (E=5/5)	Si	90
(Scholtz et al., 2018)	1,44	0.54	Q2	Ninguna	Ninguna	Si	95
(Dadkhah & Lagzian, 2018)	-	0.26	Q3	Ninguna	Ninguna	Si	75
(Martins et al., 2018)	3,71	1.21	Q1	Ninguna	Ninguna	No	80
(Medina et al., 2018)	-	0.18	Q2	16 (E=16/16)	10 (E=10/10)	Si	95
(Moscoso-Zea et al., 2019)	3,75	0.78	Q1	Ninguna	Ninguna	No	95

(Joshua & Moge, 2020)	-	0.15	Q4	Ninguna	Ninguna	Si	95
(Yulianto & Kasahara, 2020)	0,17	0.19	Q3	7 (E=7/7)	0 (E=1/2)	Si	95
(Combata Niño et al., 2020)	14,09	2.77	Q1	Ninguna	Ninguna	Si	90
(Berges et al., 2021)	3,37*	0.93	Q1	0 (E=9/9)	0 (E=1/1)	Si	95
(Neamah, 2021)	-	0.21	Q4*	Ninguna	Ninguna	No	90
(Musa et al., 2021)	-	0.20	Q4	Ninguna	Ninguna	No	95
(Najm et al., 2022)	2,69*	0.61*	Q3*	0 (E=5/5)	0 (E=1/1)	Si	90

\*Aún no se ha publicado el valor correspondiente al año, se toma el más reciente

Se observa que más de la mitad (61,5%) de los estudios se encuentran en revistas presentes en el JCR. Respecto al SJR, el 15,4% de las revistas tienen un Indicador de Prestigio mayor a 1 y están distribuidas de forma homogénea entre los cuartiles (un 31% en el primer cuartil y un 23% respectivamente en cada uno de los cuartiles restantes). Estos resultados muestran que la mayoría de las revistas en las cuales se publica este tipo de estudios están presentes en el JCR y manejan un prestigio homogéneo en el SJR. Otro resultado de la Tabla 5 se relaciona con la presentación de las medidas y dimensiones, solo se representan y/o explican en cinco de los estudios primarios. En cuanto a implementaciones el resultado es positivo, el 69% de los estudios primarios aplican en una IES real el modelo que proponen. Por otro lado, aplicando los criterios de (Kitchenham & Brereton, 2013) se encuentra que las falencias más predominantes se relacionan con la omisión del usuario final en el diseño de la DW y una descripción incompleta del contexto en el que se realiza la investigación.

Después de medir la calidad de los estudios primarios publicados en revistas presentes en el JCR/SJR se procede a dar respuesta a las preguntas de investigación.

**Pregunta 1: ¿Qué particularidades se están presentando en los estudios sobre el modelado de una bodega de datos para educación superior?**

Durante la revisión se encontró que trece (13) de los treinta y ocho (38) estudios primarios presentan un enfoque general del diseño de una DW para educación superior, más allá de centrarse en algún proceso particular hacen un análisis holístico de su diseño y aplicación. De los estudios primarios de literatura gris siete (7) siguieron este mismo enfoque general. La Tabla 6 muestra las particularidades que se están presentando en DW con un enfoque general, agrupadas en tres temas: Dificultades en un 25%, Recomendaciones también en un 25% e Innovaciones en un 50%.

**Tabla 6.**

*Particularidades de Literatura sobre DW generales para IES*

Temas	Síntesis
Dificultades	<b>Literatura publicada:</b> (Scholtz et al., 2018) hacen un estudio de las dificultades de la implementación de herramientas de análisis en un país con dificultades en ancho de banda y acceso a internet. (Berges et al., 2021) propusieron un Framework donde descubrieron tres aspectos: se deben considerar un gran número de variables, los datos están segregados en diferentes sistemas de información y el procesamiento de datos a menudo requiere muchos recursos.
	<b>Literatura gris:</b> (Jayashree & Priya, 2020) descubrieron que es escaso el uso de datos agregados (datos combinados de varias mediciones) durante la implementación de DW.
Recomendaciones	<b>Literatura publicada:</b> Los autores de (Martins et al., 2018) destacan la importancia de la definición precisa de los conceptos, características y el alcance de los sistemas de información, donde es clave analizar el grado de satisfacción de la institución y la etapa posterior a la adopción del sistema.
	<b>Literatura gris:</b> (Rahutomo et al., 2018) enfatizan en la importancia del uso de encuestas durante el diseño de una DW, ya sea en la recolección de requerimientos o en la evaluación del grado de satisfacción. (B Akki & M. N., 2018) se enfocan en la importancia de definir un dashboard que logre proporcionar información de valor con un muy bajo nivel de granularidad.
Innovaciones	<b>Literatura publicada:</b> para manejar la gran cantidad de datos que se acumula en las IES (Moscoso-Zea et al., 2019) proponen usar una infraestructura híbrida entre DW y gestión del conocimiento. (Combata Niño et al., 2020) justifican la importancia de incluir dentro del proceso de implementación un centro de competencias en inteligencia de negocios con habilidades en negocios, análisis y tecnologías de información
	<b>Literatura gris:</b> los autores (Somya et al., 2018) recomiendan soluciones de Inteligencia de Negocios orientadas a servicios ya que permiten hacer la Extracción, Transformación y Carga de una forma más simple comparada con el procedimiento tradicional. Otro enfoque similar es el propuesto por (Khan B et al., 2018), donde se define un eDWaaS que significa DW como servicio para aumentar la rentabilidad evitando implementaciones internas o

---

subcontratadas. En cuanto a vías de acceso a una DW (Lapura et al., 2018) opta por usar API REST con la intención de favorecer los testeos de usabilidad. (Sorour et al., 2020) por su parte propone una arquitectura basada en el aseguramiento de calidad educativa.

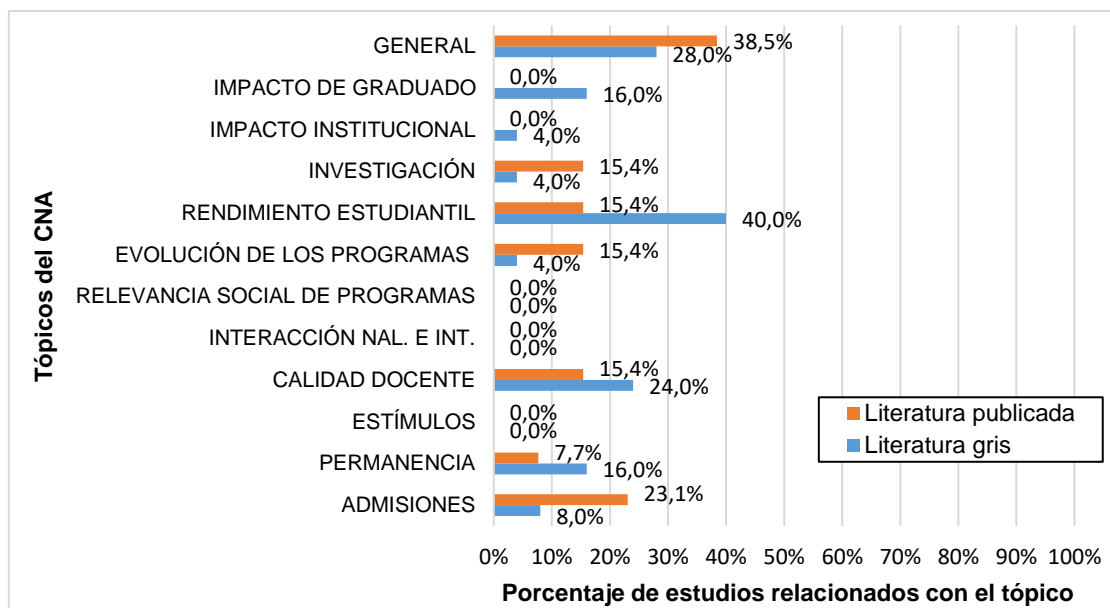
---

**Pregunta 2: ¿Qué tópicos del CNA se están abordando en investigaciones sobre bodegas de datos para soportar procesos académicos, investigativos y de seguimiento a graduados?**

Como se mencionó anteriormente los cinco (5) factores de la Tabla 1 correspondientes a los procesos académicos, investigativos y de seguimiento a graduados son: estudiantes (Factor 2), profesores (3), procesos académicos (4), Investigación y creación artística y cultural (6), Pertinencia e impacto social (7). Estos factores son definidos por el CNA en los once (11) tópicos que se aprecian en la Figura 4.; para los factores 2, 3 y 4 los tópicos son: evolución de los programas, relevancia social de los programas, calidad docente, admisiones, permanencia, estímulos para estudiantes y rendimiento estudiantil; para el factor 6 solo se tuvo en cuenta si el estudio aborda investigación; y para el 7 fueron definidos dos enfoques: impacto social de la institución e impacto social de los graduados.

**Figura 4.**

*Distribución de la literatura*



La Figura 4 muestra que la literatura publicada se está centrando en mostrar el diseño general de una DW para IES, la cual representa el 38,5% de los estudios publicados en revistas del JCR/SRJ y el 28% de la literatura gris, esto quiere decir que tanto en la literatura publicada como en la gris casi la tercera parte de los estudios no se particularizan en un determinado proceso. Sin embargo, ocurre un caso particular con el rendimiento estudiantil, un tópico que se aborda de forma recurrente en la literatura gris, alcanzando un 40%. También se puede apreciar que, aunque la relevancia social de los programas, la interacción nacional e internacional y los estímulos estudiantiles hacen parte de los tópicos evaluados por el CNA, no hay literatura que contribuya al apoyo de toma de decisiones por medio de DW en estos temas, de manera que solo ocho de los once tópicos contenidos en los factores del CNA han sido compatibles con el uso de DW y podrían fortalecerse con este tipo de herramientas de análisis de datos.

Otro aspecto para resaltar es que en comparación con procesos académicos existe poca literatura relacionada con investigación (15,4% de la literatura publicada y el 4% de la literatura gris) y seguimiento a graduados (16% de la literatura gris). Después de

analizar los temas abordados por la literatura se sintetizan los aportes de cada trabajo encontrado, llegando a los resultados presentados en la Tabla 7.

**Tabla 7.**

*Aportes por procesos encontrados en la literatura*

Proceso	Síntesis
Académico	<p><b>Literatura publicada:</b> Boulila et al. (2018) buscan detectar la eliminación repentina de cursos, inasistencias, retrasos en el plan de estudios y bajo rendimiento por medio de DW; Medina et al. (2018) desarrolla una DW de indicadores de productividad docente para conocer la distribución de los docentes y las horas de clase impartidas. Por otro lado, Yulianto &amp; Kasahara (2020) diseñan un modelo enfocado en el análisis de tasas de matrícula. El estudio de Neamah (2021) analiza el aporte de una bodega de datos en la mejora de procesos de admisiones y rendimiento académico. Musa et al. (2021) presentan nueve factores y tres dimensiones que se deben tener en cuenta en Inteligencia de Negocios. Najm et al. (2022) se enfocan en OLAP Mining para predecir el rendimiento de los estudiantes usando un data mart educacional.</p>
	<p><b>Literatura gris:</b> Urrea et al. (2018) se enfocan su DW en el desempeño académico enriqueciendo el modelo con características sociodemográficas y académicas de estudiantes; Jiménez et al. (2018) proponen una DW para la evaluación docente teniendo en cuenta características sociodemográficas y académicas tanto de profesores como estudiantes. Sutedja et al. (2018) afirman que una bodega de datos que apoye la gestión activa de estudiantes es de gran ayuda para la toma de decisiones de la IES, manteniendo un uso positivo de los datos de los estudiantes activos. Venter &amp; Myburgh (2018) proponen el uso de los principios de la investigación de diseño de acción (ADR), que se relacionan con formulación de problema, creación de artefacto, evaluación, reflexión y aprendizaje. Chandra et al. (2018) proponen un proceso donde se abarque el progreso del estudiante y los cursos perdidos. Pérez-Pérez et al. (2018) identificaron para el uso de herramientas y tipos de procesos tres categorías: manejo administrativo, calidad educativa y rendimiento académico. El estudio de Peralta (2018) clasifica y predice la aprobación de los estudiantes en la asignatura de Básica Computación. Un trabajo que también se relaciona con el rendimiento académico es el de Villegas-Ch et al. (2018) donde se busca encontrar grupos de estudiantes con similitudes por medio del uso de Inteligencia de Negocios. Maia et al. (2018) proponen el uso de Web Intelligence, una tecnología basada en cuadros de mando y reportes de rendimiento. Pesonen et al. (2018) hacen un análisis de probabilidad para determinar la retención del estudiante y analizar si lograra graduarse. Ortiz &amp; Hallo (2019) busca utilizar lineamientos de acreditación para el diseño de una DW siguiendo la metodología de Kimball. Urrea Pito et al. (2020) proponen una DW para el análisis del desempeño académico de universidades por medio de la implementación de dos modelos. Joshua &amp; Mogeia (2020) proponen implementar una DW de registro y admisión con métodos ágiles. Serasinghe et al., (2021) proponen un modelo de registro de estudiantes replicable a otras áreas funcionales de la institución. Mulyawan et al. (2022) utilizan lineamientos de rendimiento institucional para definir una DW de actividad estudiantil, calificaciones, cursos, programas y aprendizaje en el aula. En cuanto al rendimiento académico Najm et al. (2022) proponen implementar una serie de data marts que se complementan con minería de datos.</p>

Investigativo	<p><b>Literatura publicada:</b> Dadkhah &amp; Lagzian (2018) muestran el uso de herramientas de análisis de datos para bibliotecarios, solucionando interrogantes relacionados con colaboraciones investigativas, literatura popular de cada país, citas y contribuciones del estado del arte; Medina et al. (2018) presentan una DW de indicadores de productividad científica.</p> <hr/> <p><b>Literatura gris:</b> Mulyawan et al. (2022) utilizan lineamientos de rendimiento institucional para definir una DW de investigación</p>
Seguimiento a graduados	<p><b>Literatura gris:</b> Irfan et al. (2018) plantean una DW en etapas de formulación, agregación y validación. Pesonen et al. (2018) buscan determinar la probabilidad de graduación y el tiempo de graduación permitiendo categorizar a los estudiantes al finalizar sus estudios. Joshua &amp; Moge (2020) proponen implementar una DW de graduados con métodos ágiles. Mulyawan et al. (2022) utilizan lineamientos de rendimiento institucional para definir una DW de seguimiento a graduados.</p>

El resultado final de esta revisión de la literatura permite tener en cuenta las dificultades, recomendaciones e innovaciones que ya se han presentado en anteriores trabajos relacionados con DW para IES y así reconocer el contexto en el cual puede encajar el presente trabajo. Otro resultado es el obtenido de la segunda pregunta de investigación de la revisión, donde se subdivide la literatura restante en procesos académicos (se menciona en el 60,5% de los estudios), investigativos (7,9%), y de seguimiento a graduados (13,2%). Este resultado muestra que la literatura de DW para IES se está enfocando en procesos académicos. Los estudios sobre investigación y el seguimiento a graduados son escasos, situación que brinda la oportunidad de que el presente trabajo sea un aporte a estos temas.

Teniendo presente el contexto actual de la literatura sobre DW para IES relacionadas con procesos académicos investigativos y de seguimiento a egresados, en el siguiente capítulo se empieza da inicio a la construcción del modelo siguiendo la metodología de (Mendoza et al., 2010).

## **Capítulo 3**

### **3.1 Metodología**

Para el desarrollo de este proyecto se usa como guía el Patrón de Investigación Iterativa (PII) propuesto por Pratt (2009), una metodología para proyectos de investigación que involucran soluciones computacionales. Según Pratt (2009) deben definirse ciclos compuestos por etapas, que son: observación, identificación del problema, desarrollo de la solución y prueba de la solución; en cada ciclo debe entregarse un producto.

La etapa de observación consiste en una revisión sistemática que representa la actualidad en DW para Educación Superior, en la identificación del problema se hace una recolección de requerimientos, para la etapa de desarrollo de la solución se establece un conjunto de modelos para procesos académicos, investigativos y de seguimiento a graduados, pero debido al alcance del proyecto se implementa uno de los modelos de *Procesos Académicos* denominado *Matrícula Estudiantil*, para la etapa de prueba de la solución se aplica un test de satisfacción con la métrica ISO/IEC 25022: Medidas de calidad de uso, definida para la sub-característica utilidad. Se detalla a continuación el manejo de cada etapa.

#### **3.1.1. Etapa 1: Observación**

Esta etapa consiste en el planteamiento de una Revisión Sistemática de la Literatura, la cual ha sido presentada en el Capítulo 2.

#### **3.1.2. Etapa 2: Identificación del problema**

Esta etapa consiste en la recolección de requerimientos del modelo dimensional con base en el Acuerdo 03 de 2014 establecido por el CNA como la base de la acreditación

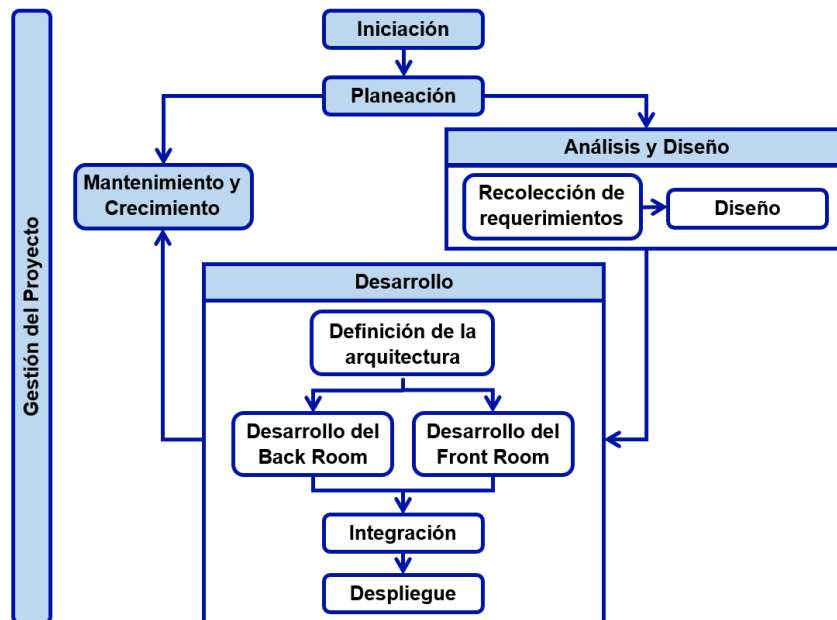
institucional y la opinión de expertos en Acreditación Institucional. Este será explicado con más detalle en el Etapa 3: Desarrollo de la solución durante la Fase de Análisis y Diseño.

### **3.1.3. Etapa 3: Desarrollo de la solución**

Este trabajo desarrolla la DW utilizando la metodología de desarrollo de bodegas de datos para MiPymes (MBD) (Mendoza et al., 2010), la cual se basa en la metodología de Ralph Kimball y una metodología ágil de DW. MBD está diseñada para proyectos conformados por grupos de trabajo pequeños, otras características con las que cuenta son: tener un enfoque iterativo e incremental, alta participación del usuario, actividades detalladas, menor creación de artefactos y responsabilidades concretas para los grupos de trabajo. En la metodología MBD se presentan las seis fases que se proponen en MBD para cada proceso, que son: *Iniciación, Planeación, Análisis y Diseño, Desarrollo, Mantenimiento y Crecimiento, y Gestión del Proyecto*. En la fase de Desarrollo se define la *arquitectura* y de forma iterativa se llevan a cabo las actividades de *desarrollo del Back Room* (donde se encapsula todos los procedimientos y recursos necesarios para llevar a cabo los procesos ETL), *desarrollo del Front Room* (donde se encapsula todos los procedimientos y recursos necesarios para presentar los datos ya tratados al usuario final), e *integración* y finalmente se realiza la fase de *despliegue*. La fase de *Gestión del proyecto* tiene cabida en todo su ciclo de vida.

**Figura 5.**

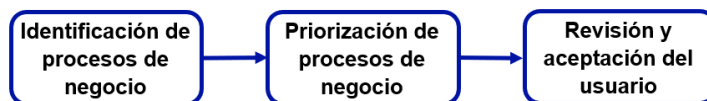
*Ciclo de vida. Adaptado de Mendoza et. al (2010)*



Exceptuando la fase de mantenimiento y crecimiento por estar fuera del alcance del proyecto, se describen a continuación el manejo de cada una de las fases del ciclo de vida de la metodología MBD.

### 3.1.3.1. Fase de Iniciación

Esta fase está conformada por las actividades que se muestran en la Figura 6. Se identifican los procesos de negocio, se identifican los de mayor viabilidad e impacto y al final de la fase se realiza la actividad de revisión y aceptación del usuario.

**Figura 6.***Fase de iniciación*

El tercer objetivo de este trabajo consiste en construir un prototipo para procesos académicos, estos se abordan por el CNA en tres factores: estudiantes, profesores y procesos académicos. El modelo de procesos académicos que se va a implementar es el de matrícula estudiantil, el cual fue seleccionado por ser el que aborda la deserción estudiantil, un tópico en el que a inicios del proyecto la Dirección de Admisiones y Registro Académico de la Universidad Industrial de Santander manifestó tener alto interés. Por esta razón, no se realizaron las actividades y artefactos de la fase de iniciación.

**3.1.3.2. Fase de Planeación**

Esta fase ya fue abordada en la propuesta de investigación que antecede el presente trabajo, donde se tuvo que definir la identidad del proyecto, el alcance, la justificación, elaboración del plan de proyecto, reunión de iniciación y revisión del usuario. La actividad de asignación de roles no es aplicable en este caso debido a que el autor fue el encargado realizar todas las actividades, bajo la orientación de los directores.

**3.1.3.3. Fase de Análisis y Diseño**

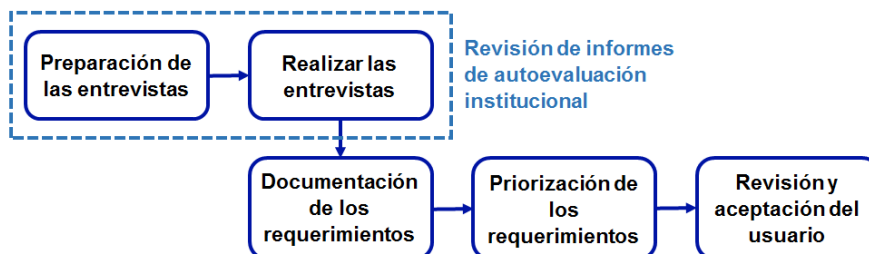
Esta fase se compone de las sub-fases de Recolección de requerimientos y de Diseño.

### 3.1.3.3.1. Recolección de Requerimientos

Tal y como se aprecia en la Figura 7 la metodología propone preparar entrevistas de extracción de requerimientos, aplicarlas, documentar los resultados, hacer una priorización y someterlos a la revisión del usuario. En este caso, se hace inviable manejar entrevistas para una necesidad a nivel nacional, como lo es el soporte al proceso de acreditación institucional de IES públicas. Por tal razón, la forma de definir los requerimientos fue haciendo una revisión de dieciocho informes de autoevaluación presentados por IES públicas en sus procesos de acreditación institucional ante el CNA, lo cual también se señala en la Figura 7. Se busca identificar las necesidades analíticas de las IES en procesos académicos, investigativos y de seguimiento a graduados, las cuales se abordan al presentar el informe de autoevaluación para acreditarse en alta calidad.

**Figura 7.**

*Sub-fase de recolección de requerimientos*



A partir de la revisión de los informes de autoevaluación institucional de las IES públicas que participaron en procesos de acreditación, se realiza la tercera actividad de la Figura 7 que consiste en obtener un artefacto denominado *Documento Inicial de Requerimientos*. En este documento se realiza la actividad de *Priorización de los Requerimientos* a partir de cantidad de veces en que se menciona cada requerimiento en los informes de autoevaluación: si se menciona en más de cinco informes es denominado *Requerimiento más frecuente*; si esta entre cinco y dos informes será un *Requerimiento*

*menos frecuente* y si a pesar de no estar en los informes se ve viable usarlo para cumplir los aspectos a evaluar del acuerdo 03 de 2014 pertenecerá a la categoría de *Requerimiento Propuesto*.

La actividad de *Priorización de Requerimientos* también constituye la base para otorgar al modelo dimensional propuesto la característica de adaptabilidad definida en la norma ISO 25010. Esto se debe a que según la norma ISO 25010 el modelo dimensional tiene esta característica si puede adaptarse de manera efectiva y eficiente a entornos de uso diferentes o en evolución (ISO/IEC, 2011). En este sentido, el hecho de priorizar requerimientos a través de las tres categorías antes explicadas permite definir un modelo adaptable tanto a las IES que decidan implementar una DW con los *Requerimientos más frecuentes* como a aquellas que opten por también tener en cuenta los *Requerimientos menos frecuentes y propuestos*. La decisión de la institución dependerá de sus disposiciones internas y la disponibilidad de los datos.

La última actividad de la sub-fase de Recolección de Requerimientos consiste en la *Revisión y aceptación del usuario*. Luego de tener el *Documento Inicial de Requerimientos* con una categorización en *más frecuentes, menos frecuentes y propuestos*, se debe validar que el resultado sea concordante con los procesos llevados a cabo por el CNA, para ello se acudió a tres expertas en acreditación institucional cuya trayectoria se resume en la tabla.

**Tabla 8.**

*Trayectoria de expertas en acreditación institucional*

<b>ID</b>	<b>Ocupación</b>	<b>Estudios</b>	<b>Experiencia</b>
PAC1	Coordinación de Evaluación de la Calidad Académica – Universidad Industrial de Santander.	PhD en Ingeniería de Telecomunicaciones.	Cinco (5) años de experiencia como consejera del CNA a nivel nacional.
PAC2	Coordinación de Evaluación de la Calidad Académica – Universidad Industrial de Santander.	Ing. de Sistemas.	Más de diez (10) años de experiencia en Evaluación de Calidad Educativa.
PAC3	Docente Planta - Universidad Industrial de Santander.	PhD en Gestión de la Innovación Tecnológica.	Más de diez (10) años de experiencia en acreditación de programas e institucional.

Luego de la consulta a las expertas en acreditación se realizó un nuevo artefacto llamado Documento con los Requerimientos Finales, el cual se encuentra en el anexo.

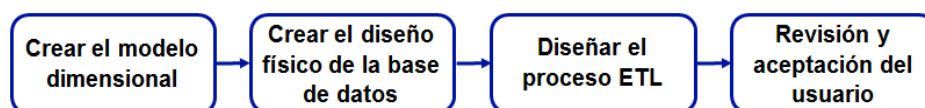
En este documento se tienen en cuenta los ajustes sugeridos y se añaden nuevos requerimientos sugeridos, que han sido ubicados en la categoría de *Requerimientos propuestos*.

**3.1.3.3.2. Diseño**

Para esta sub-fase la metodología establece las actividades presentadas en la Figura 8, que son: crear el modelo dimensional, crear el diseño físico de la base de datos, diseñar el proceso ETL y la revisión y aceptación del usuario.

**Figura 8.**

*Sub-fase de diseño*



La primera actividad consiste en crear el modelo dimensional, proceso que se abordará al detalle en el Capítulo 4. Las subsecuentes tres actividades de la Figura 8 se relacionan con la implementación del prototipo y serán explicadas en el Capítulo 5.

### 3.1.3.4. Fase de Desarrollo

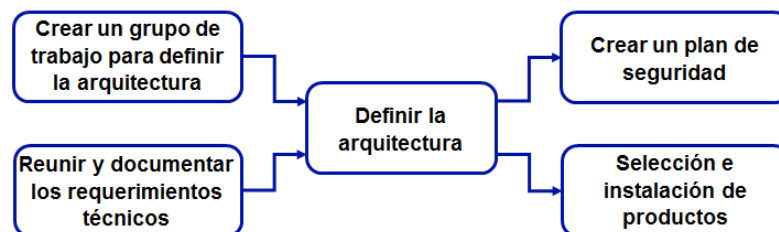
Esta fase se compone de las sub-fases presentadas en la Figura 5, que son: *Definición de la arquitectura, Desarrollo del Back Room, Desarrollo del Front Room, Integración y Despliegue*. Cada sub-fase será descrita a continuación.

#### 3.1.3.4.1. Definición de la Arquitectura

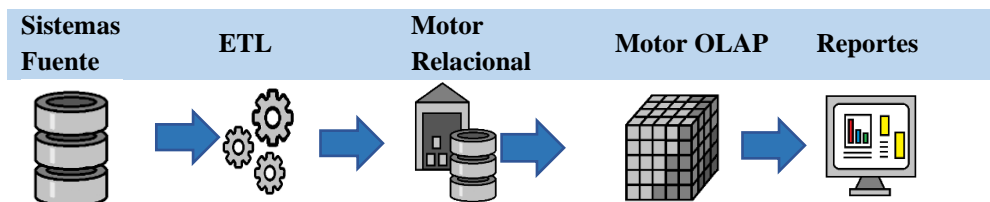
Se presenta en la Figura 9 las actividades que esta sub-fase, que son: *Crear un grupo de trabajo para definir la arquitectura, Reunir y documentar los requerimientos técnicos, Definir la arquitectura, Selección e instalación de productos y Crear un plan de seguridad*.

**Figura 9.**

*Sub-fase de definición de arquitectura*



La *Creación de un grupo de trabajo* no es aplicable en este trabajo por estar siendo realizado por un único autor. La actividad de *Reunir y documentar los requerimientos técnicos* es realizada a partir de los requerimientos obtenidos en la sección 3.1.3.3.1. Teniendo en cuenta estos requerimientos se *Define la arquitectura*, que es presentada en la Figura 10.

**Figura 10.***Arquitectura del prototipo de DW*

La arquitectura mostrada en la Figura 10 se subdivide en los siguientes elementos:

- **Sistemas fuente:** son las fuentes de almacenamiento de los datos de transacciones del proceso de negocio.
- **Área de preparación de datos:** proceso de Extracción, Transformación y Carga (ETL) de los datos hacia la DW.
- **DW relacional:** consolidado de datos centralizados que integra los datos provenientes de los diferentes sistemas fuente
- **Cubo OLAP:** estructura de almacenamiento que permite el análisis de información desde todas las perspectivas del proceso de negocio.
- **Visualización:** es el conjunto de herramientas de front-end que permiten la visualización de datos por medio de reportes o consultas Ad-hoc que acceden al cubo creado.

La actividad de *Selección e Instalación de productos* que se menciona en la Figura 9 se lleva a cabo para cada componente mencionado en la Figura 10, se parte de una evaluación de las características: costo, documentación, funcionalidades generales, características técnicas y algunas características específicas por componente. Para llevar a cabo la comparación se seleccionaron las herramientas disponibles en el mercado, se descargaron, instalaron y exploraron. Para conocer con más detalle el funcionamiento y

características de cada herramienta también se hizo una revisión de la documentación. La columna *Peso* proporciona el porcentaje de importancia se le da a la característica a partir de las necesidades del proyecto, cada herramienta tiene las columnas *Puntaje* y *Total*, que corresponden a la calificación y la calificación ponderada (multiplicar el *Puntaje* por el *Peso*) respectivamente. Las herramientas con mayor valor en la suma de valores de la columna *Total* fueron las elegidas para implementar el prototipo.

En la Tabla 9 se presenta la comparación de las herramientas ETL, donde *SQL Server Integration Services* obtiene el mayor puntaje. También se tiene en cuenta el análisis hecho por (Camilo & Mera, 2022) de estas mismas herramientas.

**Tabla 9.**

*Matriz de comparación de herramientas ETL*

Característica	Peso	SQL Server Integration Services Enterprise Edition		Pentaho Data Integration		Talend for Data Integration		Infosphere Datastage – IBM	
		Puntaje	Total	Puntaje	Total	Puntaje	Total	Puntaje	Total
Costo	0,20	90,0	8,0	30,0	6,0	30,0	6,0	20,0	4,0
Documentación	0,15	100,0	5,0	40,0	6,0	50,0	7,5	80,0	12,0
Características Técnicas	0,15	99,0	4,9	100,0	15,0	98,0	14,7	80,0	12,0
Funcionalidades	0,15	100,0	5,0	100,0	15,0	95,0	14,3	100,0	15,0
Componentes de Transformación	0,35	100,0	5,0	98,0	34,3	99,0	34,7	100,0	35,0
<b>Puntaje Total</b>	<b>1,00</b>		<b>7,9</b>		<b>76,3</b>		<b>77,2</b>		<b>78,0</b>

En la Tabla 10 se muestra la comparación de las herramientas de motor relacional, donde *SQL Server* es la de mayor puntaje. También se tiene en cuenta el análisis hecho por (Camilo & Mera, 2022) de estas mismas herramientas.

**Tabla 10.**

*Matriz de comparación de herramientas de motor relacional*

Característica	Peso	SQL Server		Oracle Database		Mysql		Posgresql	
		Puntaje	Total	Puntaje	Total	Puntaje	Total	Puntaje	Total
Costo	0,10	90,0	9,0	60,0	6,0	70,0	7,0	70,0	7,0
Documentación	0,20	100,0	20,0	100,0	20,0	100,0	20,0	100,0	20,0
Características Técnicas	0,20	99,8	9,9	100,0	20,0	93,0	18,6	100,0	20,0
Funcionalidad	0,25	100,0	5,0	100,0	25,0	100,0	25,0	100,0	25,0
Seguridad	0,25	100,0	5,0	100,0	25,0	93,0	23,3	90,0	22,5
<b>Puntaje Total</b>	<b>1,00</b>		<b>98,9</b>		<b>96,0</b>		<b>93,9</b>		<b>94,5</b>

En la Tabla 11 se muestra la comparación de las herramientas OLAP, donde *SQL Analysis Server* es la de mayor puntaje. También se tiene en cuenta el análisis hecho por (Camilo & Mera, 2022) de estas mismas herramientas.

**Tabla 11.**

*Matriz de comparación de herramientas OLAP*

Característica	Peso	SQL Server Analysis Services Enterprise Edition		Pentaho Schema Workbench Community		Iccube Enterprise		JASPERSOFT	
		Puntaje	Total	Puntaje	Total	Puntaje	Total	Puntaje	Total
Costo	0,20	90,0	18,0	100,0	20,0	90,0	18,0	90,0	18,0
Documentación	0,15	100,0	15,0	90,0	13,5	68,0	10,2	69,6	10,4
Características Técnicas	0,15	93,0	13,9	95,0	14,25	90,0	13,5	69,5	10,4
Funcionalidad	0,20	95,0	19,0	92,0	18,4	83,2	16,6	65,4	13,1
Técnicas de Diseño	0,30	100,0	30,0	94,0	28,2	88,0	26,4	80,0	24,0
<b>Puntaje Total</b>	<b>1,00</b>		<b>95,9</b>		<b>94,4</b>		<b>84,7</b>		<b>75,9</b>

En cuanto a la comparación de herramientas de visualización de la Tabla 12, donde *SQL Server Reporting Services* es la de mayor puntaje. También se tiene en cuenta el análisis hecho por (Camilo & Mera, 2022) de estas mismas herramientas.

**Tabla 12.**

*Matriz de comparación de herramientas de visualización*

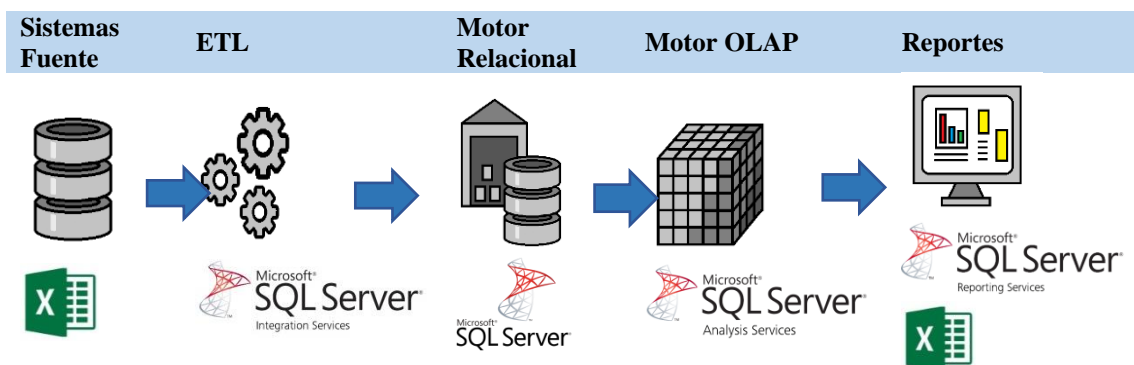
Característica	Peso	SQL Server Reporting Services		Jaspersoft		Pentaho Reporting		SQL Power Wabit	
		Puntaje	Total	Puntaje	Total	Puntaje	Total	Puntaje	Total
Costo	0,15	80,0	12,0	80,0	12,0	80,0	12,0	100,0	15,0
Documentación	0,15	100,0	15,0	88,8	13,3	66,2	9,9	58,4	8,8
Características Técnicas	0,15	75,0	11,3	100,0	15,0	100,0	15,0	45,0	6,8
Funcionalidades	0,20	100,0	20,0	82,0	16,4	98,6	19,7	63,5	12,7
Gráficos	0,20	100,0	20,0	100,0	20,0	100,0	20,0	100,0	20,0
Conexiones	0,15	100,0	15,0	92,0	13,8	100,0	15,0	92,0	13,8
<b>Puntaje Total</b>	<b>1,00</b>		<b>93,3</b>		<b>90,5</b>		<b>91,7</b>		<b>77,1</b>

Las consultas Ad-hoc se manejan con la herramienta Excel, ya que es compatible con *SQL Server Analysis Services* para acceder al cubo OLAP. Otra ventaja de Excel es ser una herramienta ampliamente conocida por funcionarios institucionales.

Se presenta en la Figura 11 la arquitectura de alto nivel del prototipo de DW para matrícula estudiantil incluyendo las herramientas seleccionadas de las comparaciones realizadas anteriormente.

**Figura 11.**

*Arquitectura del prototipo de matrícula estudiantil*



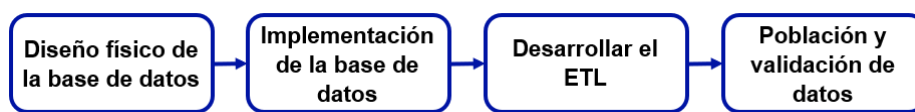
**Desarrollo del Back Room**

En esta sub-fase se realiza el diseño y ejecución del proceso de extracción, transformación y carga (ETL) de la Figura 11, esta subdividida en cuatro actividades que

se presentan en la Figura 12 y serán profundizadas más adelante en la sección 5.1 cuando se realice la implementación del prototipo, estas actividades son: *Diseño físico de la base de datos, Implementación de la base de datos, Desarrollar el ETL, Población y validación de datos.*

### Figura 12.

*Sub-fase de desarrollo del Back Room*

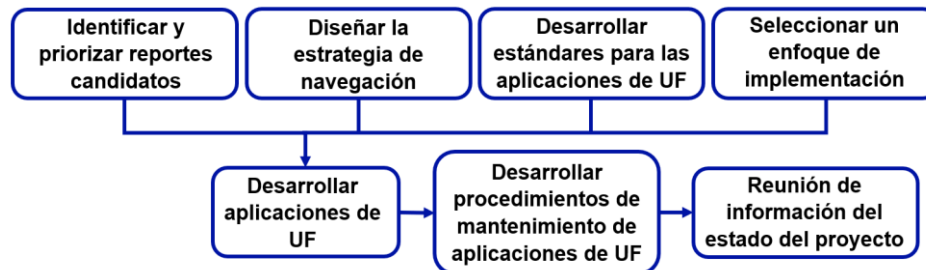


#### 3.1.3.4.2. Desarrollo del Front Room

En esta sub-fase se realiza el prototipo de la sección 5.1 con los reportes solicitados en los requerimientos, se presenta en la Figura 13 y se divide en las actividades: *Identificar y realizar reportes candidatos, Diseñar la estrategia de navegación, Desarrollar estándares para las aplicaciones de Usuario Final, Seleccionar un enfoque de implementación, Desarrollar las aplicaciones de Usuario Final, Desarrollar procedimientos de mantenimiento de aplicaciones de Usuario Final y Desarrollar documentación de las aplicaciones.*

**Figura 13.**

*Sufase de desarrollo de Front Room*

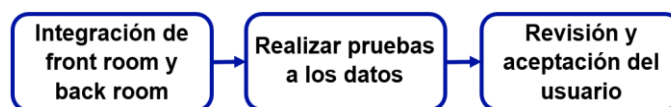


**3.1.3.4.3. Integración**

En esta sub-fase se integra el *Back Room* y *Front Room* del prototipo. Se presenta en la Figura 14 y se compone de las siguientes actividades: integración de front room y back room, realizar pruebas a los datos y revisión y aceptación del usuario. Tal y como se mencionó en los objetivos del proyecto durante la etapa de aceptación del usuario del prototipo se aplicará un test de nivel de satisfacción con base en la métrica ISO/IEC 25022: Medidas de calidad de uso, definida para la sub-característica utilidad.

**Figura 14.**

*Sub-fase de integración*



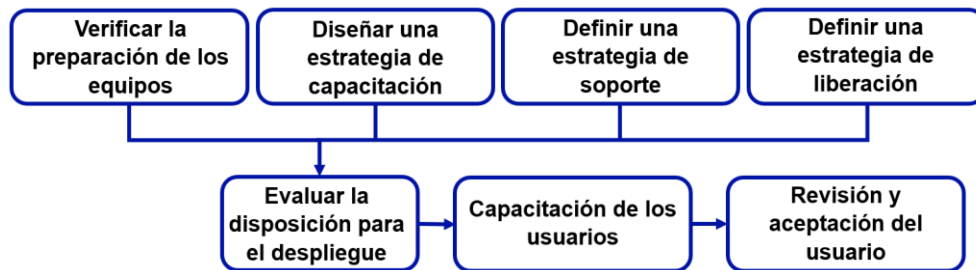
**3.1.3.4.4. Despliegue**

Esta última parte de la Fase de Desarrollo es la que pone la DW a disposición de los usuarios. Está compuesta por las actividades de la Figura 15, que son: verificar la preparación de los equipos para llevar a cabo la instalación, diseñar una estrategia de capacitación de usuarios, definir una estrategia de soporte a usuarios, definir una estrategia de liberación, evaluar la disposición para el despliegue, capacitación de los

usuarios y revisión y aceptación del usuario. Este trabajo se despliega en software on-premise utilizando un dispositivo local de la Universidad Industrial de Santander.

**Figura 15.**

*Sub-fase de despliegue*

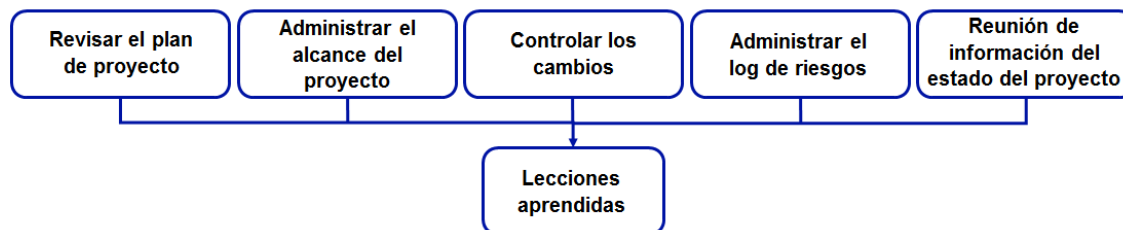


**3.1.3.5. Fase de Mantenimiento y Crecimiento**

Esta fase se sale del alcance del proyecto, razón por la cual no se describen sus actividades y artefactos.

**3.1.3.6. Fase de Gestión del Proyecto**

Para esta sub-fase la metodología establece las actividades presentadas en la Figura 16, que son: *revisar el plan de proyecto, administrar el alcance del proyecto, controlar los cambios, administrar el log de riesgos y hacer reuniones de información del estado del proyecto*. Este conjunto de actividades proporciona las *lecciones aprendidas*.

**Figura 16.***Fase de gestión del proyecto*

Esta fase se mantiene a lo largo de todo el proyecto por medio de un seguimiento semanal del proyecto. La metodología afirma que las actividades de administrar el alcance del proyecto y administrar el log de cambios son opcionales, esto se debe a que en un proyecto pequeño y con un alcance bien definido, es posible que no se soliciten cambios que tengan que registrarse. Al finalizar el proyecto se llevó a cabo una reunión en la que se documentaron las lecciones aprendidas, que serán presentadas en la sección.

#### **3.1.4. Etapa 4: Prueba de la solución**

El prototipo será evaluado con un funcionario UIS responsable de la toma de decisiones en Matrícula Estudiantil, por medio de un test de nivel de satisfacción con base en la métrica ISO/IEC 25022: Medidas de calidad de uso, definida para la subcaracterística utilidad.

## **Capítulo 4**

### **4.1. Modelado dimensional**

En esta sección se presenta la descripción de los modelos dimensionales obtenidos para procesos académicos, investigativos y de seguimiento a graduados, basados en los requerimientos extraídos en la *Recolección de Requerimientos*. El primer paso consiste en presentar la matriz bus de los procesos, donde se podrán identificar las dimensiones

conformadas. Posteriormente será presentado cada esquema estrella con las medidas, el detalle de sus dimensiones y la identificación de los casos de diseño utilizados.

#### **4.1.1. Matriz bus**

La matriz bus es una herramienta de planificación que da una perspectiva general del modelo dimensional propuesto, mostrando todas las dimensiones conformadas que son compartidas entre los diagramas estrella (Kimball & Ross, 2013). Esta matriz se presenta en la tabla con los quince modelos dimensionales del proyecto, los cuales se ubican en las filas. En las columnas se presentan las dimensiones utilizadas, la utilidad de la matriz se centra en representar con una X las relaciones entre dimensiones y modelos, señalando en azul oscuro las dimensiones conformadas, que son aquellas que son usadas por más de un modelo.



#### 4.1.2. Dimensiones

Los procesos académicos, investigativos y de seguimiento a graduados que fueron abordados en los modelos dimensionales de la matriz bus poseen dimensiones conformadas y no conformadas, las cuales son detalladas a continuación.

##### *Dimensiones conformadas:*

- *Periodo:* periodo académico, calendario o de cohorte.
- *Programa:* programa académico de la institución educativa.
- *Novedad Programa:* mini-dimensión con cambios en el programa relacionados con el rediseño, la acreditación, entre otros.
- *Facultad:* facultad de la institución.
- *Matrícula estudiantil:* características de la matrícula del estudiante.
- *Estudiante:* persona que recibe educación de la institución.
- *Profesor:* profesional de la educación dedicado a la docencia.
- *Dedicación:* vinculación, jornada y término de la dedicación docente.
- *Localización:* niveles de organización territorial del país.
- *Movilidad:* proceso que permite realizar actividades en instituciones nacionales o extranjeras.
- *Persona:* miembro de la comunidad académica de la institución.
- *Indicador Persona:* características que pueden cambiar con el tiempo en un miembro de la comunidad académica de la institución.
- *Grupo de Investigación:* conjunto de personas que interactúan para investigar y generar productos de conocimiento.

- *Indicador Grupo de Investigación:* características que pueden cambiar con el tiempo en un grupo de investigación.
- *Proyecto de Investigación:* idea planteada por estudiantes y profesores para dar respuesta a alguna laguna del conocimiento.
- *Fecha:* día, mes y año en que ocurre alguna acción.

***Dimensiones no conformadas:***

- *Demografía:* características sociales del estudiante.
- *Apoyo estudiantil:* beneficio otorgado al estudiante por parte de la institución.
- *Deserción:* características del abandono del programa por parte de un estudiante.
- *Idioma:* se refiere a las características de idioma.
- *Indicador Profesor:* características que pueden cambiar con el tiempo en un profesor.
- *Profesor Visitante:* profesional de la educación dedicado a la docencia y perteneciente a otra institución.
- *Indicador Profesor Visitante:* características que pueden cambiar con el tiempo en un profesor visitante.
- *Categoría:* categoría otorgada al profesor.
- *Premios:* premios otorgados a profesor.
- *Evento:* encuentro en el que se lleva a cabo un ambiente de discusión.
- *Actividad Académica:* operaciones que se realizan dentro del marco de la enseñanza.
- *Curso:* capacitación dirigida a impartir un conocimiento.

- *Semillero*: espacios extracurriculares de formación en los que participan estudiantes bajo el acompañamiento de profesores.
- *Laboratorio*: espacio provisto de aparatos y utensilios adecuados para realizar experimentos científicos.
- *Indicador Laboratorio*: características que pueden cambiar con el tiempo en un laboratorio.
- *Producto*: resultado del trabajo académico realizado por los profesores-investigadores en algún área de conocimiento.
- *Indicador Producto*: características que pueden cambiar con el tiempo en un producto de investigación.
- *Entidad*: entidad financiadora de proyectos de investigación.
- *Participación*: modalidades de participación del graduado con la institución.
- *Graduado*: persona que ya se ha graduado de la institución.
- *Indicador Graduado*: características que pueden cambiar con el tiempo en un graduado.
- *Pregunta*: se detalla los atributos de las preguntas de las encuestas.
- *Respuesta*: se detalla los atributos de las respuestas de las encuestas.
- *Encuesta*: características de la encuesta realizada para acreditación.

#### **4.1.3. Modelos dimensionales**

En esta sección se presentan los quince modelos dimensionales propuestos para dar soporte a la acreditación institucional en procesos académicos, investigativos y de seguimiento a graduados, catorce de los modelos se centran en estos temas y uno está

enfocado en el manejo de encuestas hechas por la institución para la acreditación institucional. La distribución de los modelos para el cumplimiento de los requerimientos se presenta en la Tabla 14, los modelos *Matrícula Estudiantil*, *Actividad Profesoral*, y *Participación Cursos* cumplen los requerimientos de más de un Factor del CNA.

**Tabla 14.**

*Distribución de los modelos dimensionales*

Proceso	Factor del CNA	Modelos dimensionales usados para el Factor del CNA	
Académico	<b>Factor 2. Estudiantes</b>	Matrícula Estudiante	Vinculación
	<b>Factor 3. Profesores</b>	Actividad Profesoral Actividad Profesor Visitante	<i>Participación Cursos (también usado en Factor 4)</i>
	<b>Factor 4. Procesos Académicos</b>	Ambientes Discusión Actividades Académicas	Participación Cursos <i>Matrícula Estudiante (también usado en Factor 2)</i>

**Nota:** Los modelos Vinculación, Profesor Visitante, Ambientes Discusión y Actividades Académicas serán detallados en los anexos.

Investigativo	<b>Factor 6. Investigación y creación artística y cultural</b>	Participación Semillero <i>Matrícula Estudiante (también usado en Factor 2)</i>	Participación Grupos Inv. Producción Intelectual <i>Actividad Profesoral (también usado en Factor 3)</i>
		Participación Proyectos Uso Laboratorios	Financiación Proyectos

**Nota:** Los modelos Participación Semillero, Participación Grupos Investigación, Participación Proyectos, Uso Laboratorios y Financiación de Proyectos serán detallados en los anexos

Seguimiento a Egresados	<b>Factor 7. Pertinencia e impacto social</b>	Graduación
Encuestas		Encuesta

**Nota:** El modelo de Encuestas será detallados en los anexos

Cada uno de los modelos mencionados en la Tabla 13 y Tabla 14 tiene la característica de adaptabilidad que fue mencionada en la sección Recolección de Requerimientos, permitiendo a la institución escoger si quiere implementar un modelo con los requerimientos más utilizados en informes de autoevaluación institucional o si tiene la capacidad de utilizar el modelo que también tiene en cuenta los requerimientos

menos frecuentes y los propuestos por el autor y expertos en acreditación. Los modelos dimensionales serán detallados uno por uno, pero antes es necesario presentar las dimensiones conformadas, esto se debe a que puede ser redundante nombrarlas constantemente en la exposición de cada modelo.

**Dimensiones Conformadas**

Antes de detallar cada modelo dimensional, se presentará a continuación el conjunto de dimensiones conformadas con las jerarquías correspondientes, las cuales serán mencionadas en cada uno de los modelos donde fueron utilizadas.

**Tabla 15.**

*Dimensión Periodo*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Periodo</b>	Periodo Completo	Parte de un año académico	2021-1, 2021-2, 2022-1...
	Año	Año Académico	2021, 2022, ...
	Periodo	Periodo de seis meses dentro del año	1, 2
	<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Jerarquía Periodo	Jerarquía compuesta por: año, periodo	2021 -> 1

**Tabla 16.**

*Dimensión Programa*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Programa</b>	Sigla	Sigla del programa académico	E3T
	Nombre	Nombre del programa académico	Ingeniería Electrónica
	Nivel programa	Nivel del programa académico	Pregrado, Maestría, ...
	Modalidad	Modalidad del programa académico	Presencial, Virtual
	Sede	Sede del programa académico	Socorro, Cúcuta...

**Tabla 17.**

*Dimensión Novedad Programa*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Novedad Programa</b>	Estado Programa	Situación en la que se encuentra el programa	Nuevo, Rediseñado, ...
	Registro Calificado	Obtención de registro Calificado por parte del programa	Primer RC, Con RC, ...
	Acreditación	Obtención de Acreditación por parte del programa	Sin Acred, En proceso, ...

**Tabla 18.**

*Dimensión Facultad*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Facultad</b>	Sigla	Sigla la Facultad	FIFMe
	Nombre	Nombre de la Facultad	Facultad de Ingenierías Físico-mecánicas

**Tabla 19.**

*Dimensión Matrícula Estudiantil*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Matrícula Estudiantil</b>	Cuartil Saber PRO inglés	Cuartil de resultado en inglés en examen Saber PRO	Q1, Q2, ...
	Tipo inclusión	Tipo de inclusión de población vulnerable	especial, regular
	Matrícula especial	Nombre de población vulnerable	Indígena, Municipio difícil acceso, ...
	Estado Matrícula	Nombre del estado en que se encuentra la matrícula del estudiante	Inscrito, admitido, matriculado1, graduado, joven investigador
	<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Jerarquía Matrícula	Jerarquía compuesta por: Tipo inclusión, matrícula especial	especial -> desplazado

**Tabla 20.**

*Dimensión Estudiante*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Estudiante</b>	Identificación	Identificación del estudiante	1098667788, ...
	Nombre	Nombre del estudiante	David Antonio Fuentes Vargas, ...
	Género	Género del estudiante	M, F, Otro

**Tabla 21.**

*Dimensión Profesor*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Profesor</b>	Identificación	Identificación del profesor	1987654368, ...
	Nombre	Nombre del profesor	Pedro Díaz Jiménez, ...
	Género	Género del profesor	M, F, Otro

**Tabla 22.**

*Dimensión Dedicación*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Dedicación</b>	Vinculación	Vinculación de la dedicación del profesor	Planta, Ocasional, Cátedra
	Jornada	Jornada de la dedicación del profesor	Tiempo Completo, Medio Tiempo
	Término	Término de la dedicación del profesor	Indefinido, Fijo

**Tabla 23.**

*Dimensión Localización*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Localización</b>	Continente	Nombre del continente de localización	América, ...
	País	Nombre del país de localización	Colombia, ...
	Departamento	Nombre del departamento de localización	Santander, ...
	Municipio	Nombre del municipio de localización	Bucaramanga, ...
	<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Jerarquía	Jerarquía compuesta por: Continente, país, departamento, municipio	Europa -> España -> Andalucía -> Sevilla

**Tabla 24.**

*Dimensión Movilidad*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Movilidad</b>	Rango Duración	Rango de Duración de la movilidad	< 1 semana, 1-2 semanas, 2semanas-1mes, ...
	Alcance	Grado de cobertura de la movilidad	Nacional, Internacional
	Área de conocimiento	Área de conocimiento	Ciencias de la Salud, ...

**Tabla 25.**

*Dimensión Persona*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Persona</b>	Identificación	Identificación de la persona	1987654368, ...
	Nombre	Nombre de la persona	Pedro Díaz Jiménez, ...
	Género	Género de la persona	M, F, Otro

**Tabla 26.**

*Dimensión Indicador Persona*

<b>Indicador Persona</b>	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Rol	Rol de la persona	Estudiante, Profesor, ...
	Clasificación Minciencias	Clasificación de Minciencias	No aplica, Emérito, Senior, ...
	Nivel de formación	Nivel de formación	Pregrado, maestría, ...

**Tabla 27.**

*Dimensión Grupo de Investigación*

<b>Grupo Investigación</b>	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Sigla	Sigla del Grupo de Investigación	STI, GTI, ...
	Nombre	Nombre del Grupo de Investigación	Grupo en Sistemas y Tecnologías de la Información, ...

**Tabla 28.**

*Dimensión Indicador Grupo de Investigación*

<b>Indicador Grupo Investigación</b>	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Categoría Minciencias	Categoría del Grupo en Minciencias	A1, A, B, C, sin categoría
	Pertenece Red de inv nal	Pertenencia del Grupo a Alguna Red Nacional	Pertenece, No pertenece
	Pertenece Red de inv intl	Pertenencia del Grupo a Alguna Red Internacional	Pertenece, No pertenece

**Tabla 29.**

*Dimensión Proyecto de Investigación*

<b>Proyecto Investigación</b>	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Nombre	Nombre del Proyecto de Investigación	Modelo Dimensional aplicado a Consultorios Jurídicos, ...
	Área de conocimiento	Área de conocimiento	Ciencias de la Salud, ...
	Financiación	Existencia de la financiación	Financiado, No financiado

**Tabla 30.**

*Dimensión Fecha*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Fecha</b>	Fecha completa	Día, mes y año	20/05/2022, ...
	Año	Año que pertenece a la fecha	2022, ...
	Mes	Mes que pertenece a la fecha	05, ...
	Día	Día que pertenece a la fecha	20, ...
	Jerarquía	Jerarquía compuesta por: año,	2022 -> junio
	Fecha	mes, día	-> 22

**1. Procesos académicos: Modelo Dimensional Matrícula Estudiantil**

El modelo *Matrícula Estudiante* de la Figura 17 es diseñado a partir de los requerimientos que se presentan en la Tabla 31, los cuales fueron obtenidos en la sección Recolección de Requerimientos.

**Tabla 31.**

*Requerimientos extraídos del Factor Estudiantes*

<b>PROCESOS ACADÉMICOS: Requerimientos del Factor de Estudiantes</b>
<p style="text-align: center;"><b>Requerimientos más frecuentes</b></p> <ol style="list-style-type: none"> <li>1. Número de estudiantes inscritos, admitidos y de primera matrícula por estrato, programa, nivel de programa, sede, modalidad, periodo y año.</li> <li>2. Índice de selectividad y absorción por género, programa, nivel de programa, sede, periodo y año.</li> <li>3. Número de indígenas inscritos por sede, periodo y año.</li> <li>4. Porcentaje de deserción por motivo, genero, facultad, nivel de programa, sede, modalidad, periodo, cohorte y año.</li> <li>5. Número de estudiantes desertores por motivo, facultad, nivel de programa, sede, periodo y cohorte.</li> <li>6. Número de estudiantes con matrícula de honor, becas por deporte y cultura, becados por estrato, facultad, nivel de programa, sede, modalidad, periodo y año.</li> </ol> <p style="text-align: center;"><b>Requerimientos menos frecuentes</b></p> <ol style="list-style-type: none"> <li>7. Número de estudiantes afrocolombianos, provenientes de municipios de difícil acceso, reinsertados, desplazados y discapacitados por estrato, programa, sede, periodo y año.</li> <li>8. Porcentaje de estudiantes por estrato y periodo.</li> <li>9. Porcentaje de retención por periodo y cohorte.</li> <li>10. Número de graduados por facultad, nivel de programa y periodo.</li> <li>11. Número de estudiantes matriculados por facultad, nivel de programa, sede, periodo y cohorte.</li> </ol>
<b>Requerimientos propuestos</b>

- 
12. Número de estudiantes de población negra, palenquera y raizal (archipiélago de San Andrés, Providencia y Santa Catalina), departamentos donde no hay IES, víctimas o desmovilizados del conflicto armado por facultad y año.
  13. Porcentaje de estudiantes de población negra, afrocolombianos, palenquera y raizal (archipiélago de San Andrés, Providencia y Santa Catalina), departamentos donde no hay IES, provenientes de municipios de difícil acceso, víctimas o desmovilizados del conflicto armado por facultad y año.
  14. Porcentaje de estudiantes con estímulos académicos o apoyos de bienestar por año.
  15. Número de estudiantes con estímulos académicos (internos: asistencia de investigación, matrícula de honor, estudiante distinguido, crédito condonable, monitor, summa cum laude, cum laude, excelencia SABER PRO; externos: Generación E-Excelencia) y apoyos de bienestar (internos: estímulos de alimentación, residencia, subsidios en pago de matrícula; externos: Generación E-Equidad) por estrato, facultad, nivel de programa, sede, periodo y año.
- 

El modelo dimensional de la Figura 17 es adaptable a las decisiones institucionales y la disponibilidad de los datos al priorizar los requerimientos, esto se representa gráficamente asignando los colores negro y gris para los atributos y medidas del modelo dimensional, se señala en color negro los elementos correspondientes a los requerimientos más frecuentes y en color gris los que se crearon a partir de requerimientos menos frecuentes o propuestos.

En la Figura 17 también se hace una señalización en color verde para describir la forma en que se cumplen los requerimientos, tomando como ejemplo el primer requerimiento más frecuente de la Tabla 31, que dice:

*Número de estudiantes inscritos, admitidos y de primera matrícula por estrato, programa, nivel de programa, sede, modalidad, periodo y año.*

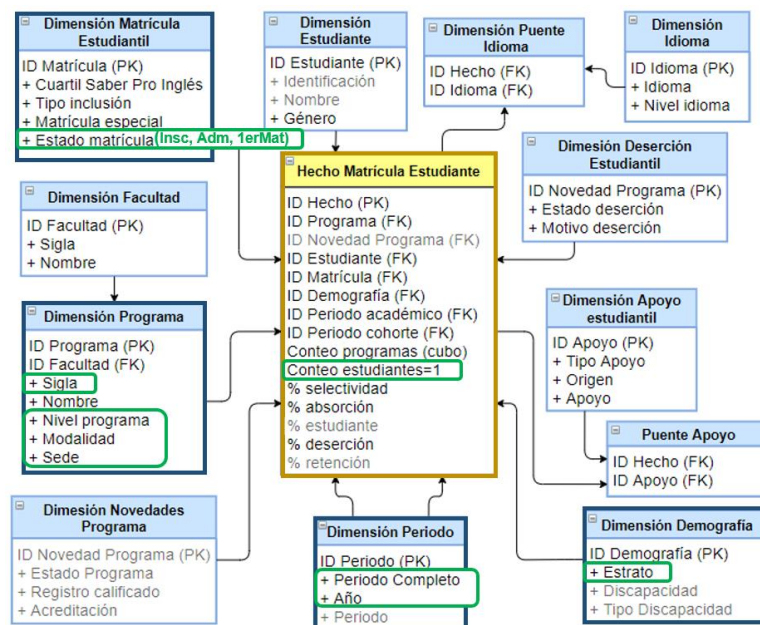
El número de estudiantes se representa con la medida conteo estudiante presente en la tabla de hechos, para tener en cuenta el estado de la matrícula del estudiante se usa el atributo *Estado matrícula* en la *Dimensión Matrícula Estudiantil* señalando si se trata de la matrícula de un inscrito, admitido o de primera matrícula. Para las restantes dimensiones de la consulta, que son *estrato, programa, nivel de programa, sede,*

modalidad, periodo y año, se utilizan las dimensiones *Demografía*, *Programa* y *Periodo* como se señala en la Figura 17.

Cada uno de los requerimientos de la Tabla 31 se ha representado en el modelo *Matrícula Estudiantil*, teniendo en cuenta la característica de adaptabilidad que fue explicada anteriormente.

**Figura 17.**

*Modelo dimensional de matrícula estudiantil*



**Lista de medidas**

El esquema estrella de *Matrícula Estudiantil* de la Figura 17 presenta las dos medidas básicas y cinco medidas derivadas extraídas de la Tabla 31.

Las medidas básicas son:

- *Conteo programas*: cantidad de programas distintos en una consulta

- *Conteo estudiantes:* cantidad de estudiantes que realizan un proceso de matrícula.

Las medidas derivadas son:

- *Porcentaje selectividad:* razón entre los estudiantes inscritos que terminan siendo admitidos.
- *Porcentaje absorción:* razón entre los estudiantes que se matriculan por primera vez y los admitidos.
- *Porcentaje estudiantes:* razón entre los estudiantes con algún atributo particular respecto al total de estudiantes.
- *Porcentaje deserción:* razón entre los estudiantes desertores y el total de matriculados.
- *Porcentaje retención:* razón entre los estudiantes que no han desertado y el total de estudiantes.

La Tabla 32 presenta un resumen de las medidas del modelo dimensional de la Figura 17.

**Tabla 32.**

*Resumen de medidas del Modelo Dimensional de Matrícula Estudiantil*

<b>Medida</b>	<b>Función de Agregación</b>	<b>Visibilidad</b>
Conteo programas	Conteo distinto de programas	Visible
Conteo estudiantes	Conteo de estudiantes	Visible
<b>Medida Derivada</b>	<b>Función de Cálculo</b>	<b>Visibilidad</b>
Porcentaje selectividad	$\%selectividad = \frac{conteo\ admitidos}{conteo\ inscritos} * 100$	Visible

Porcentaje absorción	$\%absorción = \frac{\text{conteo primer matrícula}}{\text{conteo admitidos} + \text{conteo primer matrícula}} * 100$	Visible
Porcentaje estudiantes	$\%estudiantes = \frac{\text{conteo por atributo de dimensión}}{\text{conteo total}} * 100$	Visible
Porcentaje deserción	$\%deserción = \frac{\text{conteo desertores}}{\text{conteo desertores} + \text{conteo matriculados}} * 100$	Visible
Porcentaje retención	$\%retención = 1 - \%deserción$	Visible

**Detalle de dimensiones**

Las consultas de *Matrícula Estudiantil* de este modelo tienen un contexto descrito por diez dimensiones, de las cuales serán presentadas a continuación las cuatro no conformadas con sus atributos, una breve descripción y un ejemplo de los datos almacenados, también son mencionadas las jerarquías definidas. Las restantes dimensiones ya fueron detalladas en la sección Dimensiones Conformadas.

Dimensiones no conformadas:

**Tabla 33.**

*Dimensión Demografía*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Demografía</b>	Estrato	Estrato	1, 2, ...
	Discapacidad	Existencia de la discapacidad	Si, No
	Tipo Discapacidad	Nombre de la discapacidad	Sordo, ...
	<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Jerarquía Discapacidad	Jerarquía compuesta por: Discapacidad, Tipo Discapacidad	Sí->Visual

**Tabla 34.**

*Dimensión Apoyo Estudiantil*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Apoyo Estudiantil</b>	Tipo Apoyo	El enfoque del apoyo de bienestar es mantener la permanencia, el académico motivar a los estudiantes sobresalientes.	bienestar, académico
	Origen Apoyo	Específica si es el apoyo es otorgado por la IES. Nombre del apoyo.	Interno, externo Matrícula de honor, ...
	<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Jerarquía Apoyo	Jerarquía compuesta por: Tipo Apoyo, Origen, Apoyo.	Bienestar -> interno -> alimentación

**Tabla 35.**

*Dimensión Deserción Estudiantil*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Deserción Estudiantil</b>	Estado deserción	Deserción con el periodo en el cual decide desertar	Deserta semestre 1, Deserta semestre 2, ...
	Motivo Deserción	Motivo de la deserción	Económico, Familiar, Salud, Laboral...
	<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Jerarquía Deserción	Jerarquía compuesta por: Estado deserción, motivo deserción	Deserta semestre 1 -> Familiar

**Tabla 36.**

*Dimensión Idioma*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Idioma</b>	Idioma	Idioma que maneja el estudiante	No aplica, inglés, francés, alemán, ...
	Nivel de Idioma	Nivel de Idioma que maneja el estudiante	A1, A2, B1, ...
	<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Jerarquía Idioma	Jerarquía compuesta por: Estado deserción, motivo deserción	Inglés -> B2

El esquema de vinculación también relacionado con el Factor de Estudiantes se explica en el ANEXO B.

## 2. PROCESOS ACADÉMICOS: Modelo Dimensional Actividad Profesorial

El modelo de *Actividad Profesorial* de la Figura 18 es diseñado a partir de los requerimientos que se presentan en la Tabla 37, los cuales fueron obtenidos en la Recolección de Requerimientos.

**Tabla 37.**

*Requerimientos extraídos del Factor Profesores*

<b>PROCESOS ACADÉMICOS: Requerimientos del Factor de Profesores</b>
<b>Requerimientos más frecuentes</b>
1. Número de profesores y profesores TCE por tipo de vinculación, jornada, término, nivel de estudios, facultad, sede, periodo y año.
2. Relación: estudiantes/profesor y estudiantes/profesor TCE por tipo de vinculación, nivel de estudios, periodo y año.
3. Porcentaje promedio de tiempo dedicado a docencia, investigación, administración y proyección social por vinculación, nivel de estudios, categoría, periodo y año.
4. Número de profesores con formación extranjera por nivel de estudios, país, continente y año.
5. Número de profesores con movilidades por área de conocimiento, sede y año.
6. Porcentaje de profesores por vinculación, nivel de estudios, categoría, periodo y año.
<b>Requerimientos menos frecuentes</b>
7. Número de profesores ocasionales por periodo y año.
8. Porcentaje de profesores y profesores TCE por vinculación, nivel de estudios y año.
9. Número de profesores por resultado de evaluación docente, categoría, vinculación, facultad, periodo y año.
10. Número de profesores por ascensos de categoría por facultad y año.
11. Número de profesores con docencia excepcional y reconocimiento de emérito por sede y año.
12. Número de profesores con comisiones de estudio posdoctoral por año.
13. Número de profesores con comisiones doctorales por sede, periodo y año.
<b>Requerimientos propuestos</b>
14. Porcentaje de profesores con formación extranjera por nivel de estudios, país, continente y año.
15. Porcentaje de profesores por ascensos de categoría por facultad y año.
16. Porcentaje de profesores con docencia excepcional y reconocimiento emérito por sede y año.
17. Porcentaje de profesores con comisiones de estudio posdoctoral, doctoral, maestría, especialización médico-quirúrgica y pasantía empresariales por sede, periodo y año.

En la Figura 18 se hace una señalización en color verde para describir la forma en que se cumplen los requerimientos, tomando como ejemplo el primer requerimiento más frecuente de la Tabla 37, que dice:

*Número de profesores y profesores TCE por tipo de vinculación, jornada, término, nivel de estudios, facultad, sede, periodo y año.*

El número de profesores se representa con la medida conteo profesores y profesores TCE, presentes en la tabla de hechos. En cuanto al contexto que envuelve el hecho se usan las dimensiones *Indicador Profesor, Facultad, Programa, Periodo y Dedicación* como se señala en la Figura 18.

Cada uno de los requerimientos de la Tabla 37 se ha representado en el modelo *Actividad Profesoral*, teniendo en cuenta la característica de adaptabilidad que fue explicada al principio de este capítulo.

#### ***Lista de medidas***

El esquema estrella de *Actividad Profesoral* de la Figura 18 presenta las ocho medidas básicas y doce medidas derivadas extraídas de la Tabla 37.

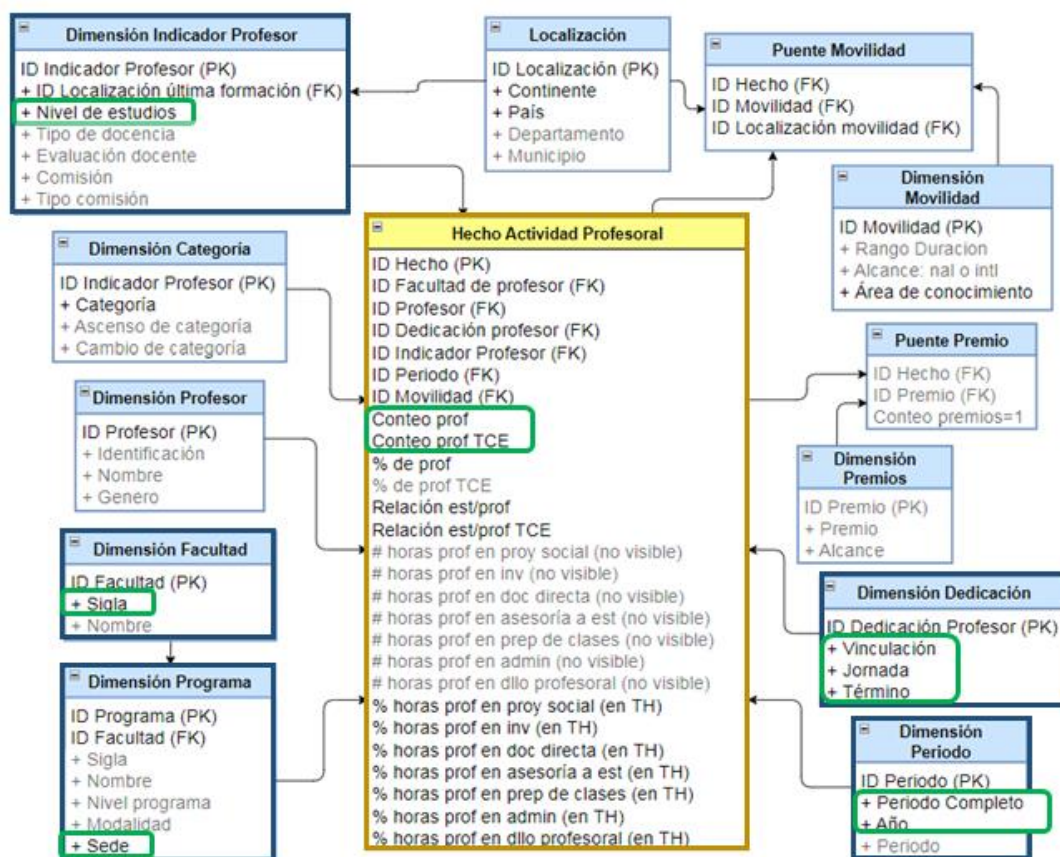
Las medidas básicas son:

- *Conteo profesor*: cantidad de profesores en una consulta
- *Número horas proyección social*: cantidad de horas semanales dedicadas por los profesores a proyección social.
- *Número horas investigación*: cantidad de horas semanales dedicadas por los profesores a investigación.
- *Número horas docencia directa*: cantidad de horas semanales dedicadas por los profesores a docencia directa.
- *Número horas asesoría a estudiantes*: cantidad de horas semanales dedicadas por los profesores a asesoría a estudiantes.

- *Número horas preparación de clases:* cantidad de horas semanales dedicadas por los profesores a preparación de clases.
- *Número horas de administración:* cantidad de horas semanales dedicadas por los profesores a administración.
- *Número horas desarrollo profesoral:* cantidad de horas semanales dedicadas por los profesores a desarrollo profesoral.

Figura 18.

Modelo Dimensional Actividad Profesoral



Las medidas derivadas son:

- *Conteo profesor TCE:* cantidad de profesores en Tiempo Completo Equivalente (TCE).

- *Porcentaje profesor*: razón entre los profesores con un atributo particular entre el total de profesores.
- *Porcentaje profesor TCE*: razón entre los profesores TCE con un atributo particular entre el total de profesores TCE.
- *Relación estudiante/profesor*: razón entre la cantidad de estudiantes y profesores.
- *Relación estudiante/profesor TCE*: razón entre la cantidad de estudiantes y profesores TCE.
- *Porcentaje horas proyección social*: razón entre el número de horas semanales dedicadas por los profesores a proyección social entre las horas totales.
- *Porcentaje horas investigación*: razón entre el número de horas semanales dedicadas por los profesores a investigación entre las horas totales.
- *Porcentaje horas docencia directa*: razón entre el número de horas semanales dedicadas por los profesores a docencia directa entre las horas totales.
- *Porcentaje horas asesoría a estudiantes*: razón entre el número de horas semanales dedicadas por los profesores a asesoría de estudiantes entre las horas totales.
- *Porcentaje horas preparación de clase*: razón entre el número de horas semanales dedicadas por los profesores a preparación de clase entre las horas totales.
- *Porcentaje horas administración*: razón entre el número de horas semanales dedicadas por los profesores a administración entre las horas totales.
- *Porcentaje horas desarrollo profesoral*: razón entre las horas semanales dedicadas por los profesores a desarrollo profesoral entre las horas totales.

La Tabla 38 presenta un resumen de las medidas del modelo dimensional de la Figura 18.

**Tabla 38.**

*Lista de medidas de Modelo Dimensional Actividad Profesoral*

<b>Medida</b>	<b>Función de Agregación</b>	<b>Visibilidad</b>
Conteo profesor	Conteo	Visible
Número horas proyección social	Suma	No Visible
Número horas investigación	Suma	No Visible
Número horas docencia directa	Suma	No Visible
Número horas asesoría a estudiantes	Suma	No Visible
Número horas preparación de clase	Suma	No Visible
Número horas administración	Suma	No Visible
Número horas desarrollo profesoral	Suma	No Visible
Conteo profesor TCE	[conteo Tiempo Completo + (conteo Medio Tiempo/2) + (conteo cátedra*Número horas docencia directa/40)]	Visible
Porcentaje profesor	[conteo por atributo de dimensión / conteo total]*100	Visible
Porcentaje profesor TCE	[conteo por atributo de dimensión / conteo total]*100	Visible
Relación estudiante/profesor	[conteo estudiante / conteo profesor]	Visible
Relación estudiante/profesor TCE	[conteo estudiante / conteo profesor TCE]	Visible
Porcentaje horas proyección social	[suma horas proyección social / suma horas totales]*100	Visible
Porcentaje horas investigación	[suma horas investigación / suma horas totales]*100	Visible
Porcentaje horas docencia directa	[suma horas docencia directa / suma horas totales]*100	Visible
Porcentaje horas asesoría a estudiantes	[suma horas asesoría a estudiantes / suma horas totales]*100	Visible
Porcentaje horas preparación de clase	[suma horas preparación de clase / suma horas totales]*100	Visible
Porcentaje horas administración	[suma horas administración / suma horas totales]*100	Visible
Porcentaje horas desarrollo profesoral	[suma horas desarrollo profesoral / suma horas totales]*100	Visible

*Detalle de dimensiones*

Las consultas de jornada profesoral tienen un contexto descrito por diez dimensiones, de las cuales siete son dimensiones conformadas y ya fueron detalladas al principio del capítulo. A continuación, las dimensiones no conformadas serán detalladas con sus atributos, una breve descripción y un ejemplo de los datos almacenados, también son mencionadas las jerarquías definidas.

Dimensiones no conformadas:

**Tabla 39.**

*Dimensión Categoría*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Categoría</b>	Categoría	Categoría del profesor	Auxiliar, Asistente, ...
	Ascenso categoría	Ascenso del profesor en su categoría	Estable, aux-asis, aux-aso, aux-titu, ...
	Cambio categoría	Cantidad de cambios hechos	Sin cambio, 1 cambio, 2 cambios, ...

**Tabla 40.**

*Dimensión Premio*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Premio</b>	Premio	Nombre del premio	Docencia excepcional, ...
	Alcance	Grado de alcance del premio	Institucional, Nacional, Internacional
	<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Jerarquía Premio	Jerarquía compuesta por: alcance del premio, premio	Institucional -> Docencia excepcional

**Tabla 41.**

*Dimensión Indicador Profesor*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Indicador Profesor</b>	Nivel estudios	Nivel de estudios del profesor	Doctorado, Maestría, ...
	Tipo docencia	Nivel de estudiantes de docencia	Ninguna, solo pregrado, ...
	Evaluación docente	Rango de resultado de evaluación docente	[<3.5], [3.6-3.9], ...

Comisión	Comisión de profesor	En comisión, No está en comisión
Tipo comisión	Tipo de comisión del profesor	Posdoctoral, doctoral, maestría...
<b>Jerarquía</b>	<b>Descripción</b>	<b>Ejemplo</b>
Jerarquía Comisión	Jerarquía compuesta por: Comisión, Tipo de Comisión	En comisión -> doctoral

El esquema de Actividad de Profesor Visitante también relacionado con el Factor de Profesores se explica en el ANEXO B.

### 3. Procesos académicos: Modelo Dimensional Participación Cursos

El modelo de *Participación en Cursos* de la Figura 19 es diseñado a partir de los requerimientos que se presentan en la Tabla 42, los cuales fueron obtenidos en la Recolección de Requerimientos.

**Tabla 42.**

*Requerimientos extraídos del Factor Procesos Académicos*

<b>PROCESOS ACADÉMICOS: Requerimientos del Factor de Procesos Académicos</b>	
<b>Requerimientos más frecuentes</b>	
1.	Número de estudiantes en cursos de lengua extranjera por nivel de formación, facultad, sede, modalidad, periodo y año.
2.	Número de cursos de lengua extranjera por idioma, sede, modalidad, periodo y año.
<b>Requerimientos menos frecuentes</b>	
3.	Número de profesores en cursos de lengua extranjera por idioma, sede, modalidad, periodo y año.
<b>Requerimientos propuestos</b>	
4.	Porcentaje de estudiantes que aprueban cursos de lengua extranjera por nivel de formación, facultad, sede, modalidad, periodo y año.
5.	Porcentaje de profesores en cursos de lengua extranjera por idioma, sede, modalidad, periodo y año.

En la Figura 19 se hace una señalización en color verde para describir la forma en que se cumplen los requerimientos, tomando como ejemplo el primer requerimiento más frecuente de la Tabla 42, que dice:

*Número de estudiantes en cursos de lengua extranjera por nivel de formación, facultad, sede, modalidad, periodo y año.*

El número de estudiantes participantes en cursos se representa con la medida conteo participantes en curso, presentes en la tabla de hechos. En cuanto al contexto que envuelve el hecho se usan los atributos *Sigla de Facultad, Modalidad, Sede, Rol de estudiante, Nivel de Formación, Tipo de Curso (en Idioma), Idioma, Periodo* y *Año* como se señala en la Figura 19.

Cada uno de los requerimientos de la Tabla 42 se ha representado en el modelo Participación en Cursos, teniendo en cuenta la característica de adaptabilidad que fue explicada anteriormente ubicando elementos en color negro y gris según corresponda.

#### ***Lista de medidas***

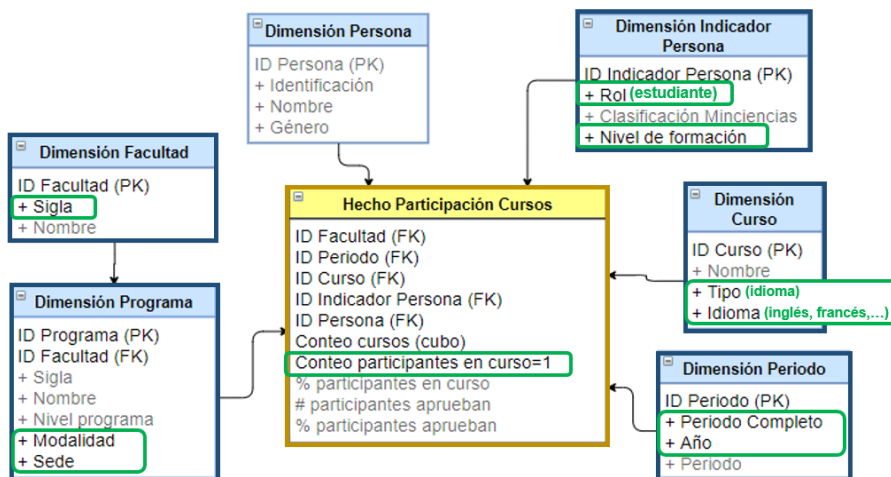
El esquema estrella de *Participación Cursos* de la Figura 19 presenta las tres medidas básicas y dos medidas derivadas extraídas de la Tabla 42.

Las medidas básicas son:

- *Conteo cursos*: cantidad de cursos en una consulta
- *Conteo de participantes en curso*: cantidad de participantes en un curso.

**Figura 19.**

*Modelo Dimensional de Participación en Cursos*



Las medidas derivadas son:

- *Conteo de participantes que aprueban:* cantidad de participantes que aprueban el curso.
- *Porcentaje participantes en curso:* razón entre participantes de un curso con un atributo y el total de participantes.
- *Porcentaje participantes aprueban:* razón entre participantes en un curso que aprueban un curso y el total de participantes.

La Tabla 43 presenta un resumen de las medidas del modelo dimensional de la Figura 19.

**Tabla 43.**

*Lista de medidas de modelo de participación en cursos*

Medida	Función de Agregación	Visibilidad
Conteo cursos	Conteo distinto de cursos	Visible
Conteo participantes en curso	Conteo	Visible
Conteo participantes aprueban	Conteo	Visible
Medida Derivada	Fórmula de cálculo	Visibilidad

Porcentaje participantes en curso	[conteo por atributo de dimensión / conteo total]*100	Visible
Porcentaje participantes aprueban	[conteo por atributo de dimensión / conteo total]*100	Visible

**Detalle de dimensiones**

Las consultas de cursos de este modelo tienen un contexto descrito por seis dimensiones, de las cuales solo es no conformada la de curso, que se describe a continuación.

Dimensiones no conformadas:

**Tabla 44.**

*Dimensión Curso*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Curso</b>	Nombre	Nombre del curso	Inglés 1, ...
	Tipo	Tipo de Curso	Idiomas, Desarrollo Profesional, ...
	Idioma	Idioma que ofrece la institución	No aplica, inglés, Francés, ...

En este punto se han descrito los modelos que cumplen los requerimientos de procesos académicos desde la perspectiva de los factores de acreditación del CNA relacionados con estudiantes, profesores y procesos académicos. Sin embargo, los requerimientos de la

Tabla 45 hacen parte del factor de procesos académicos pero hacen uso del esquema de *Matrícula Estudiantil* en la Figura 17, que ya se explicó en la sección Recolección de Requerimientos.

**Tabla 45.**

*Requerimientos del Factor de Procesos Académicos*

<b>PROCESOS ACADÉMICOS: Requerimientos del Factor de Procesos Académicos</b>
<p><b>Requerimiento menos frecuente</b></p> <p>1. Número de profesores capacitados en desarrollo profesoral y en TIC por año.</p>
<p><b>Requerimiento propuesto</b></p> <p>2. Porcentaje de profesores capacitados en desarrollo profesoral y en TIC por año.</p>

Los esquemas de Ambientes de Discusión y Actividades Académicas también relacionados con el Factor de Procesos Académicos se explican en el ANEXO B.

**4. Procesos investigativos: Modelo Dimensional Producción Intelectual**

El modelo de *Producción Intelectual* de la Figura 20 es diseñado a partir de los requerimientos que se presentan en la Tabla 46, los cuales fueron obtenidos en la sección Recolección de Requerimientos.

**Tabla 46.**

*Requerimientos extraídos del Factor de Investigación y creación artística y cultural*

<b>PROCESOS INVESTIGATIVOS: Requerimientos del Factor de Investigación y creación artística y cultural</b>
<p><b>Requerimientos más frecuentes</b></p> <p>1. Número de ponencias docentes por alcance, categoría de grupo de investigación y año.</p> <p>2. Número de artículos por alcance de profesores planta por categoría de indexación, facultad y año.</p> <p>3. Número de libros realizados por categoría de grupo de investigación facultad y año.</p> <p>4. Número de publicaciones en Scopus por facultad, área de conocimiento y año.</p> <p>5. Número de software de autoría de profesores por año.</p> <p>6. Número de capítulos de libro de autoría de profesores por categoría de grupo de investigación, facultad y año.</p>
<p><b>Requerimientos menos frecuentes</b></p> <p>7. Número de traducciones de profesores por año.</p> <p>8. Número de patentes nacionales e internacionales de profesores planta por año.</p> <p>9. Número de obras artísticas hechas por profesores al año.</p> <p>10. Número de publicaciones en WoS por facultad y año.</p>

---

11. Número de trabajos dirigidos de maestría y doctorado por profesores planta por grupo de investigación según categoría y año.

**Requerimientos propuestos**

12. Número de libros de ensayo por año.

13. Número de estudios posdoctorales (9, 10, 11 o 12 meses) por año.

14. Número de reseñas críticas por categoría de revista y año.

---

En la Figura 20 se hace una señalización en color verde para describir la forma en que se cumplen los requerimientos, tomando como ejemplo el primer requerimiento más frecuente de la Tabla 46, que dice:

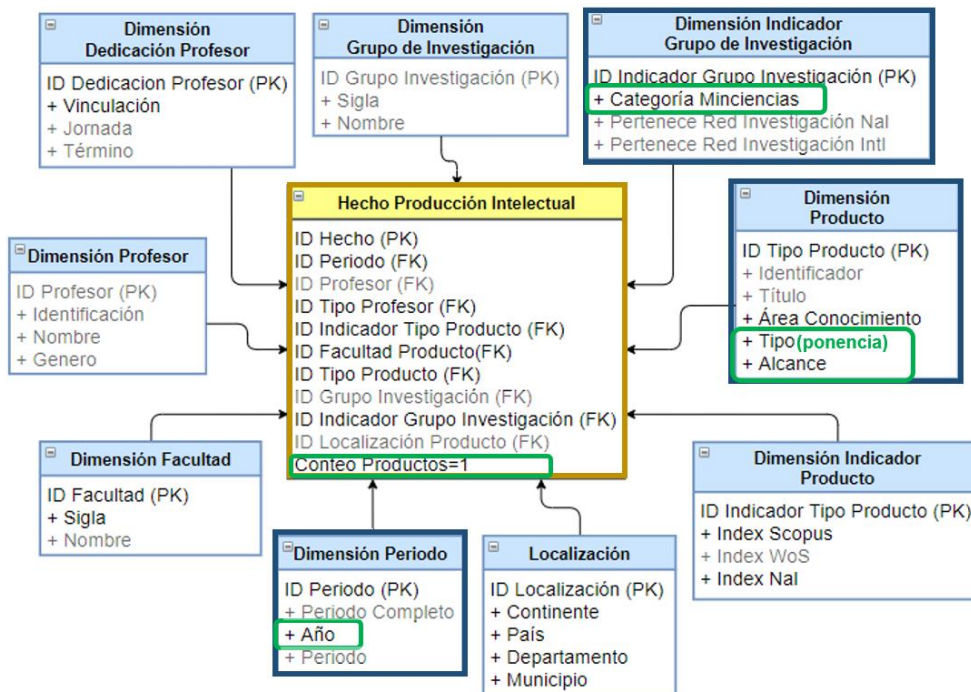
*Número de ponencias docentes por alcance, categoría de grupo de investigación y año.*

El número de profesores se representa con la medida conteo de productos, presente en la tabla de hechos. En cuanto al contexto que envuelve el hecho se usan las dimensiones *Indicador Grupo de Investigación, Producto y Periodo* como se señala en la Figura 20.

Cada uno de los requerimientos de la Tabla 46 se ha representado en el modelo *Producción Intelectual*, en cuanto al requerimiento de ejemplo se usan atributos y medidas en color negro por ser un requerimiento frecuente.

**Figura 20.**

*Modelo Dimensional de Producción Intelectual*



**Lista de medidas**

En el esquema estrella de *Producción Intelectual* de la Figura 20 se presenta la medida conteo productos, con la cual se puede consultar la cantidad de *Producción Intelectual* respecto a diversas dimensiones de contexto, esta medida se muestra en la Tabla 47.

**Tabla 47.**

*Medida del Modelo Producción Intelectual*

Medida	Función de Agregación	Visibilidad
Conteo de Productos	Conteo	Visible

**Detalle de dimensiones**

Las consultas de *Producción Intelectual* de este modelo tienen un contexto descrito por nueve dimensiones, de las cuales dos son no conformadas y siete conformadas. Debido a que las dimensiones conformadas ya fueron explicadas al principio del capítulo se detallan a continuación las dos dimensiones no conformadas.

**Tabla 48.**

*Dimensión Producto*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Producto</b>	Identificador	Identificador del Producto	1432,...
	Título	Nombre del producto	Revisión Sistemática de la Literatura...
	Área de conocimiento	Área de conocimiento	Ciencias de la Salud,...
	Tipo	Tipo de producto	Ponencia, traducción,...
	Alcance	Alcance del producto	Local, Regional, Nacional,...

**Tabla 49.**

*Dimensión Indicador Producto*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Indicador Producto</b>	Index Scopus	Indexación del producto en Scopus	Si, No
	Index WoS	Indexación del producto en Web of Science	Si, No
	Index Nal	Indexación del producto en pubindex	No, A2, A1,...

Los esquemas de Participación en Semilleros, Participación en Proyectos, Uso Laboratorios, Participación en Grupos de Investigación y Financiación de Proyectos también relacionados con el Factor de Investigación se explica en el ANEXO B.

**5. Procesos de seguimiento a graduados: Modelo Dimensional Graduados**

El modelo de *Graduación* de la Figura 21 es diseñado a partir de los requerimientos que se presentan en la Tabla 50, los cuales fueron obtenidos en la sección Recolección de Requerimientos.

**Tabla 50.**

*Requerimientos extraídos del Factor de Pertinencia e Impacto Social*

<b>PROCESOS DE SEGUIMIENTO A GRADUADOS: Requerimientos del Factor de Pertinencia e Impacto Social</b>
<p><b>Requerimientos menos frecuentes</b></p> <p>1. Número de graduados inscritos en la plataforma institucional por país en que se encuentra, facultad y año.</p>
<p><b>Requerimientos propuestos</b></p> <p>2. Número de graduados que contestan encuestas por facultad y año.</p> <p>3. Porcentaje de egresados inscritos en la plataforma institucional por facultad y año.</p> <p>4. Porcentaje de egresados que contestan encuestas por facultad y año.</p> <p>5. Número de egresados que hacen parte de organismos de decisión institucionales por facultad y año.</p> <p>6. Número de egresados que participan en procesos de autoevaluación institucional, autoevaluación de programas y reformas curriculares por facultad y año.</p> <p>7. Número de egresados destacados en programa institucional de reconocimiento a graduados al año.</p> <p>8. Número de egresados ponentes en eventos académicos de los programas por facultad y año.</p> <p>9. Número de egresados que participan en encuentros o congresos de egresados por facultad y año.</p> <p>10. Número de egresados pertenecientes a asociaciones de egresados por facultad y año.</p>

En la Figura 21 se hace una señalización en color naranja para describir la forma en que se cumplen los requerimientos, tomando como ejemplo el primer requerimiento menos frecuente de la Tabla 50, que dice:

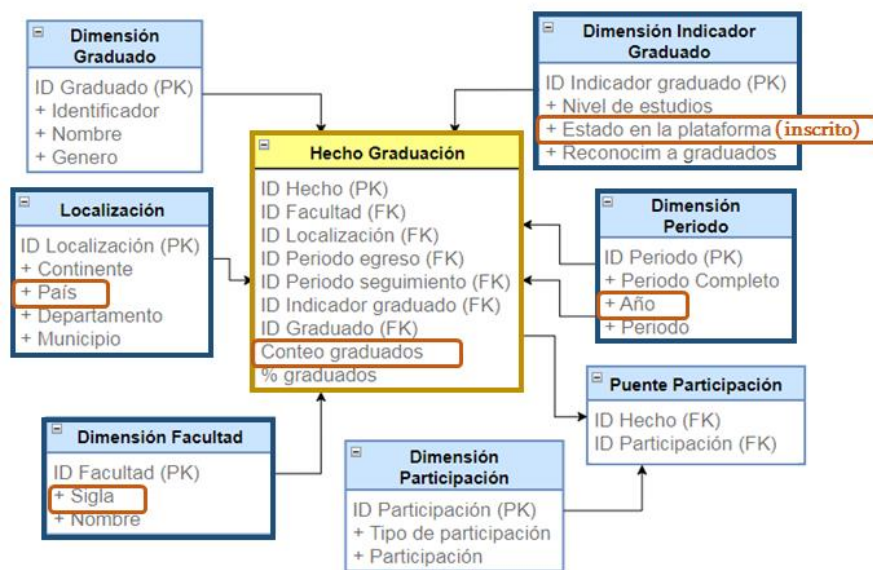
*Número de graduados inscritos en la plataforma institucional por país en que se encuentra, facultad y año.*

El número de graduados inscritos se representa con la medida conteo de graduados, presente en la tabla de hechos. En cuanto al contexto que envuelve el hecho

se usan las dimensiones *Localización*, *Facultad*, *Periodo* e *Indicador Graduado* como se señala en la Figura 21. Por la característica de adaptabilidad, todos los elementos del modelo están en color gris por solamente involucrar requerimientos menos frecuentes y propuestos.

**Figura 21.**

*Modelo Dimensional de Graduación*



**Lista de medidas**

El esquema estrella de *Graduación* tiene una medida básica y una derivada extraídas de la Tabla 51.

La medida básica es:

- *Conteo graduados*: cantidad de graduados en una consulta

Las medidas derivadas son:

- *Porcentaje graduados*: razón entre la cantidad de graduados con un atributo dimensional y el total de graduados.

La Tabla 51 presenta un resumen de las medidas del modelo dimensional de la Figura 21.

**Tabla 51.**

*Lista de medidas del modelo de graduación*

<b>Medida</b>	<b>Función de Agregación</b>	<b>Visibilidad</b>
Conteo graduados	Conteo	Visible
<b>Medida Derivada</b>	<b>Fórmula de cálculo</b>	<b>Visibilidad</b>
Porcentaje graduados	[conteo por atributo de dimensión / conteo total]*100	Visible

***Detalle de dimensiones***

Las consultas de graduación de este modelo tienen un contexto descrito por seis dimensiones, de las cuales tres son no conformadas y se detallan a continuación.

Dimensiones no conformadas:

**Tabla 52.**

*Dimensión Graduado*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Graduado</b>	Identificador	Identificación del egresado	1987654368, ...
	Nombre	Nombre del egresado	Pedro Díaz Jiménez,...
	Género	Género del egresado	M, F, Otro

**Tabla 53.**

*Dimensión Participación*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Participación</b>	Tipo participación	Sigla del Grupo de Investigación	STI, GTI,...
	Participación	Nombre del Grupo de Investigación	Grupo en Sistemas y Tecnologías de la Información,...

**Tabla 54.**

*Dimensión Indicador Graduado*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Indicador Graduado</b>	Nivel	Nivel del programa de egreso	Doctorado, Maestría,...
	Estado en plataforma	Especifica si el egresado está inscrito en la plataforma institucional de egresados	Inscrito, No inscrito
	Reconocimiento graduados	Pertenece a programa de reconocimiento a graduados	Si, No

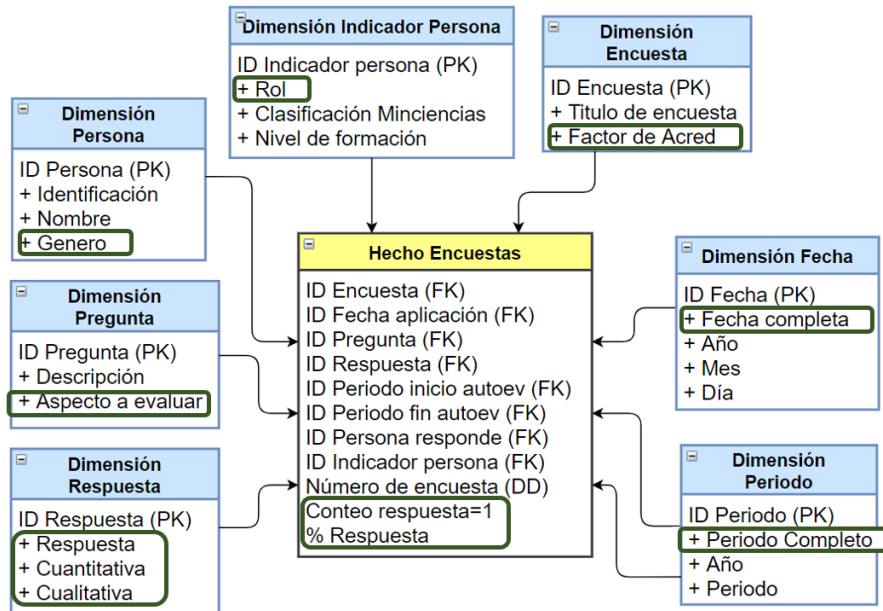
**6. Modelo Dimensional Encuestas**

El modelo de *Encuestas* de la Figura 22 es diseñado a partir de los aspectos a evaluar de los procesos académicos, investigativos y de seguimiento a graduados donde se pidieron percepciones de la comunidad académica.

Se utiliza el conteo de respuestas de las encuestas pedidas en los factores de acreditación del CNA usando como contexto siete dimensiones.

**Figura 22.**

*Modelo Dimensional de Encuestas*



**Lista de medidas**

El esquema estrella de *Encuestas* de la Figura 22 presenta una medida básica y una derivada.

La medida básica es:

- *Conteo graduados*: cantidad de graduados en una consulta

La medida derivada es:

- *Porcentaje graduados*: razón entre los graduados con un atributo de dimensión y el total de graduados.

La Tabla 55 presenta un resumen de las medidas del modelo dimensional de la Figura 22.

**Tabla 55.**

*Lista de medidas para modelo encuestas*

<b>Medida</b>	<b>Función de Agregación</b>	<b>Visibilidad</b>
Conteo graduados	Conteo	Visible
<b>Medida Derivada</b>	<b>Fórmula de cálculo</b>	<b>Visibilidad</b>
Porcentaje graduados	[conteo por atributo de dimensión / conteo total]*100	Visible

***Detalle de dimensiones***

Las consultas de graduación de este modelo tienen un contexto descrito por siete dimensiones, de las cuales las tres no conformadas serán presentadas a continuación con sus atributos, una breve descripción y un ejemplo de los datos almacenados.

Dimensiones no conformadas:

**Tabla 56.**

*Dimensión Encuesta*

<b>Encuesta</b>	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Título encuesta	Título de la encuesta	Apreciación sobre estímulos estudiantiles
	Factor de Acreditación	Factor de Acreditación del Acuerdo 03 de 2014 que es cubierto por la encuesta	Estudiantes,...

**Tabla 57.**

*Dimensión Pregunta*

<b>Pregunta</b>	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
	Descripción	Pregunta	¿La Institución otorga las condiciones de bienestar necesarias para la actividad académica?
	Aspecto a evaluar	Aspecto a evaluar del Acuerdo 03 de 2014	F1C2A3: Factor1-Característica2-AspectoaEvaluar3

**Tabla 58.**

*Dimensión Respuesta*

	<b>Atributo</b>	<b>Descripción</b>	<b>Ejemplo</b>
<b>Respuesta</b>	Respuesta	Respuesta	En alto grado,...
	Cuantitativa	Se especifica si la pregunta de la encuesta es cuantitativa	Si, No
	Cualitativa	Se especifica si la pregunta de la encuesta es cualitativa	Si, No

**4.1.4. Casos de diseño utilizados en los modelos dimensionales**

Luego de explicar los quince modelos propuestos para procesos académicos, investigativos y de seguimiento a graduados, en esta sección se presenta a continuación los casos de diseño que fueron utilizados durante la elaboración de dichos esquemas, acompañados de su descripción y la parte del modelo en que fue aplicada.

**4.1.4.1. Dimensión que Cambia Lentamente Tipo 4: Mini-dimensión**

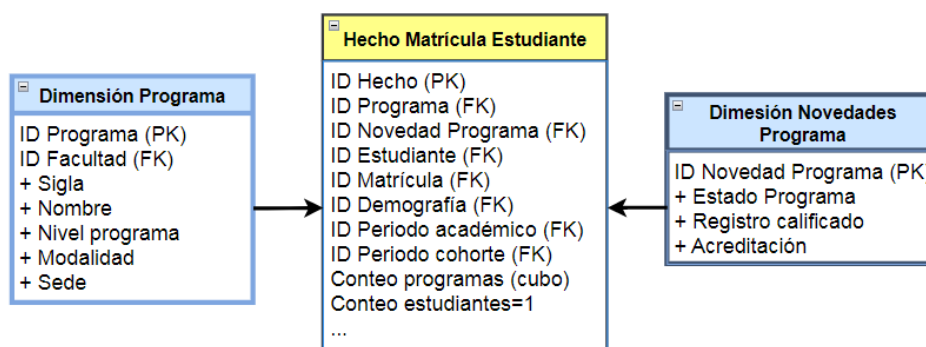
Este caso de diseño está presente en aquellos los modelos dimensionales donde la dimensión tiene atributos que cambian en el tiempo y se necesita almacenar el histórico de estos cambios.

Los modelos en los cuales se hizo necesario utilizar mini-dimensiones fueron: *Matrícula Estudiantil* con las dimensiones *Matrícula Estudiantil*, *Novedades Programa* (ver Figura 23), y *Deserción Estudiantil*; *Vinculación* con la dimensión *Novedades Programa*; *Jornada Profesor* con las dimensiones *Indicador Profesor*, *Categoría* y *Dedicación*; *Profesor Visitante* con la dimensión *Indicador Profesor Visitante*; *Actividades Académicas* con la dimensión *Indicador Persona*; *Cursos* con también con la dimensión *Indicador Persona*; *Participación en semillero* con la dimensión *Matrícula Estudiantil*; *Participación en Proyectos de Investigación* con las dimensiones *Indicador*

*Grupo de Investigación e Indicador Persona; Uso de Laboratorios con la dimensión Indicador Laboratorio; Participación en Grupos de Investigación con las dimensiones Indicador Persona e Indicador Grupo de Investigación, Producción Intelectual con las dimensiones Indicador Producto, Dedicación Profesor e Indicador Grupo de Investigación; Financiación de Proyectos con la dimensión Indicador Grupo de Investigación; Graduación con la dimensión Indicador Graduado; y finalmente, Encuestas con la dimensión Indicador Persona.*

**Figura 23.**

*Mini-dimensión Novedades Programa en el Modelo Matrícula Estudiantil*



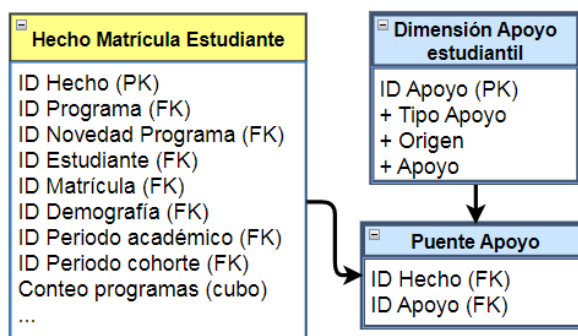
**4.1.4.2. Dimensión multi-valuada y tabla puente**

Este caso de diseño consiste en crear una relación muchos a muchos entre una tabla de hechos y una dimensión mediante una tabla puente, la cual tiene las claves primarias de la tabla de hechos (clave primaria ID Hecho) y la dimensión involucrada.

Este caso de diseño se identificó en los modelos: *Matrícula Estudiantil* con los puentes *Apoyo Estudiantil* (ver Figura 24) e *Idioma; Jornada Profesor* con los puentes *Movilidad y Premios*; y finalmente, *Participación en Proyectos de Investigación* con el puente *Grupos Participantes*. La Figura 24 muestra la relación muchos a muchos entre la tabla de hechos y la dimensión de apoyo estudiantil.

**Figura 24.**

*Dimensión multi-valuada y tabla puente en modelo de Matrícula Estudiantil*

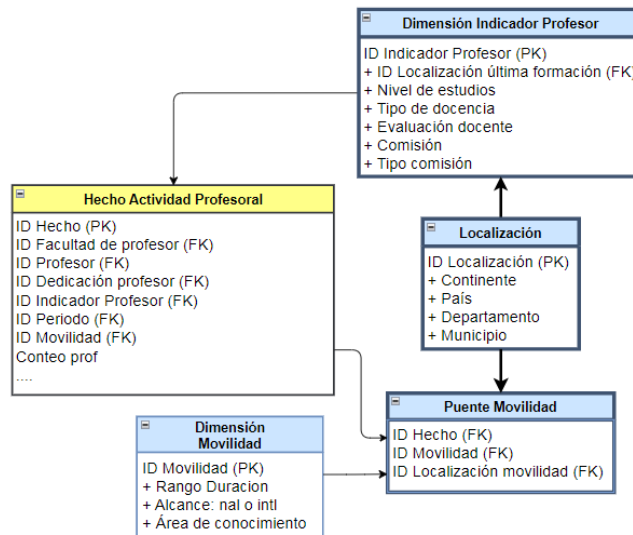


#### 4.1.4.3. Sub-dimensión

Este caso de diseño representa una relación uno a muchos entre dimensiones, busca dividir dimensiones grandes en estructuras más pequeñas. En los modelos del presente trabajo se tuvo que separar las dimensiones *Facultad* y *Programa* debido que algunos de los requerimientos pedía consultas dadas por facultad y no por programa, este comportamiento permitió plantear una sub-dimensión *Programa-Facultad* para los seis modelos donde se requieren consultas por *Programa*. Otro caso del uso de mini-dimensiones es el que se presenta en el modelo de *Jornada profesor*, el cual se aprecia en la Figura 25, la mini-dimensión *Localización* se hizo necesaria para representar el lugar geográfico donde el profesor tuvo su última formación y donde pudo haber hecho alguna movilidad.

**Figura 25.**

*Sub-dimensión en modelo de Jornada Profesor*

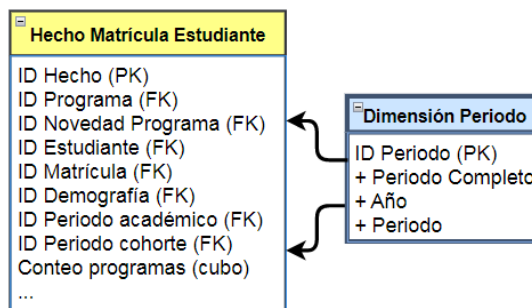


**4.1.4.4. Juego de Roles**

Este caso de diseño se presenta cuando existe una relación múltiple entre dos tablas. Se observa en la Figura 26 la relación de juego de roles entre la dimensión *Periodo* y la tabla de hechos *Matrícula Estudiante*, la dimensión periodo es utilizada para realizar consultas de estudiantes por periodo académico o por cohorte, el juego de roles de *Periodo* también se utilizó en los modelos de *Vinculación*, *Participación Grupos de Investigación*, *Graduación* y *Encuestas*. Otros juegos de roles utilizados fueron: *Dimensión Programa* en el *Modelo Participación Semillero*; y finalmente, Dimensiones *Indicador Persona*, *Facultad* y *Fecha* en el modelo *Proyectos de Investigación*.

**Figura 26.**

*Juego de roles en modelo Matrícula Estudiantil*



**4.2. Validación del modelo dimensional**

Para validar los quince modelos dimensionales planteado en el presente proyecto se utilizan dos herramientas: el grupo focal y la norma ISO-25010 (ISO/IEC, 2011). El grupo focal de expertos sirve de validación, ya que es un espacio de discusión que permite obtener las percepciones de un grupo específico sobre un área de interés (Kontio et al., 2008). El grupo focal no solo valida el diseño del modelo dimensional, también evalúa la adaptabilidad del mismo definida bajo la norma ISO-25010 (ISO/IEC, 2011), establecida como el grado en el que un producto o sistema puede adaptarse de manera efectiva y eficiente a los entornos operativos o de uso.

Autores como Kontio et al. (2008) y Mendoza-Moreno et al. (2013) definen cuatro pasos para el planteamiento de un grupo focal, los cuales son: planificación del grupo focal, definición del grupo focal, construcción de la sesión de grupo focal, análisis de la información y reporte de resultados. A continuación, se presenta el proceso de validación de adaptabilidad del modelo dimensional planteado en el presente proyecto con base en los pasos antes mencionados.

### 4.2.3. Planificación del grupo focal

Los recursos de planteamiento son: el objetivo del grupo focal y la preparación de materiales y procedimientos a seguir.

**Objetivo del grupo focal:** conocer la opinión de profesionales con experiencia en bodegas de datos y bases de datos sobre el modelo dimensional propuesto y su grado de adaptabilidad al contexto de una Institución de Educación Superior. Esta actividad permitirá someter a prueba el modelo con base en la norma 25011 (ISO/IEC, 2011) y descubrir oportunidades de mejora.

**Preparación de materiales y procedimientos a seguir:** los elementos de ejecución del grupo focal son: 1) Estructura del protocolo del grupo focal, 2) Los instrumentos que serán empleados, 3) Definición de métodos de captura y registro de información y 4) definición de métodos de análisis de la información obtenida en el debate.

#### 1. Estructura del protocolo del grupo focal

**Tabla 59.**

*Protocolo del grupo focal*

<b>Elemento</b>	<b>Descripción</b>
Agenda de trabajo	Actividades que se llevan a cabo por cada participante
Cuestionario	Preguntas con las cuales se extrae información relevante para la evaluación del modelo
Propuesta a evaluar	Documento con la descripción del modelo dimensional propuesto

#### 2. Instrumentos que serán empleados

**Tabla 60.**

*Elementos para realizar el grupo focal*

<b>Elemento</b>	<b>Información</b>
Fecha de realización	7 junio de 2022
Hora de inicio	6:06 p.m.

---

Duración	2 horas
Lugar	Online – Vía <i>Google Meets</i>
Tema	Validación de los Modelos Dimensionales del proyecto
Moderador	David Antonio Fuentes Vargas – <i>conectado desde Bucaramanga</i>
Relatora	PhD. Martha Eliana Mendoza Becerra – <i>conectada desde Popayán</i> P1: PhD. Elizabeth León Guzmán – <i>conectada desde Bogotá</i>
Participantes	P2: PhD. Alexander Bustamante Martínez – <i>conectado desde Santa Martha</i> P3: MSc. Diego Armando Quintero Quiñonez – <i>conectado desde Cali</i>

---

### 3. *Métodos de captura y registro de información*

Para realizar la ejecución del grupo focal se contó con las notas de las apreciaciones de cada participante, tomadas por la codirectora del proyecto. También, se hizo entrega a cada participante de un cuestionario y un material de lectura con los modelos en evaluación. Además, se tomó grabación en vídeo de toda la reunión.

### 4. *Métodos de análisis de la información*

Se realiza un análisis estadístico de la información obtenida de los cuestionarios y un análisis cualitativo de las observaciones y oportunidades de mejora recogidas.

#### 4.2.4. **Diseño del grupo focal**

En este paso se definen las estrategias de selección de los participantes del grupo focal. Esta actividad estuvo a cargo del autor e incluyó las siguientes actividades:

***Definición del perfil del participante.*** Para la selección de expertos se definieron los siguientes criterios:

- Tener conocimientos en el diseño de modelos de esquema estrella para bodegas de datos.

- Estar activo en la industria de análisis de datos y/o la academia ya sea como profesor o estudiante.

**Identificación de los participantes.** Para la selección de expertos se definieron los siguientes criterios:

**Tabla 61.**

*Perfil profesional del grupo focal*

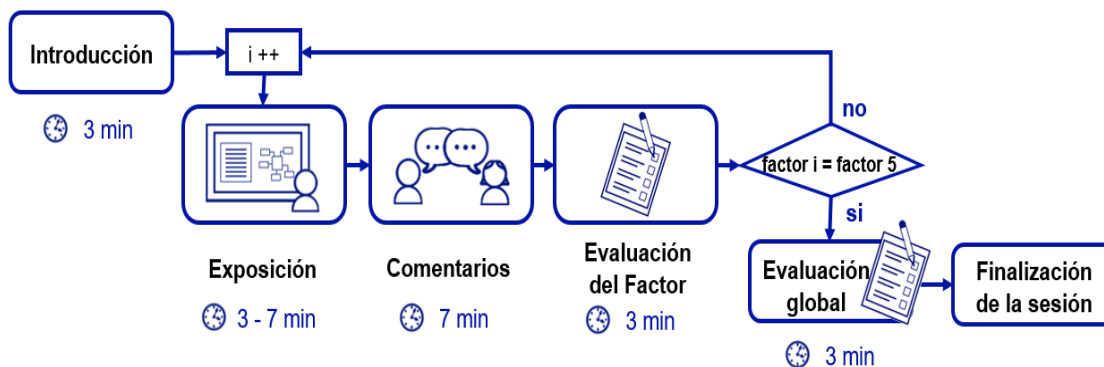
ID	Ocupación	Estudios	Experiencia
P1	Docente de Bodegas de Datos – Universidad del Magdalena	PhD en Computer Science	Más de 5 años de experiencia en diseño e implementación de proyectos de Bodegas de Datos
P2	Docente de Bodegas de Datos – Universidad Nacional de Colombia	PhD en Computer Science	
P3	Arquitecto de Datos – PRAGMA	Master en Data Science	

#### 4.2.5. Construcción de la sesión de grupo focal

Durante el grupo focal se sigue el orden presentado en la Figura 27. El proceso inicia con una introducción de 3 minutos donde se sintetiza el proceso de modelado; se explica el factor correspondiente en un tiempo entre 3 y 7 minutos (6 minutos para el factor estudiantes, 5 minutos para profesor, 4 minutos para procesos académicos, 7 minutos para investigación y 3 minutos para graduados y el modelo de encuestas); luego se hace una sesión de 7 min donde los expertos presentan su realimentación y finalmente se disponen a llenar el cuestionario de evaluación durante 3 min. Este proceso se repite cíclicamente para cada factor de acreditación del CNA, cuando ya se evalúen los cinco factores y el modelo de encuestas se realizara una evaluación final de todo lo visto a través de un cuestionario global.

**Figura 27.**

*Agenda del Grupo Focal*



**4.2.6. Análisis de los datos y reporte de resultado**

Durante el proceso de captura de información se tomó nota de las observaciones realizadas por los participantes. Además, se pidió a los participantes responder un cuestionario al final de la discusión del modelo propuesto, el cual se basa en la norma ISO-25010 en la sub-característica de adaptabilidad. Las primeras cuatro preguntas del cuestionario se diseñaron para obtener una respuesta en la Escala de Likert (Totalmente en desacuerdo, En desacuerdo, Ni de acuerdo ni en desacuerdo, De acuerdo, Totalmente De acuerdo) y las dos últimas preguntas fueron abiertas. El cuestionario global consiste únicamente en dos preguntas abiertas. A continuación, se enuncian en la Tabla 62 y Tabla 63 las preguntas de los cuestionarios.

**Tabla 62.**

*Cuestionario para la evaluación del modelo dimensional de cada factor*

En relación al Factor expuesto (Estudiantes, Profesores, Procesos Académicos, Investigación o Graduados), responda las siguientes preguntas:	
ID	Preguntas cerradas
1	¿Considera que el modelo dimensional propuesto para este factor se adapta a los requerimientos que se repiten con más frecuencia en los informes de acreditación?

2	¿Considera que el modelo dimensional propuesto para este factor se adapta a los requerimientos que se repiten con menos frecuencia en los informes de acreditación?
3	¿Considera que el modelo dimensional propuesto para este factor se adapta a los requerimientos propuestos por el autor y los expertos en acreditación institucional?
4	¿Considera que el modelo permite apoyar la toma de decisiones estratégicas institucionales?
<b>ID</b>	<b>Preguntas abiertas</b>
5	¿Considera que se deben agregar elementos al modelo dimensional presentado? En caso afirmativo ¿cuáles y por qué?
6	¿Considera que se deben eliminar elementos del modelo dimensional presentado? En caso afirmativo ¿cuáles y por qué?

**Tabla 63.**

*Cuestionario para la evaluación global del modelo dimensional*

<b>En relación con los modelos dimensionales expuestos, responda las siguientes preguntas:</b>	
<b>ID</b>	<b>Pregunta</b>
1	¿Tiene algún comentario adicional acerca de los modelos dimensionales presentados?
2	¿Tiene alguna sugerencia adicional acerca de los modelos dimensionales presentados?

Luego de realizar el grupo focal, se realizó un análisis de los aportes obtenidos tanto de los cuestionarios diligenciados como de la sesión de lo dicho en la discusión de expertos.

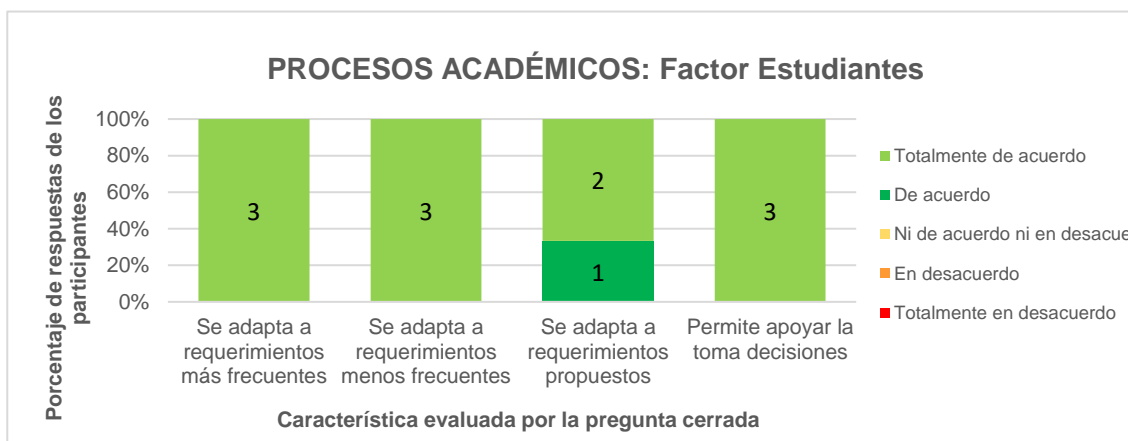
**Análisis de las preguntas cerradas**

Las primeras cuatro preguntas de la Tabla 62 son realizadas a los expertos por medio de un cuestionario para cada uno de los factores, esto se hace después de hacer la exposición del factor y llevar a cabo la realimentación verbal por parte de los participantes. Los resultados de cada factor se presentan de la Figura 28 a la Figura 32, especificando el número de expertos que catalogaron cada pregunta en un nivel de la escala de Likert en su color correspondiente (verde claro, verde oscuro, amarillo, naranja o rojo), en estas figuras se observa que los modelos se están ajustando a la característica de adaptabilidad de la norma ISO-25010, que se cumple si el modelo puede adaptarse de manera efectiva y eficiente a entornos de uso diferentes o en evolución, lo cual se analiza a través del planteamiento de requerimientos *más frecuentes, menos frecuentes y propuestos* para favorecer el diseño de un modelo adaptable a las necesidades analíticas

de la institución, al permitirle decidir si va a implementar modelos de requerimientos más frecuentes o versiones más amplias que agreguen requerimientos menos frecuentes y propuestos.

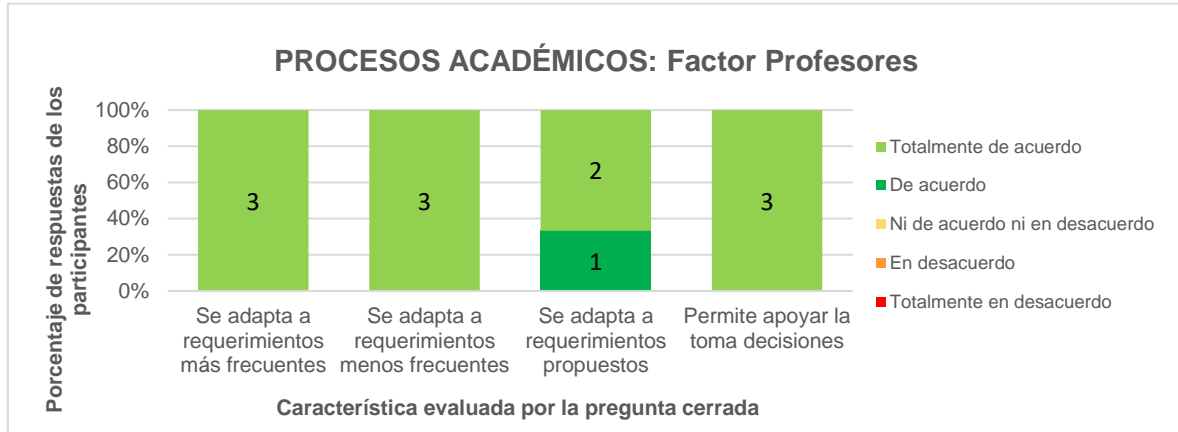
**Figura 28.**

*Resultados de preguntas cerradas para Factor de Estudiantes*



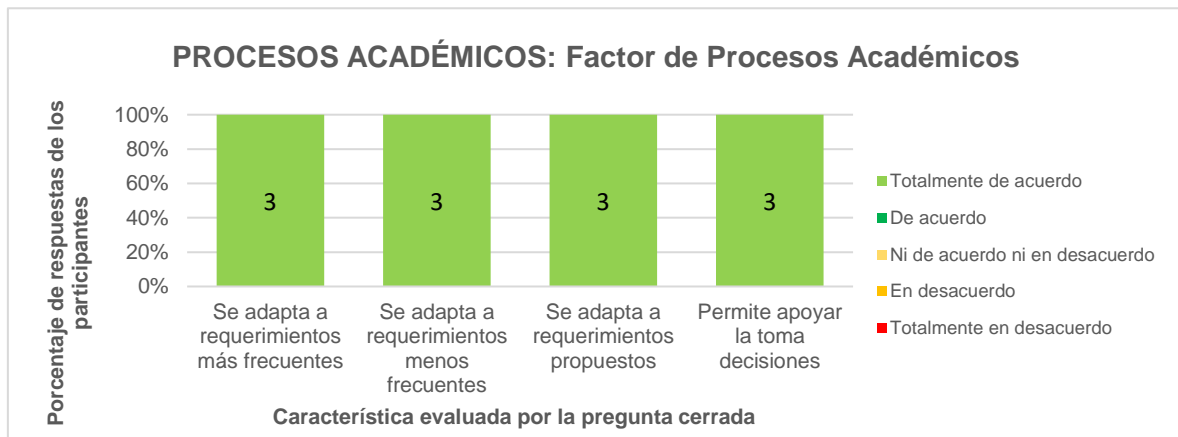
**Figura 29.**

*Resultados de preguntas cerradas para el Factor de Profesores*



**Figura 30.**

*Resultados de preguntas cerradas para el Factor de Procesos Académicos*



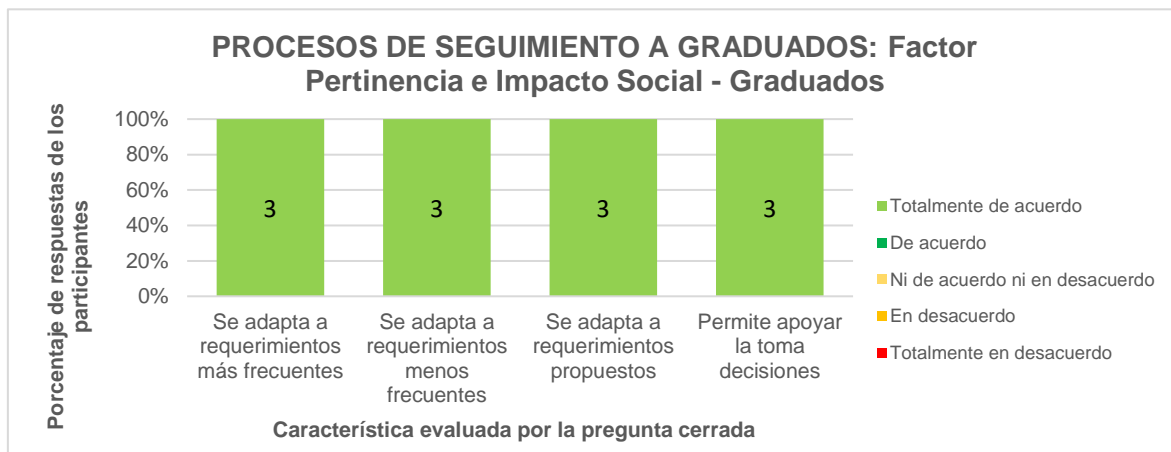
**Figura 31.**

*Resultados de preguntas para el Factor de Investigación*



**Figura 32.**

*Resultados de preguntas cerradas ara el Factor de Graduados y Encuestas*



**Análisis de las preguntas abiertas**

En la Tabla 64 se presentan las sugerencias que los expertos realizaron sobre el trabajo realizado.

**Tabla 64.**

*Respuestas de pregunta abierta por parte de expertos en Modelado Dimensional*

<b>Factor</b>	<b>Partic.</b>	<b>Respuesta</b>	<b>Acción de mejora/Justificación</b>
Procesos Académicos: Factor Estudiante	P1	Almacenamiento de medidas derivadas	Las medidas derivadas son manejadas con reportes estándar para evitar afectar el resultado
	P2	Agregar Estado de reserva de cupo, tener cuidado en adición de medidas para distintas dimensiones	Se agrega el atributo reserva de cupo al modelo de matrícula estudiantil
	P3	Ninguna	No aplica
Procesos Académicos: Factor Profesores	P1	Renombrar la tabla de hechos para aportar mayor claridad	Se renombran las tablas de hecho de forma concordante a los procesos de negocio
	P2	Cambiar nombre a la tabla de hechos	Se renombran las tablas de hecho de forma concordante a los procesos de negocio
	P3	Cambiar nombre a la tabla de hechos	Se renombran las tablas de hecho de forma concordante a los procesos de negocio
Procesos Académicos: Factor Procesos Académicos	P1	Ninguna	No aplica
	P2	Revisar los nombres de las tablas de hechos para que reflejen la "actividad" a la que se refiere	Se renombran las tablas de hecho de forma concordante a los procesos de negocio
	P3	Ninguna	No aplica
Procesos Investigativos: Factor Investigación	P1	Ninguna	No aplica
	P2	Agregar geografía de la publicación a la tabla de hechos de producto	Se agrega la dimensión Localización al modelo de Producción Intelectual (ver Figura 20)
	P3	-Quizás se podría profundizar en el uso de las tablas "Puentes" muchos a muchos y verificar la posibilidad de omitirlas, a futuro sería una relación entre tablas con demasiados registros. -Tener cuidado con las medidas porcentuales. -La tabla puente de participantes en proyecto tiene medidas, es recomendable manejarla como tabla de hechos.	-Se evita el uso de dimensiones varios a varios en los modelos, solo se mantiene la tabla puente de grupos de investigación en el modelo de proyectos de investigación ya que los grupos de investigación no presentan cambios con alta frecuencia. -Las medidas porcentuales son manejadas con reportes estándar para dejar la medida preparada para visualización. -La medida de número de participantes en proyecto que se encontraba en la dimensión puente es trasladada a la tabla de hechos de la. De esta forma, la granularidad queda por participantes y se puede consultar el número de proyectos a partir de un Distinct definido en el cubo OLAP.
Procesos de seguimiento a	P1	Ninguna	No aplica
	P2	Ninguna	No aplica

<p>Graduados y Modelo de Encuestas</p>	<p>P3</p>	<p>Eliminaría la relación de la tabla <i>IndicadorEgresado</i> con la tabla de hecho y más bien dejaría una relación de la dimensión Egresado a Indicador Egresado, al final la información es del egresado y la relación de los hechos a esas 2 tablas siempre tiene que ser a la misma persona.</p>	<p>La dimensión <i>IndicadorEgresado</i> maneja el concepto de minidimensión definido por Ralph Kimball, es individualizada en una dimensión aparte debido a que permite manejar registros que cambian rápidamente.</p>
--	-----------	---	---

## Capítulo 5

### 5.1. Prototipo para procesos académicos

Se describe el desarrollo del prototipo de bodega de datos para matrícula estudiantil de la **¡Error! No se encuentra el origen de la referencia.**, el cual ya fue detallado en la sección 4.1.3, el proceso involucra cuatro fases, que son: Back Room, Front Room, Integración y Despliegue.

#### 5.1.1. Desarrollo del Back Room

Esta sub-fase se divide en las actividades *Diseño físico de la base de datos*, *Implementación de la base de datos*, *Desarrollar el ETL*, *Población y validación de datos*.

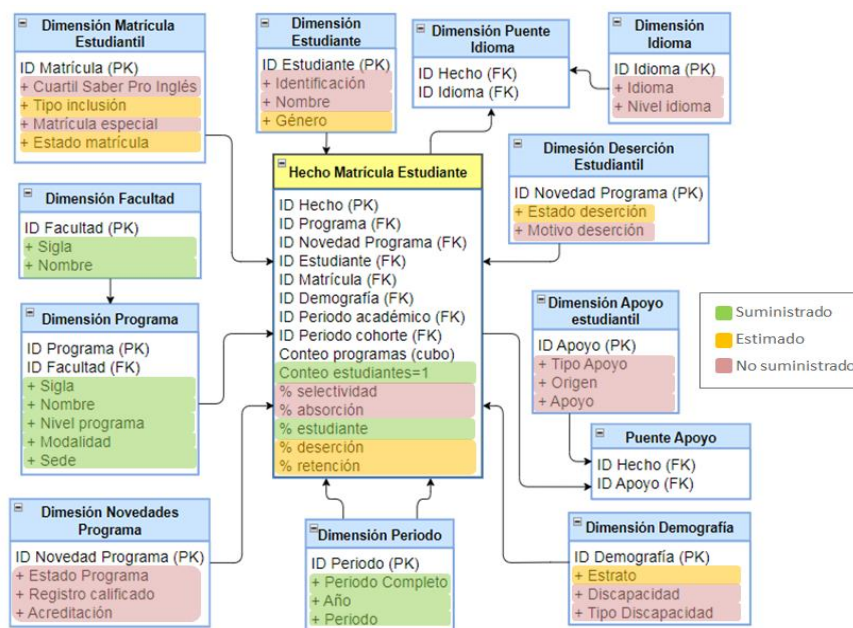
#### Diseño físico de la base de datos

Se recibe por parte de la Universidad un conjunto de 31166 registros de datos anonimizados y relacionados con fechas de grados y matrícula estudiantil por periodo y programa, en el cual el rango de temporal está comprendido entre el 2008 y el 2018. Para implementar el modelo de matrícula estudiantil se hizo necesario estimar el género, el estrato, estado de deserción, cohorte y el número de matrícula. En la Figura 33 se puede observar en color verde los datos suministrados, en naranja los atributos que fueron estimados y en rojo los que no fueron suministrados. Respecto a la cantidad de medidas y atributos correspondientes a los requerimientos más frecuentes y que hacen parte de la

versión reducida del modelo adaptable, la disponibilidad y estimación de los datos permitió implementar el 59% de los atributos y el 40% de las medidas.

**Figura 33.**

*Datos suministrados por la Universidad*



***Estimación del género***

A pesar de que los datos de graduados contienen el género, este atributo no está presente en las tablas de estudiantes que siguen cursando sus estudios. La estimación de los géneros que hacen falta se realiza con los reportes públicos que hizo la universidad para el periodo 2018, manteniendo la distribución por género de cada programa en el conjunto de estudiantes matriculados y otorgando la distribución de género que tienen todos los estudiantes de la universidad a los estudiantes que cursaron más de una carrera.

### ***Estimación del estrato***

Para estimar el estrato se utilizan los reportes públicos que hizo la universidad para el periodo 2018, manteniendo la distribución por estrato de cada programa en el conjunto de estudiantes matriculados y otorgando la distribución de estrato de toda la universidad a la porción de estudiantes que cursaron más de una carrera.

### ***Estimación de la deserción***

Para estimar la deserción se hizo necesario filtrar los estudiantes que dejaron de matricularse dentro del rango temporal de los datos, en el caso de los estudiantes que no se siguieron matriculando en los dos últimos periodos del rango de los datos (2018-1 y 2018-2) se estableció que eran desertores. En caso de que el estudiante no matriculara el último periodo del rango (2018-2) no se toma como desertor debido a la posibilidad de que no haya matriculado ese periodo, pero volviera en el 2019-1. Para corroborar que los estudiantes que dejaron de matricularse realmente sean desertores y no sea por su graduación se realiza la búsqueda del identificador asignado al estudiante dentro de la tabla de graduación, si el estudiante deja de matricularse y no se encuentra en la tabla de graduados se consigna su estado como desertor especificando el nivel en el cual dejó de matricular.

### ***Estimación de la cohorte***

La estimación de la cohorte se realiza con los estudiantes que pertenecen al primer nivel dentro del rango temporal de los datos suministrados, se calcula el periodo más bajo en el cual se matricularon, el cual se establece como cohorte. La cohorte no se pudo establecer para los estudiantes cuyo primer periodo de matrícula no se encontraba dentro del rango de los datos suministrados por la Universidad.

### ***Estimación del número de matrícula***

La estimación del número de matrícula del estudiante se realiza por medio de un conteo incremental que inicia en el periodo de cohorte y culmina cuando el estudiante se haya graduado, haya desertado o quede fuera del rango de los datos suministrados por la universidad (periodo 2018-2).

### **Implementación de la base de datos**

Se implementó la base de datos en el motor relacional SQL Server seleccionado en el CICLO 3 del capítulo de 3.1 Metodología, donde también se realizaron las estimaciones mencionadas en la sección anterior. Luego de establecer el entorno de la base de datos se realiza el desarrollo del proceso de Extracción, Transformación y Carga (ETL) que será detallado en la sección a continuación.

### **Desarrollo del proceso ETL**

A continuación, se describe el flujo de datos en el proceso de Extracción, Transformación y Carga (ETL) de las tablas Estudiante, Programa y Hechos Matricula Estudiantil. Las dimensiones restantes (Deserción, Demografía, Periodo y Facultad) serán abordadas en el ANEXO C debido a que tienen un procedimiento de cargue análogo al descrito para la Dimensión Estudiante. Los cargues serán totales, se propone como trabajo futuro una implementación con cargas parciales, la cual permite almacenar y actualizar información de forma eficiente en grandes conjuntos de datos.

#### **1. Tabla dimensión Estudiante**

En la Figura 34 se presenta el cargue de la dimensión Estudiante que se describe en el siguiente flujo:

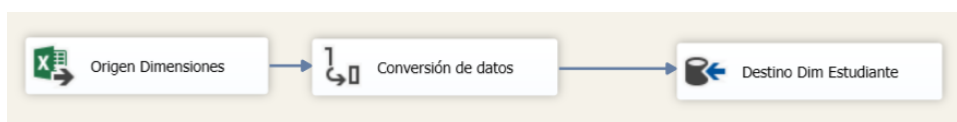
**Extracción:** se obtienen los datos de la Fuente de Excel proporcionada por la Universidad con las estimaciones calculadas en SQL Server.

**Transformación:** se realizó una revisión de SQL Server sin encontrar valores inconsistentes, razón por la cual la transformación necesaria consistió en la conversión de datos al formato de destino.

**Carga:** la salida del flujo de transformación es asociada con la dimensión Estudiante de la DW.

### Figura 34.

*Cargue de Dimensión Estudiante*



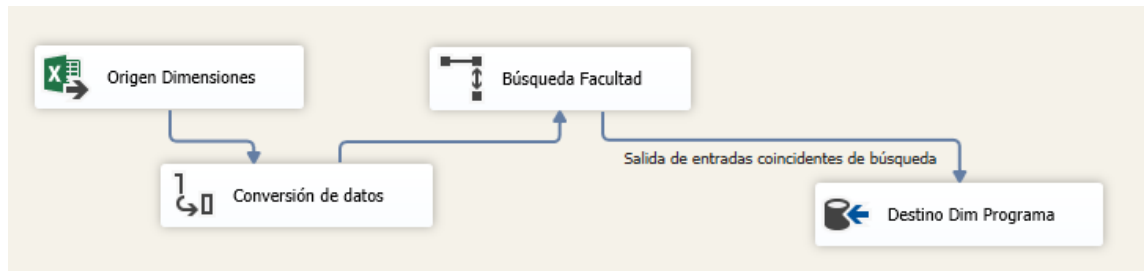
## 2. Tabla dimensión Programa

En la Figura 35 se presenta el cargue de la dimensión Programa que se describe en el siguiente flujo:

**Extracción:** se obtienen los datos de la Fuente de Excel proporcionada por la Universidad con las estimaciones calculadas en SQL Server.

**Transformación:** se realizó una revisión de SQL Server sin encontrar valores inconsistentes, razón por la cual la transformación necesaria consistió en la conversión de datos al formato de destino y la adquisición de la llave foránea de la Dimensión Facultad (cuyo cargue se presenta en el ANEXO C).

**Carga:** la salida del flujo de transformación es asociada con la dimensión Programa de la DW.

**Figura 35.***Cargue de Dimensión Programa***3. Tabla de Hechos Matrícula Estudiantil**

En la Figura 36 se presenta el cargue de la tabla de hechos Matrícula Estudiantil que se describe en el siguiente flujo:

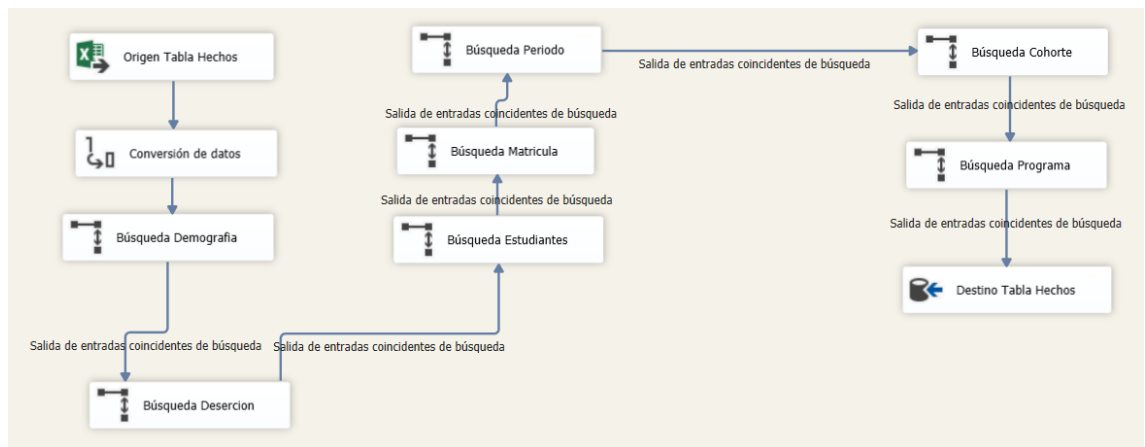
**Extracción:** mediante una sentencia SQL se combinan las tablas de las dimensiones y se exportan los datos en a las tablas de Excel que son leídas en el proceso de extracción.

**Transformación:** se realizó una revisión de SQL Server sin encontrar valores inconsistentes, razón por la cual la transformación necesaria consistió en la conversión de datos al formato de destino. También se llaman las llaves foráneas de los Cargues de dimensión.

**Carga:** la salida del flujo de transformación es asociada con la Tabla de Hechos de la DW.

**Figura 36.**

*Cargue de Tabla de Hechos*



**Población y validación de datos**

Después de finalizar la etapa de paquetes de cargue, se ejecutaron para poblar la DW. Se ejecuta cada dimensión teniendo precaución de dar prioridad a la dimensión de facultad por ser un sub-dimensión de la que depende la dimensión programa; luego de cargar las dimensiones se ejecuta el cargue de la Tabla de Hechos. Finalmente, se validó la consistencia de los datos a partir de hacer comparaciones con los sistemas fuente.

**5.1.2. Desarrollo del Front Room**

En esta sección se describen las actividades realizadas para generar la visualización de los reportes y las consultas.

**Identificación y priorización de reportes candidatos**

Se seleccionan los reportes relacionados con los requerimientos que cumplieron con la disponibilidad de datos descrita en la sección de Diseño físico de la base de datos. Los requerimientos que se van a presentar al usuario se pueden observar en la Tabla 65,

han sido priorizados de acuerdo a la disponibilidad de datos y el proceso selección llevado a cabo en la etapa de requerimientos del CICLO 3 del Capítulo 3

**Tabla 65.**

*Requerimientos para reportes y consultas*

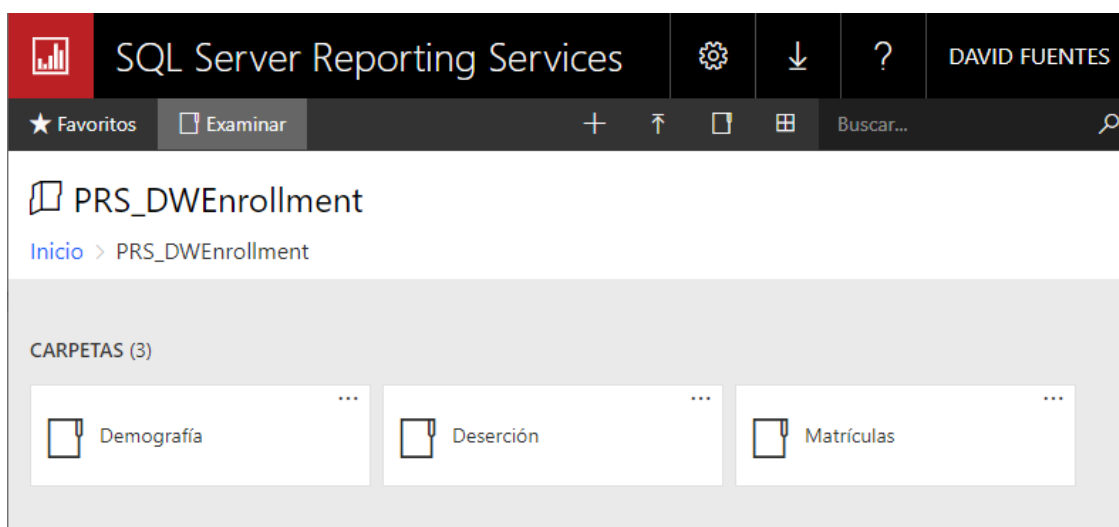
<b>PROCESOS ACADÉMICOS: MATRÍCULA ESTUDIANTIL</b>	
a)	<p><i>Aplicación equitativa de los criterios para la admisión y permanencia de los estudiantes</i></p> <p><b>Requerimiento para reporte:</b> Número de estudiantes de primera matrícula por estrato, nivel de programa, sede, modalidad, periodo y año.</p>
b)	<p><i>Estrategias que garanticen la integración de los estudiantes a la institución en consideración a su heterogeneidad social y cultural</i></p> <p><b>Requerimiento para reporte:</b> porcentaje de estudiantes por estrato y periodo.</p>
c)	<p><i>Deserción de estudiantes, análisis de causas y estrategias de permanencia en condiciones de calidad</i></p> <p><b>Requerimiento para reporte 1:</b> Porcentaje de deserción por motivo, facultad, nivel de programa, sede, modalidad, genero, periodo, cohorte y año.</p> <p><b>Requerimiento para reporte 2:</b> Número de estudiantes desertores por motivo, facultad, nivel de programa, sede, periodo y cohorte.</p> <p><b>Requerimiento para reporte 3:</b> Porcentaje de retención por periodo y cohorte.</p>

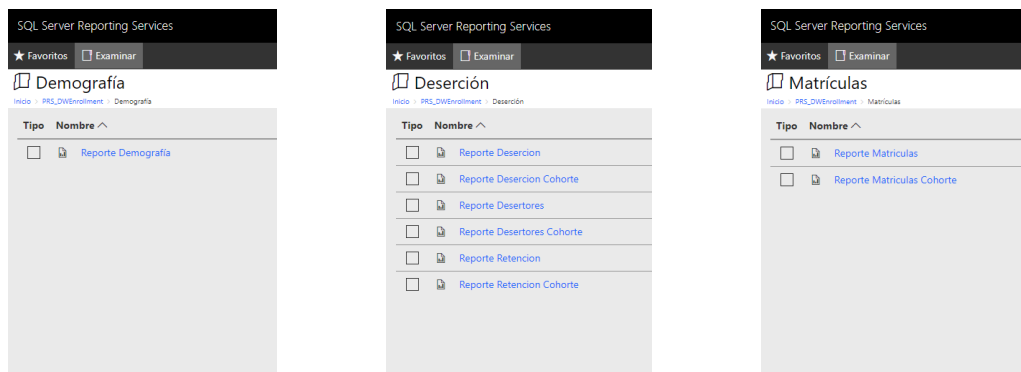
**Diseño de la estrategia de navegación**

En esta fase se agrupan los reportes creando una carpeta por tema, tal y como se visualiza en la Figura 37.

**Figura 37.**

*Reportes organizados por tema*





Las consultas Ad-hoc se visualizan desde un archivo Excel permitiendo al usuario arrastrar las medidas y dimensiones que vaya necesitando según sus necesidades analíticas.

### Estándar para la aplicación de Usuario Final

Los reportes mantienen una estructura estándar que se puede observar en la Selección del enfoque de implementación

Se selecciona el portal web del servidor de informes de SQL Server para los reportes, en el caso de las consultas Ad-hoc se utilizó Excel debido a que es una herramienta con que la gran mayoría de usuarios ya se encuentran familiarizados.

### Desarrollo de la aplicación de Usuario Final

Realizar la aplicación de Usuario Final requirió los siguientes pasos:

#### Creación de cubo

La creación del cubo se desarrolló con SQL Analysis Services (SSAS), herramienta definida en el CICLO 3 del Capítulo 3. Se configura la conexión con el origen de datos de la DW para crear las vistas y medidas calculadas correspondientes, conforme a los datos disponibles se crean las medidas calculadas de deserción y retención, las cuales

quedan relacionadas con el estado de deserción (no deserta, deserta en nivel 1, deserta en nivel 2...) establecido en la Dimensión Deserción.

*Figura 38*, la cual se compone de los siguientes elementos:

**Cabecera:** logos de los grupos de investigación STI con el escudo de la Universidad Industrial de Santander y GTI con el escudo de la Universidad del Cauca. También se presenta el título correspondiente al reporte.

**Cuerpo:** se compone de los siguientes elementos:

- **Tabla:** se presenta una tabla con los grupos individuales en las filas y el periodo académico o cohorte en las columnas.
- **Gráfica:** se muestra una gráfica que resumen de los datos mostrados en la tabla.
- **Parámetros:** se permite la selección de variables como la facultad, el programa o un intervalo específico de periodos en que se desee analizar la información.

### **Selección del enfoque de implementación**

Se selecciona el portal web del servidor de informes de SQL Server para los reportes, en el caso de las consultas Ad-hoc se utilizó Excel debido a que es una herramienta con que la gran mayoría de usuarios ya se encuentran familiarizados.

### **Desarrollo de la aplicación de Usuario Final**

Realizar la aplicación de Usuario Final requirió los siguientes pasos:

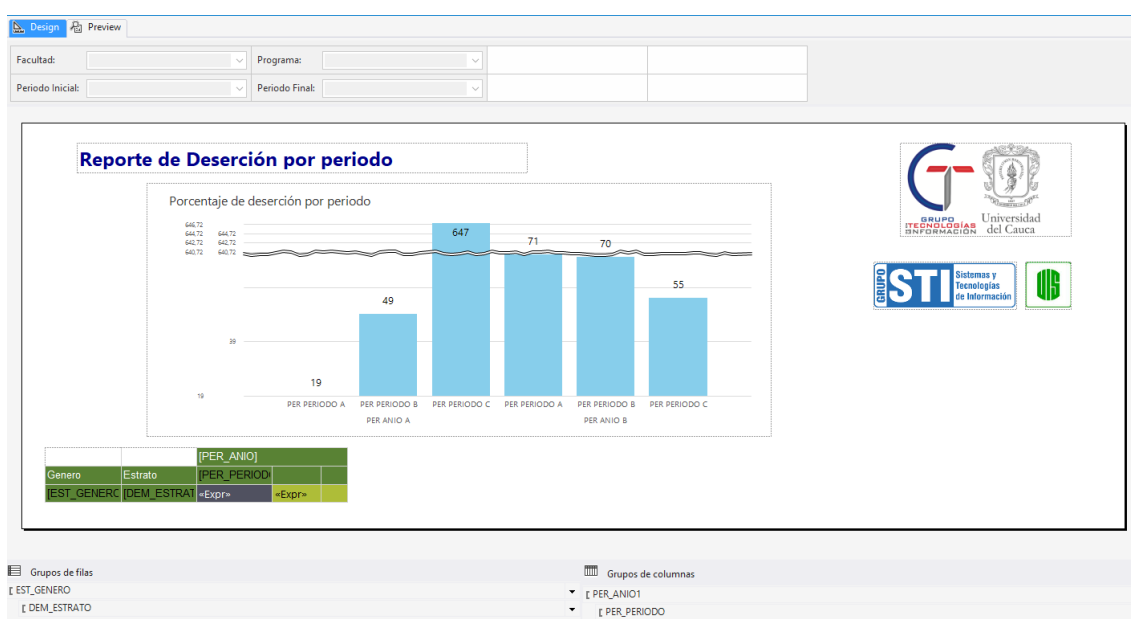
#### **Creación de cubo**

La creación del cubo se desarrolló con SQL Analysis Services (SSAS), herramienta definida en el CICLO 3 del Capítulo 3. Se configura la conexión con el origen de datos de la DW para crear las vistas y medidas calculadas correspondientes, conforme

a los datos disponibles se crean las medidas calculadas de deserción y retención, las cuales quedan relacionadas con el estado de deserción (no deserta, deserta en nivel 1, deserta en nivel 2...) establecido en la Dimensión Deserción.

**Figura 38.**

*Esquema de Reportes de Aplicación de Usuario*



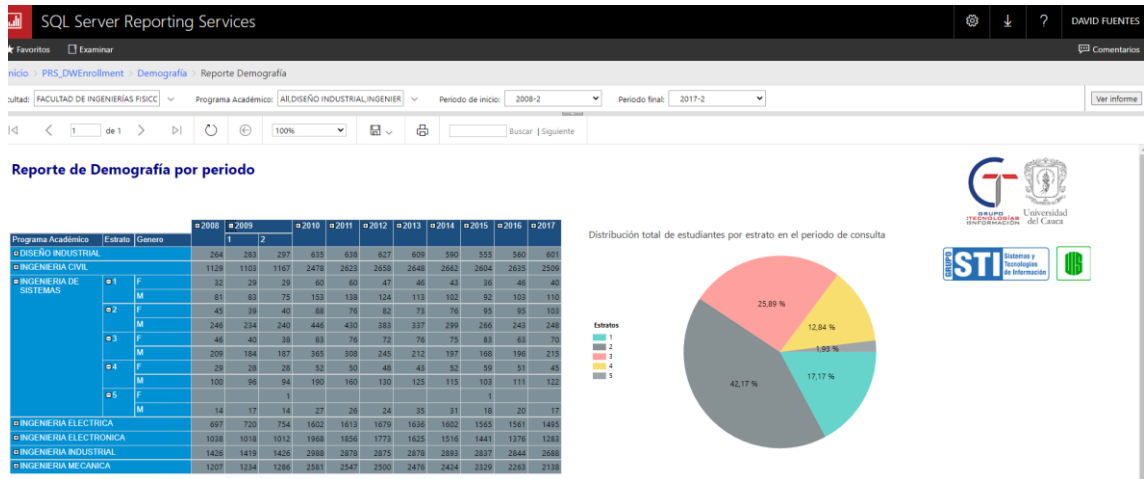
**Creación de reportes estándar**

Los reportes fueron diseñados en SQL Reporting Services (SSRS), herramienta definida en el CICLO 3 del Capítulo 3. Se configura la conexión con el cubo y se sigue el Estándar para la aplicación de Usuario Final que fue establecido con anterioridad.

Se presenta a continuación uno de los reportes OLAP que fueron definidos a partir de los requerimientos, el cual corresponde a la demografía de las matrículas, se muestra en la Figura 39 el despliegue de la carrera Ingeniería de sistemas, con el respectivo estrato y genero de los estudiantes; así mismo se despliegan los periodos para el año 2009.

**Figura 39.**

*Reporte de Demografía*

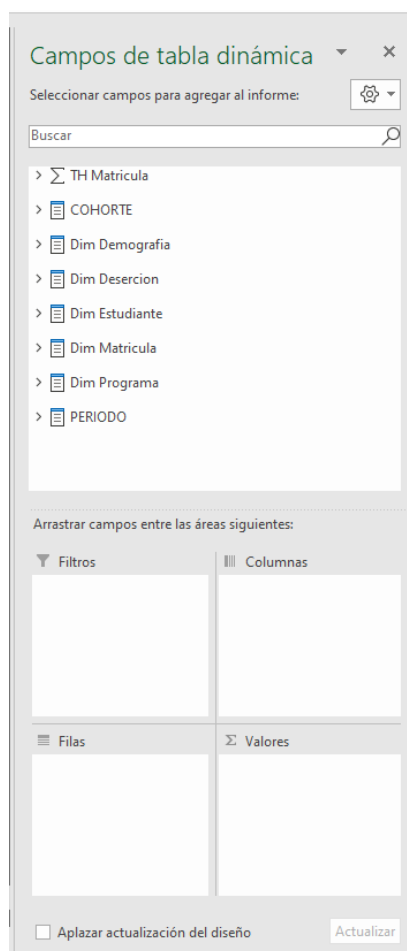


**Creación de consultas Ad-hoc**

Las consultas fueron diseñadas consultando el cubo desde Excel con tablas dinámicas, herramienta también definida en el CICLO 3 del Capítulo 3. Las jerarquías y atributos del cubo dimensional se presentan desde Excel en la Figura 40. En la Figura 41 se presenta una consulta Ad-hoc que tiene los desertores por periodo, presentando en particular los valores para matriculas regulares.

**Figura 40.**

*Campos de tabla dinámica*



**Figura 41.**

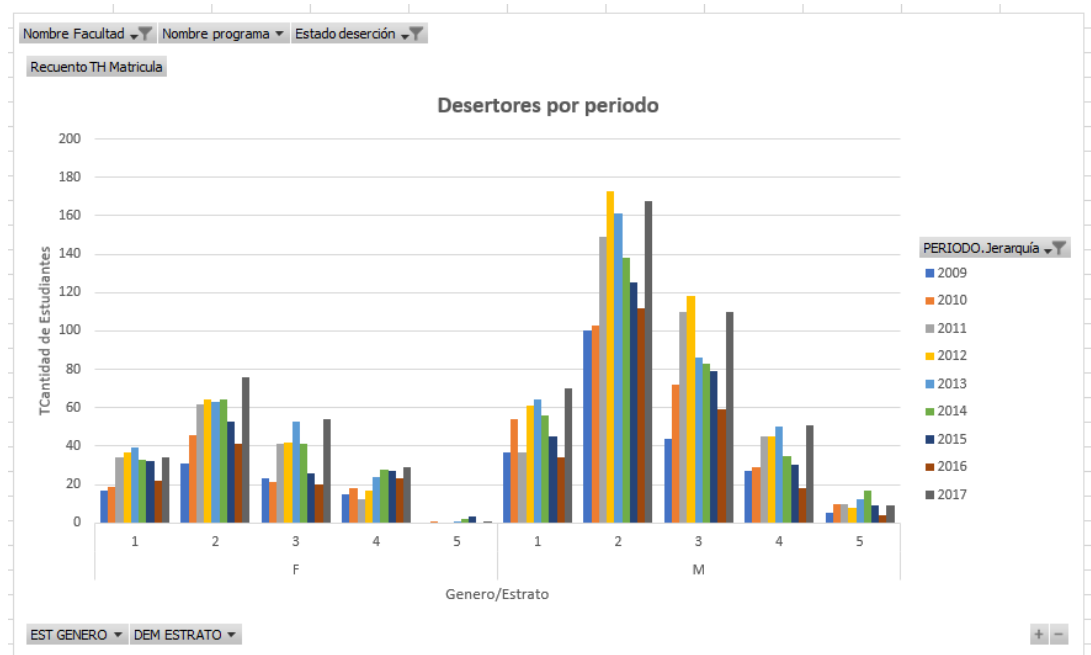
*Tabla de desertores por periodo*

Desertores por periodo																					
Nombre Facultad	FACULTAD DE INGENIERÍAS FISICOMECAÑICAS																				
Nombre programa	All																				
Estado deserción	(Varios elementos)																				
Recuento TH Matricula	Año-Periodo																				
Genero	2009		2010		2011		2012		2013		2014		2015		2016		2017		Total general		
F																					
1	17	7	12	34	37	39	33	32	22	34										267	
2	31	27	19	62	64	63	64	53	41	76										500	
3	23	12	9	41	42	53	41	26	20	54										321	
4	15	7	11	12	17	24	28	27	23	29										193	
5			1			1	2	3			1										8
M																					
1	37	21	33	37	61	64	56	45	34	70										458	
2	100	52	51	149	173	161	138	125	112	168										1229	
3	44	27	45	110	118	86	83	79	59	110										761	
4	27	15	14	45	45	50	35	30	18	51										330	
5	5	5	5	10	8	12	17	9	4	9										84	
Total general	299	173	200	500	565	553	497	429	333	602										4151	

La Figura 42 muestra el grafico de barras de la tabla dinámica, donde se observa que un alto conjunto de estudiantes

**Figura 42.**

*Desertores por año*



## **Creación de procedimientos de mantenimiento para la aplicación de Usuario Final**

Esta actividad no se realizó debido a que SSRS permite tanto cambiar las fuentes de datos como crear, eliminar y actualizar las plantillas.

En esta sub-fase se realiza el prototipo de la sección 5.1 con los reportes solicitados en los requerimientos, se presenta en la Figura 13 y se divide en las actividades: *Identificar y realizar reportes candidatos, Diseñar la estrategia de navegación, Desarrollar estándares para las aplicaciones de Usuario Final, Seleccionar un enfoque de implementación, Desarrollar las aplicaciones de Usuario Final, Desarrollar procedimientos de mantenimiento de aplicaciones de Usuario Final y Desarrollar documentación de las aplicaciones.*

### **5.1.3. Integración**

Se realiza la integración del Back Room y Front Room, a partir de los siguientes pasos:

#### **5.3.1 Procesar el cubo**

Esta actividad carga los datos de la base de datos relacional.

#### **5.3.2 Conexión al cubo y a la DW con la herramienta de reportes**

En esta actividad se conectan los reportes del servidor informes con los datos de la fuente OLAP o relacional.

#### **5.3.3 Pruebas sobre el prototipo**

Se ejecutan consultas sobre la DW para identificar inconsistencias, haciendo comparaciones con los datos del sistema fuente.

#### **5.3.4 Presentación del prototipo a los usuarios**

Se presenta el prototipo al usuario explicando el manejo de las consultas Ad-hoc en Excel y mostrando los reportes de SRSS.

### **5.3.5 Proceso de aceptación de usuario**

Se realizó un test de aceptación, elaborado con base en la característica de satisfacción, Modelo para la calidad en uso de la norma ISO/IEC 25022. Esta fase será detallada más adelante en la sección de Evaluación de la utilidad del prototipo .

#### **5.1.4. Despliegue**

Esta etapa no fue realizada debido a que el prototipo no cuenta con la totalidad de los datos reales, sin embargo, se presenta la herramienta como un instrumento de análisis útil para dar soporte a los procesos de acreditación institucional con el uso de la totalidad de los datos disponibles por la Universidad. Las pruebas fueron realizadas en una máquina local personal.

## **5.2. Evaluación de la utilidad del prototipo**

### **5.2.1. Evaluación del prototipo**

Se describe el proceso de evaluación del prototipo de bodegas de datos realizado para medir el nivel de satisfacción de los usuarios finales con los reportes estándar y las consultas Ad-hoc del prototipo de matrícula, mediante la aplicación de un test de nivel de satisfacción.

Primero se realiza una descripción acerca del objetivo de la evaluación, seguido de la normativa, descripción y aplicación del test; luego se describe el proceso de evaluación y se presentan los resultados obtenidos.

### **Objetivo de la evaluación**

La finalidad es medir el grado de satisfacción del usuario con los reportes estándar y consultas analíticas ad-hoc de la Bodega de Datos para procesos académicos en el área de Matricula Estudiantil.

### **Normativa**

La norma ISO (Organización Internacional de Estandarización) e IEC (Comisión Electrotécnica Internacional) conforman un sistema especializado en estandarización internacional el cual en la norma ISO/IEC 25022 mide la calidad con base en el valor práctico de un producto software con un enfoque en medidas de utilidad que evalúen el grado en que un usuario está satisfecho con el logro percibido de los objetivos pragmáticos, incluidos los resultados y consecuencias del uso. (International Standard Organization & International Electrotechnical Commission, 2016). Debido a que no se pudo realizar el despliegue en la Universidad ni se requirió escribir un código fuente que requiriera un análisis se utilizó la norma antes mencionada para centrarse en medir la calidad de uso enfocándose en la sub-característica utilidad.

Según la norma ISO/IEC 25022 se mide la utilidad en los cinco enfoques de la Tabla 66, este trabajo evaluará el prototipo en el enfoque de satisfacción de usuario (señalado en color verde) debido a que los otros enfoques están estrechamente relacionados con las fases de despliegue, que como se describió anteriormente se salen del alcance del proyecto.

**Tabla 66.**

*Medidas de utilidad (International Standard Organization & International Electrotechnical Commission, 2016)*

Nombre	Descripción	Método
Satisfacción con las características	La satisfacción del usuario con características específicas del sistema	Cuestionario
<b>Nota:</b> suele ser un cuestionario que utiliza la escala de Likert. Si los elementos del cuestionario se combinan para dar una puntuación general, se pueden ponderar, ya que diferentes preguntas pueden tener diferente importancia.		
Uso Discrecional	La proporción de usuarios potenciales que eligen usar un sistema o función	Medir el comportamiento del usuario o la recopilación automatizada de datos
Utilización de características	La proporción de un conjunto identificado de usuarios del sistema que utilizan una característica particular	Medir el comportamiento del usuario o la recopilación automatizada de datos
Proporción de usuarios que se quejan	La proporción de usuarios que presentan quejas	Medir el comportamiento del usuario
Proporción de quejas de los usuarios sobre una función en particular	La proporción de quejas de los usuarios sobre una función en particular	Medir el comportamiento del usuario

**Descripción**

Para la validación del prototipo se utiliza un cuestionario con escala de Likert siguiendo lo establecido en la medida de utilidad de satisfacción de usuario presentada en la Tabla 66. El test se compone de tres secciones:

- **Primera sección:** se describe el objetivo de la evaluación.
- **Segunda sección:** se presenta la evaluación de los reportes estándar con preguntas formuladas para evaluar la importancia analítica y la facilidad de navegación.
- **Tercera sección:** se presenta la evaluación de las consultas Ad-hoc con preguntas formuladas para evaluar la importancia analítica y la facilidad de navegación.

Siguiendo la instrucción de la norma ISO/IES 25022 que se presenta en la Tabla 66 se utiliza la escala de Likert, la cual se describe en la Tabla 67.

**Tabla 67.**

*Escala de Likert*

<b>Valor numérico</b>	<b>Conformidad</b>
5	Totalmente de acuerdo
4	Parcialmente de acuerdo
3	Indiferente
2	Algo en desacuerdo
1	Totalmente en desacuerdo

**Aplicación del test de nivel de satisfacción**

El test fue resuelto por dos funcionarios UIS, cuya descripción se detalla en la Tabla 68.

**Tabla 68.**

*Usuarios evaluadores*

<b>ID</b>	<b>Ocupación</b>	<b>Experiencia relacionada con acreditación</b>
U1	Coordinadora de Posgrados de la Escuela de Ingeniería de Sistemas e Informática – Universidad Industrial de Santander.	Participación en comités de acreditación del CNA
U2	Coordinadora Académica – Universidad Industrial de Santander.	Líder de competencia en Comité de Renovación de Acreditación

Se acuerda una reunión con cada usuario, tanto para los reportes estándar como para las consultas ad-Hoc se hace una capacitación previa en la que se explica el funcionamiento del sistema, se realizan consultas ideadas por el usuario y posteriormente se brinda el espacio para llenar el test de satisfacción.

**Proceso de evaluación**

El proceso de evaluación se realiza siguiendo la norma ISO/IEC 25022 y el procedimiento definido por (Balseca Chisaguano, 2014) el cual establece la ecuación:

$$X = \frac{A}{B} \quad (1)$$

Donde  $X$  es la valoración otorgada por el evaluador,  $A$  el número de preguntas con respuestas satisfactorias y  $B$  es la cantidad de preguntas realizadas en el cuestionario.

La interpretación del resultado total obtenido de la ecuación 1 se realiza con base en la tabla formulada por (Balseca Chisaguano, 2014)

**Tabla 69.**

*Grado de satisfacción según nivel de satisfacción*

Escala de medición	Niveles de puntuación	Grado de satisfacción
87,5%-100%	Cumple con los requisitos	Muy satisfactorio
50%-87,4%	Aceptable	Satisfactorio
27,5%-49%	Mínimamente aceptable	Insatisfactorio
0%-27,4%	Inaceptable	

**Resultados obtenidos**

El resultado se basa en la cantidad de respuestas del funcionario, que se presenta en el ANEXO D. En la Tabla 70 se presenta el total de preguntas evaluadas y las respuestas obtenidas.

**Tabla 70.**

*Distribución de respuestas obtenidas en el test*

Funcionario	Cantidad de preguntas	Calificación				
		5	4	3	2	1
U1	35	29	6			
U2	35	35				
<b>Total</b>	<b>70</b>	<b>64</b>	<b>6</b>			

Con base en la matriz de calidad establecida por (Balseca Chisaguano, 2014) se realiza una adaptación a la característica de utilidad. Para analizar el nivel de satisfacción se tienen en cuenta las respuestas Totalmente de acuerdo (5, que corresponde a un 91,43%) y Parcialmente de acuerdo (4, que corresponde a un 8,57%), donde  $X_{MS1}$

representa la razón de respuestas en 5 y  $X_{S2}$  la razón de respuestas en 4. Los resultados se presentan en la tabla.

**Tabla 71.**

*Matriz de calidad. Adaptado de (Balseca Chisaguano, 2014) y (Tapia & Cruz, 2020)*

Métrica	Formula	Valor obtenido por variable	Valor total obtenido	Valor deseado
Nivel de satisfacción	$X_{MS1} = \frac{A_1}{B}$	$A_1 = 64$ $B = 70$ $X_{MS1} = 0,9143$	$X = X_{MS1} + X_{S2}$ $X = 1 \rightarrow 100\%$	1
	$X_{S2} = \frac{A_2}{B}$	$A_2 = 6$ $B = 70$ $X_{S2} = 0,0857$		

El resultado de la validación mostrado en la Tabla 71 es de 1, lo cual según la Tabla 69 corresponde a un grado de satisfacción de 100% equivalente a **Muy Satisfactorio**.

De esta manera se evidencia que los reportes estándar y las consultas analíticas definidas en el proyecto son importantes para cumplir necesidades analíticas y representan un apoyo en los procesos de toma de decisiones institucionales.

### Conclusiones

- Los procesos académicos, investigativos y de seguimiento a graduados han sido abordados con soporte de Bodegas de Datos (DW) en ocho tópicos base: evolución de los programas, investigación, calidad docente, admisiones, permanencia, rendimiento estudiantil, impacto social de la institución e impacto social de los graduados. Estos temas son tenidos en cuenta en los cinco factores de acreditación institucional que fueron abordados en el proyecto.

- Se determina el 38,5% de la literatura publicada en revistas reconocidas por en el JCR/SJR presenta el tema de DW para IES de una forma general, sin profundizar en algún tópico específico. Las particularidades abordadas en estos estudios se dividen en tres áreas, que son: dificultades (25%), recomendaciones (25%) e innovaciones (50%). Este resultado muestra que la mayoría de los estudios de este tipo están apuntando a la innovación, agregando conceptos como gestión de conocimiento, habilidades de negocio, Software como Servicio, entre otros.
- La literatura de DW para IES puede ser agrupada en procesos académicos (se menciona en el 60,5% de los estudios), investigativos (7,9%), y de seguimiento a graduados (13,2%). La literatura sobre investigación y el seguimiento a graduados es escasa, situación que brinda la oportunidad de hacer estudios enfocados en estos procesos.
- La mayoría (el 61,5%) de los trabajos sobre procesos académicos (relacionados con los factores 2, 3 y 4 del CNA), investigativos (factor 6) y de seguimiento a graduados (factor 7), presentan falencias relacionadas con el nivel de detalle de los componentes (dimensiones, atributos y medidas) del modelo.
- Se proponen siete modelos para procesos académicos: dos para el factor estudiantes, dos para el factor profesores y tres para el factor de procesos académicos. Para investigación se proponen seis, para graduados un modelo y otro más para encuestas. Se incluyen los requerimientos identificados de los procesos de acreditación, los propuestos a partir de los lineamientos del CNA y los sugeridos por las expertas en acreditación.
- Los modelos fueron validados por un grupo focal conformado por tres expertos en Bodegas de Datos y Bases de Datos siguiendo la norma ISO-25010: Medidas

de calidad del producto, definida en la sub-característica adaptabilidad, los expertos consultados calificaron en un 100% (totalmente de acuerdo y de acuerdo) que los modelos se adaptan a los requerimientos menos frecuentes, más frecuentes y propuestos, y también en un 100% que los modelos permiten apoyar la toma de decisiones.

- Se construye un prototipo basado en el modelo de procesos académicos relacionado con matrícula estudiantil el cual permite visualizaciones por consultas Ad-hoc y reportes estándar.
- El prototipo se valida por funcionarios UIS encargados de toma de decisiones relacionadas con calidad educativa por medio de un test de nivel de satisfacción con base en la métrica ISO/IEC 25022: Medidas de calidad de uso, definida para la sub-característica utilidad, obteniendo una calificación de 100% (muy satisfactorio).
- Este trabajo puede sentar las bases para la implementación de tecnologías de bodegas de datos aplicadas a los subsecuentes decretos de acreditación institucional que implemente el CNA. También, puede aumentar su alcance en pruebas que involucren más expertos e Instituciones, así como ser una herramienta que permita mejorar la calidad de la educación superior a nivel nacional e internacional.

### **Referencias Bibliográficas**

ACOFI, & Confedi. (2018). *Aseguramiento de la calidad y mejora de la educación en ingeniería: Experiencias en América Latina* (L. A. Gonzalez Araujo, Ed.; 1st ed.).

- Alkhalil, A., Abdallah, M. A. E., Alogali, A., & Aljaloud, A. (2021). Applying big data analytics in higher education: A systematic mapping study. *International Journal of Information and Communication Technology Education*, 17(3), 29–51.  
<https://doi.org/10.4018/IJICTE.20210701.0a3>
- Amazon Web Services. (2022). *Conceptos relacionados con el almacenamiento de datos*. <https://aws.amazon.com/es/data-warehouse/>
- B Akki, S., & M. N, V. (2018). Design of Dashboard for University Examination Result Analysis System. *3rd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2018*.
- Balseca Chisaguano, E. A. (2014). *Evaluación De Calidad De Productos Software En Empresas De Desarrollo De Software Aplicando La Norma Iso/Iec 25000*. 209.
- Berges, A., Ramirez, P., Pau, I., Tejero, A., & Crespo, A. G. (2021). A Framework for Strategic Intelligence Systems Applied to Education Management: A Pilot Study in the Community of Madrid. *IEEE Access*, 9, 1–1.  
<https://doi.org/10.1109/access.2021.3081734>
- Boulila, W., Al-kmal, M., Farid, M., & Mugahed, H. (2018). A business intelligence based solution to support academic affairs: case of Taibah University. *Wireless Networks*, 3. <https://doi.org/10.1007/s11276-018-1880-3>
- Camilo, W., & Mera, L. (2022). *Modelo Dimensional para Proyectos y Grupos de Investigación de la VRI de la Universidad del Cauca*. Universidad del Cauca.
- CESU. (2014). Acuerdo 03 de 2014. In *Consejo Nacional de Acreditación (CNA)* (p. 32).

- Chandra, E., Suganda Girsang, A., Hadinata, R., & Muhamad Isa, S. (2018). Analysis Students' Graduation Eligibility Using Data Warehouse. *2018 International Conference on Information Management and Technology (ICIMTech)*.
- Clarivate. (2022). *InCites Journal Citation Reports*.  
<https://error.incites.clarivate.com/error/Error?DestApp=IC2JCR&Error=IPError&Params=DestApp%3DIC2JCR&RouterURL=https%3A%2F%2Flogin.incites.clarivate.com%2F&Domain=.clarivate.com&Src=IP&Alias=IC2>
- Combata Niño, H. A., Cómbita Niño, J. P., & Morales Ortega, R. (2020). Business intelligence governance framework in a university: Universidad de la costa case study. *International Journal of Information Management*, 50(October 2018), 405–412. <https://doi.org/10.1016/j.ijinfomgt.2018.11.012>
- Dadkhah, M., & Lagzian, M. (2018). A research on the use of business intelligence for academic research. *Library Hi Tech News*, 35(3), 10–12.  
<https://doi.org/10.1108/LHTN-03-2018-0016>
- Dingsøyr, T., & Dybå, T. (2008). *Empirical Studies of Agile Software Development : A Systematic Review* (Issue August). <https://doi.org/10.1016/j.infsof.2008.01.006>
- Fuentes-Vargas, D.-A., Soraca-Higuera, J.-A., Cobos-Lozada, C.-A., Mendoza-Becerra, M.-E., & Gómez-Flórez, L.-C. (2019a). Una Revisión de Bodegas de Datos para Educación Superior. *Revista Ibérica de Sistemas y Tecnologías de Informação*, 33, 309–322.
- Fuentes-Vargas, D.-A., Soraca-Higuera, J.-A., Cobos-Lozada, C.-A., Mendoza-Becerra, M.-E., & Gómez-Flórez, L.-C. (2019b). Una Revisión de Bodegas de Datos para Educación Superior. *Revista Ibérica de Sistemas y Tecnologías de Informação*, 33, 309–322.

- González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige : The SJR indicator. *Journal of Informetrics*, 4, 379–391. <https://doi.org/10.1016/j.joi.2010.03.002>
- Google. (2022). *Temas Google Cloud*. <https://cloud.google.com/learn/what-is-etl?hl=es>
- Grech, P. (2013). Introducción a la Ingeniería: un enfoque a través del diseño. In *PEARSON* (2nd ed.).
- Inmon, W. H. (2005). *Building the Data Warehouse* (4th ed., Vol. 13, Issue 401). Wiley Publishing. <https://doi.org/10.1017/CBO9781107415324.004>
- International Scientific Institute. (2018). *IMPACT FACTOR LIST*. [SCIJOURNAL.ORG](https://scijournal.org/about-us.html). <https://scijournal.org/about-us.html>
- International Standard Organization, & International Electrotechnical Commission. (2016). *International Standard ISO/IEC 25022. 2010*.
- Irfan, M., Zulfikar, W. B., Alam, C. N., Saadillah, D., & Fuadi, R. S. (2018). Conceptual model of executive information system data (A Case Study at the State Islamic University of Sunan Gunung Djati Bandung). *IOP Conference Series: Materials Science and Engineering*, 434(1). <https://doi.org/10.1088/1757-899X/434/1/012256>
- ISO/IEC. (2011). *International Standard ISO/IEC 25010*.
- Jayashree, G., & Priya, C. (2020). Comprehensive Guide to Implementation of Data Warehouse in Education. *Intelligent Computing and Innovation on Data Science*, 1–8.
- Jiménez, J., Yanza, N., & Mendoza, M. (2018). *Modelo Dimensional para Evaluación Docente de la Universidad del Cauca*. Universidad del Cauca.

- Joshua, S. R., & Moge, T. (2020). Agile analytics: Adoption framework for business intelligence in higher education. *Journal of Theoretical and Applied Information Technology*, 98(7), 1032–1042.
- Khan, A. (2012). *Business Intelligence & Data Warehousing Simplified: 500 Questions, Answers and Tips*. David Pallai.
- Khan B, A., Ghosh, S., & Ghosh, S. K. (2018). eDWaaS : A Scalable Educational Data Warehouse as a Service. *International Conference on Intelligent Systems Design and Applications, ISDA*, 998–1007. <https://doi.org/10.1007/978-3-319-76348-4>
- Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling* (3d ed.). John Wiley & Sons, Inc.
- Kitchenham, B., & Brereton, P. (2013). A systematic review of systematic review process research in software engineering. *Information and Software Technology*, 55(12), 2049–2075. <https://doi.org/10.1016/j.infsof.2013.07.010>
- Kitchenham, B., Charters, S., & Budgen, D. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. In *Keele University*.
- Kontio, J., Bragge, J., & Lehtola, L. (2008). The focus group method as an empirical tool in software engineering. *Guide to Advanced Empirical Software Engineering, 2004*, 93–116. [https://doi.org/10.1007/978-1-84800-044-5\\_4](https://doi.org/10.1007/978-1-84800-044-5_4)
- Lapura, E. V. F., Kenneth, J., Fernandez, J., Jonathan, M., Pagatpat, K., Dante, D., Lapura, E. V. F., Dinawanao, J. K. J., Jonathan, M., Pagatpat, K., & Dante, D. (2018). Development of a University Financial Data Warehouse and its Visualization Tool. *Procedia Computer Science*, 135, 587–595. <https://doi.org/10.1016/j.procs.2018.08.229>

- Ley 30. (1992). Ley 30 de Diciembre 28 de 1992. *Fundamentos de La Educación Superior*, Art. 30.
- Maia, A., Portela, F., & Santos, M. F. (2018). Web intelligence in higher education: A study on the usage of business intelligence techniques in education. *Proceedings - 2018 IEEE 6th International Conference on Future Internet of Things and Cloud Workshops, W-FiCloud 2018*, 176–181. <https://doi.org/10.1109/W-FiCloud.2018.00034>
- Mao, S. (2010). Academic Data Warehouse in Oracle. In *Interagir: pensando a extensão* (Vol. 0, Issue 15). Southern Connecticut State University.
- Martins, J., Branco, F., Gonçalves, R., Au-Yong-Oliveira, M., Oliveira, T., Naranjo-Zolotov, M., & Cruz-Jesus, F. (2018). Assessing the success behind the use of education management information systems in higher education. *Telematics and Informatics*, 38, 182–193. <https://doi.org/10.1016/j.tele.2018.10.001>
- Medina, F., Fariña, F., & Castillo-Rojas, W. (2018). Data Mart para obtención de indicadores de productividad académica en una universidad Data Mart to obtain indicators of academic productivity in a university. *Ingeniare. Revista Chilena de Ingeniería*, 26, 88–101. <https://doi.org/http://dx.doi.org/10.4067/S0718-33052018000500088>
- Mendoza, M. E., Durán, L., & Rivera, N. (2010). MBD 1.0. Metodología de Desarrollo de Bodegas de Datos Para Micro, Pequeñas y Medianas Empresas. *UIS Ingenierías*, 9(1), 85–101.
- Mendoza-Moreno, M., González-Serrano, C., & Pino, F. J. (2013). Focus group como proceso en Ingeniería de Software: Una experiencia desde la práctica. *DYNA (Colombia)*, 80(181), 51–60.

- Microsoft. (2022). *Understand star schema and the importance for Power BI*.  
<https://docs.microsoft.com/en-us/power-bi/guidance/star-schema>
- Moscoso-Zea, O., Castro, J., Paredes-Gualtor, J., & Lujan-Mora, S. (2019). A Hybrid Infrastructure of Enterprise Architecture and Business Intelligence Analytics for Knowledge Management in Education. *IEEE Access*, 7(c), 38778–38788.  
<https://doi.org/10.1109/ACCESS.2019.2906343>
- Moscoso-Zea, O., Paredes-Gualtor, J., & Lujan-Mora, S. (2018). A Holistic View of Data Warehousing in Education. *IEEE Access*, 6, 64659–64673.  
<https://doi.org/10.1109/ACCESS.2018.2876753>
- Mulyawan, B., Sutrisno, T., & Christanti, V. (2022). *An Application Performance Measurement Model for Higher Education in Indonesia*.
- Musa, S., Ali, N. M., Miskon, S., Giro, M. A., & Aljabali, R. (2021). Business Intelligence usage model for Higher Education Institutions. *Journal of Theoretical and Applied Information Technology*, 99(5), 1020–1032.
- Najm, I. A., Dahr, J. M., Hamoud, A. K., Alasady, A. S., Awadh, W. A., Kamel, M. B. M., & Humadi, A. M. (2022). OLAP Mining with Educational Data Mart to Predict Students' Performance. *Informatica (Slovenia)*, 46(5), 11–19.  
<https://doi.org/10.31449/inf.v46i5.3853>
- Neamah, A. F. (2021). Adoption of Data Warehouse in University Management: Wasit University Case Study. *Journal of Physics: Conference Series*, 1860(1).  
<https://doi.org/10.1088/1742-6596/1860/1/012027>
- Ortiz, L., & Hallo, M. (2019). Analytical Data Mart for the Monitoring of University Accreditation Indicators. *2019 IEEE World Conference on Engineering Education (EDUNINE)*.

- Peralta, M. S. (2018). Business intelligence in e-learning, a case study of an ecuadorian university. *Proceedings - 13th Latin American Conference on Learning Technologies, LACLO 2018*, 29–32. <https://doi.org/10.1109/LACLO.2018.00016>
- Pérez-Pérez, Y. M., Rosado-Gómez, A. A., & Puentes-Velásquez, A. M. (2018). Application of business intelligence in the quality management of higher education institutions. *Journal of Physics: Conference Series*, 1126(1). <https://doi.org/10.1088/1742-6596/1126/1/012053>
- Pesonen, J., Fomkin, A., & Jokipii, L. (2018). *Building Data Science Capabilities into University Data Warehouse to Predict Graduation*. <http://hdl.handle.net/10138/298769>
- Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). *Systematic Mapping Studies in Software Engineering*.
- Pratt, K. S. (2009). Design Patterns for Research Methods: Iterative Field Research. *AAAI Spring Symposium: Experimental Design for Real, 1994*, 1–7.
- Rahutomo, R., Putri, R. A., & Pardamean, B. (2018). Building Datawarehouse for Educational Institutions in 9 Steps. *The 1st 2018 INAPR International Conference*.
- Scholtz, B., Calitz, A., & Haupt, R. (2018). A business intelligence framework for sustainability information management in higher education. *International Journal of Sustainability in Higher Education*, 19(2), 266–290. <https://doi.org/10.1108/IJSHE-06-2016-0118>
- Scimago Lab. (2022). *Scimago Journal & Country Rank*. <https://www.scimagojr.com/>
- Serasinghe, C. U., Jayakody, D. C. R., Dayananda, K. T. M. N., & Asanka, D. (2021). *Design and Implementation of Data Warehouse for a Higher Educational Institute in Sri Lanka*. 1–5. <https://doi.org/10.1109/i2ct51068.2021.9417820>

- Somya, R., Manongga, D., & Pakereng, M. A. I. (2018). Service-oriented Business Intelligence (SoBI) for Academic and Financial Data Integration in University. *2018 International Seminar on Application for Technology of Information and Communication (ISemantic)*.
- Sorour, A., Atkins, A. S., Stanier, C. F., & Alharbi, F. D. (2020). The Role of Business Intelligence and Analytics in Higher Education Quality : A Proposed Architecture. *2019 International Conference on Advances in the Emerging Computing Technologies (AECT)*, 0–5.
- Sutedja, I., Yudha, P., Khotimah, N., & Vasthi, C. (2018). Building A Data Warehouse to support Active Student Management : Analysis and Design. *2018 International Conference on Information Management and Technology (ICIMTech)*, 460–465.
- Tapia, I., & Cruz, L. (2020). *Modelo Dimensional para la División de Gestión de Salud Integral y Desarrollo Humano de la Universidad del Cauca*.
- UIAES-UIS. (2018). *UIS en Cifras*.  
<https://app.powerbi.com/view?r=eyJrIjoiOGIwZTkzZGQtNmI4MCM0OGVILTgwYjAtMGVjZmJlNGFIMmM3IiwidCI6ImE1ODRhZDMyLWRjZjYtNDE1MC1hNGI1LTdmYWZmOTI0OGFhNiIsImMiOiR9>
- United Nations. (2015). *Sustainable Development Goals*.  
<https://www.un.org/sustainabledevelopment/education/>
- University of York. (2014). *Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews*. <https://www.ncbi.nlm.nih.gov/books/NBK285222/>
- Urrea, D., Cerón, E., & Mendoza, M. (2018). *Modelo Dimensional para el Registro Académico de Estudiantes de la Universidad del Cauca*. Universidad del Cauca.

- Urrea Pito, D., Cerón Moreno, E., & Mendoza-Becerra, M.-E. (2020). Bodega de datos con alta capacidad de análisis para el desempeño académico de Universidades. *Investigación e Innovación En Ingenierías*, 8, 102–118.
- Venter, C., & Myburgh, I. (2018). Adapting a DW/BI Module for Gen-Z Students: An Action Design Research Study. *17th European Conference on E-Learning*.  
<https://www.researchgate.net/publication/328803393>
- Villegas-Ch, W., Luján-Mora, S., & Buenaño-Fernandez, D. (2018). Towards the integration of business intelligence tools applied to educational data mining. *II IEEE World Engineering Education Conference: The Role of Professional Associations in Contemporaneous Engineer Careers*.
- Yulianto, A. A., & Kasahara, Y. (2020a). Data Warehouse System for Multidimensional Analysis of Tuition Fee Level in Higher Education Institutions in Indonesia. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 11, Issue 6). [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)
- Yulianto, A. A., & Kasahara, Y. (2020b). Data Warehouse System for Multidimensional Analysis of Tuition Fee Level in Higher Education Institutions in Indonesia. *International Journal of Advanced Computer Science and Applications*, 11(6), 541–550. <https://doi.org/10.14569/IJACSA.2020.0110666>