



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

**MAESTRÍA EN GESTIÓN DE SISTEMAS DE INFORMACIÓN E
INTELIGENCIA DE NEGOCIOS**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO
DE MAGISTER EN GESTIÓN DE SISTEMAS DE INFORMACIÓN E
INTELIGENCIA DE NEGOCIOS**

**TEMA: “MODELO DE GESTIÓN DE INDICADORES DE EVALUACIÓN
PARA LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR”**

AUTORA: AUCANCELA GUAMÁN, MARGARITA ALEJANDRA

DIRECTOR: ING. RAURA RUIZ, JORGE GEOVANNY

SANGOLQUI

2019

CERTIFICADO DEL DIRECTOR**ESPE****UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA****VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA****CENTRO DE POSGRADOS****CERTIFICACIÓN**

Certifico que el trabajo de titulación, “MODELO DE GESTIÓN DE INDICADORES DE EVALUACIÓN PARA LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR” fue realizado por la señora maestrante Aucancela Guamán, Margarita Alejandra, el mismo que ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustente públicamente.

Sangolquí, 24 de julio de 2019

Una firma manuscrita en tinta azul que parece decir 'JGR' con un rasgo diagonal.

Jorge Geovanny Raura Ruiz

C.C.: 0501773063

AUTORÍA DE RESPONSABILIDAD**VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA****CENTRO DE POSGRADOS****AUTORÍA DE RESPONSABILIDAD**

Yo, Aucancela Guamán, Margarita Alejandra, con cédula de ciudadanía n°0602920845, declaro que el contenido, ideas y criterios del trabajo de titulación: "MODELO DE GESTIÓN DE INDICADORES DE EVALUACIÓN PARA LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR", es de mi autoría y responsabilidad, cumpliendo con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas. Consecuentemente el contenido de la investigación mencionada es veraz.

Sangolquí, 24 de julio de 2019

Firma manuscrita en tinta azul de Margarita Aucancela Guamán.

Margarita Alejandra Aucancela Guamán

C.C.: 060292084-5

AUTORIZACIÓN



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

VICERRECTORADO DE INVESTIGACIÓN, INNOVACIÓN Y
TRANSFERENCIA DE TECNOLOGÍA

CENTRO DE POSGRADOS

AUTORIZACIÓN

Yo, Aucancela Guamán, Margarita Alejandra, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: “MODELO DE GESTIÓN DE INDICADORES DE EVALUACIÓN PARA LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR”, en el Repositorio Institucional, cuyo contenido, ideas y criterios son de mi responsabilidad.

Sangolquí, 24 de julio de 2019

Una firma manuscrita en tinta azul que dice "Margarita Aucancela G".

Margarita Alejandra Aucancela Guamán

C.C.: 060292084-5

DEDICATORIA

A mi amado Dios:

Quien día a día, con lazos de amor y bondad inunda mi vida de sueños, esperanza y nuevas oportunidades para vivir y servir a la sociedad.

A mi amado esposo Celin:

Quien me ha demostrado su amor y paciencia, quien me ha motivado a obtener este nuevo logro, te amo mi bello sol de Diciembre.

A mi hijo Mateito:

Que cada día me sorprende con su imaginación, inteligencia y locuras, además de enseñarme el valor que tiene la familia, este título que he obtenido te lo dedico a ti hijo mío, porque tuve que dejarte tan pequeño para dedicarme a estudiar, te amo hijo mío.

A mis padres

Quienes me han enseñado valores como la perseverancia, la humildad, el respeto y la consideración a los demás.

A mis hermanos

Quienes me han apoyado con sus consejos y agradable compañía

A mis maestros de la ESPE

Quienes generosamente han compartido sus conocimientos en esta área de especialización como lo es la Inteligencia de Negocios, que Dios bendiga sus vidas y sigan adelante cosechando éxitos y logros que contribuyan al desarrollo de la sociedad ecuatoriana

A mis maestros y a las autoridades de la ESPOCH

Quienes contribuyeron a mi formación profesional, así como me han permitido a través del presente trabajo aportar con soluciones que agreguen valor al desarrollo institucional, así como su participación en la sociedad ecuatoriana como ente de formación profesional

AGRADECIMIENTO

A mi amado Dios

Gracias por todo lo que haces por mi, indudablemente soy tu hija y tú eres mi Padre.

A mi amado esposo Celin:

Gracias por tu apoyo incondicional, por tu dulzura, por tu paciencia y por todo el amor que me das.

A mi hijo Mateito:

Gracias por tu paciencia hijito mío, te amo mucho

A mis padres y hermanos

Gracias por apoyarme siempre

A mi tutor, Ing. Geovanny Raura y a mi coordinadora del programa: Ing. Tatiana Gualotuña

Gracias por su apoyo en todo este proceso, gracias por la confianza y el tiempo, son excelentes gestores académicos, sigan adelante.

A mis maestros y a las autoridades de la ESPOCH

Gracias por permitirme desarrollar el presente trabajo de investigación en la institución que me dio la vida profesional y ahora la vida laboral.

ÍNDICE DE CONTENIDO

CERTIFICADO DEL DIRECTOR.....	i
AUTORÍA DE RESPONSABILIDAD	ii
AUTORIZACIÓN.....	iii
DEDICATORIA	iv
AGRADECIMIENTO	v
ÍNDICE DE CONTENIDO	vi
ÍNDICE DE TABLAS.....	x
ÍNDICE DE FIGURAS.....	xi
RESUMEN.....	xiv
ABSTRACT	xv
 CAPÍTULO I	
INTRODUCCIÓN	
1.1 Antecedentes.....	1
1.2 Planteamiento del problema	2
1.3 Justificación, Importancia y Alcance del Proyecto (Preguntas de Investigación).....	5
1.4 Objetivos.....	6
1.5 Hipótesis	6
 CAPÍTULO II	
FUNDAMENTACIÓN TEÓRICA	

2.1	Marco Teórico Referencial (Legal, Epistemológico y Conceptual).....	8
2.2	Antecedentes del Estado del Arte.....	22

CAPÍTULO III

MEMORIA TÉCNICA METODOLÓGICA

3.1	Metodología de la Investigación.....	29
-----	--------------------------------------	----

CAPÍTULO IV

DESARROLLO DEL MODELO DE GESTIÓN DE INDICADORES DE ACREDITACIÓN

4.1	Justificación.....	31
	Paso 1: Evaluación del Negocio.....	31
4.2	Planeación.....	39
	Paso 2: Evaluación de la Infraestructura.....	39
	Paso 3: Planificación del Proyecto.....	46
4.3	Análisis Del Negocio.....	47
	Paso 4: Definición de Requisitos del Proyecto.....	47
	Paso 5: Análisis de Datos.....	59
	Paso 6: Prototipo de Aplicación.....	60
	Paso 7: Análisis del Repositorio de Metadatos.....	60
4.4	Diseño.....	66
	Paso 8: Diseño del Data Warehouse.....	68

Paso 9: Diseño del ETL.....	70
Paso 10: Diseño del Repositorio de Metadatos.....	71
4.5 Construcción.....	72
Paso 11: Desarrollo del ETL.....	72
Paso 12: Desarrollo Aplicación.....	74
Paso 13: Minería de Datos	74
4.6 Desarrollo	80
Paso 15: Implementación	80
Paso 16: Evaluación de la Solución	81

CAPÍTULO V

RESULTADOS

Facilidad de Aprendizaje del Sistema	83
Eficiencia del Sistema	84
Flexibilidad del Sistema	85
Participación de los Usuarios	87
Tareas.....	88
Usabilidad.....	90
Resumen General de los Resultados Obtenidos	91
Síntesis de las Preguntas de Investigación Planteadas	92

CONCLUSIONES	95
RECOMENDACIONES	96
REFERENCIAS BIBLIOGRÁFICAS	97

ÍNDICE DE TABLAS

Tabla 1 <i>Estudios por grupo de control</i>	24
Tabla 2 <i>Construcción de la cadena de búsqueda</i>	25
Tabla 3 <i>Relación entre los objetivos, las preguntas de investigación y las metodologías</i>	29
Tabla 4 <i>Costos iniciales del proyecto</i>	37
Tabla 5 <i>Costos fijos del proyecto</i>	38
Tabla 6 <i>Proyección de gastos</i>	38
Tabla 7 <i>Análisis de la reducción en horas del proceso</i>	39
Tabla 8 <i>Requisitos de Hardware y Sistema Operativo para el Servidor</i>	44
Tabla 9 <i>Requisitos de Hardware y Sistema Operativo para el Cliente</i>	44
Tabla 10 <i>Análisis del modelo de evaluación de carreras según el tipo de indicador</i>	48
Tabla 11 <i>Análisis de indicadores, fuentes de datos y tablas relacionadas</i>	50
Tabla 12 <i>Análisis del repositorio de metadata</i>	61

ÍNDICE DE FIGURAS

<i>Figura 1.</i> Árbol de Problemas del Proyecto de Investigación	4
<i>Figura 2.</i> Relación entre las variables dependiente e independiente del marco conceptual	8
<i>Figura 3.</i> Resumen de las Técnicas de Análisis de Datos	11
<i>Figura 4.</i> Tipos de Cluster	13
<i>Figura 5.</i> Fórmula para calcular el Índice de silhouette	15
<i>Figura 6.</i> Fases del proceso de autoevaluación de la IES.....	20
<i>Figura 7.</i> Número de artículos encontrados en las bases científicas	26
<i>Figura 8.</i> Etapas de la metodología BIR	31
<i>Figura 9.</i> Fases del proceso de autoevaluación	34
<i>Figura 10.</i> Actividades del proceso de autoevaluación	35
<i>Figura 11.</i> Actividades del proceso de autoevaluación institucional	36
<i>Figura 12.</i> Diagrama de la Infraestructura Tecnológica de la ESPOCH.....	41
<i>Figura 13.</i> Arquitectura del sistema de Inteligencia de Negocios propuesto	41
<i>Figura 14.</i> Comparativa de las plataformas de integración.....	42
<i>Figura 15.</i> Comparativa de las plataformas de visualización de datos.....	43
<i>Figura 16.</i> Mapa de Procesos de la ESPOCH	46
<i>Figura 17.</i> Planificación del proyecto	47
<i>Figura 18.</i> Modelo de Entidad Relación del Sistema Académico.....	59
<i>Figura 19.</i> Arquitectura del sistema de gestión de indicadores.....	60
<i>Figura 20.</i> Modelo de Gestión de Indicadores de Evaluación y de Gestión	66
<i>Figura 21.</i> Esquema de la arquitectura del sistema propuesto	67
<i>Figura 22.</i> Arquitectura del Modelo de Gestión.....	68

Figura 23. Diseño lógico del Data Warehouse.....	69
Figura 24. Esquema para la extracción de datos de la producción científica de tipo regional	70
Figura 25. Job de carga de los datos de docentes.....	71
Figura 26. Repositorio físico de metadatos	71
Figura 27. Base de Datos Stage dsa_evaluacion.....	72
Figura 28. Componente de Pentaho utilizado para la depuración de datos.....	73
Figura 29. Data Warehouse dwh_evaluacion.....	73
Figura 30. ETL para cargar la fact_rendimiento_finanzas.....	74
Figura 31. Configuración del Tableau para acceder al Data Warehouse	74
Figura 32. Número de clusters óptimo.....	75
Figura 33. Clusters generados por el algoritmo k-means.....	76
Figura 34. Cluster generado aplicando el algoritmo PAM.....	76
Figura 35. Cluster Generado por el algoritmo jerárquico Clara.....	77
Figura 36. Análisis del índice de silhouette aplicado al cluster jerárquico	79
Figura 37. Configuración que vincula la tabla de hechos con las dimensiones	80
Figura 38. Reporte de la producción científica por periodo académico.....	80
Figura 39. Producción científica por campo amplio	81
Figura 40. Producción científica por dedicación del docente	81
Figura 41. Evaluación de los resultados mediante sentencias SQL	82
Figura 42. Dimensiones para evaluar la usabilidad del sistema propuesto.....	82
Figura 43. Respuesta que el sistema presenta frente a preguntas del usuario	83
Figura 44. Resultado de la gestión rápida y eficaz del sistema.....	84
Figura 45. Apoya el sistema al desarrollo de las actividades del usuario	84

Figura 46. Contribuye el sistema a generar los indicadores de evaluación.....	85
Figura 47. Se configura el sistema de acuerdo con las necesidades del usuario.....	86
Figura 48. Acceso al sistema en cualquier lugar	86
Figura 49. Tiene el usuario el conocimiento y las habilidades para utilizar el sistema	87
Figura 50. Conoce el usuario todas las funcionalidades del sistema	87
Figura 51. Conoce el usuario los procedimientos para alcanzar los objetivos de evaluación	88
Figura 52. Las tareas que realiza el sistema son habituales para el usuario.....	89
Figura 53. Soluciona el sistema alguna tarea complicada.....	89
Figura 54. Contribuye el sistema a generar indicadores de evaluación	90
Figura 55. Contribuye el sistema al cumplimiento rápido y fácil de tareas y objetivos	91

RESUMEN

Las universidades ecuatorianas están sujetas al cumplimiento de indicadores de acreditación que permitan el aseguramiento de la calidad de la educación superior. Sin embargo, se ha evidenciado que el tiempo empleado para generar los indicadores en general son excesivamente altos y más aún cuando los responsables de obtener esta información no tienen conocimiento de los procesos de acreditación inherentes. Este proyecto implementa un sistema de inteligencia de negocios basado en el modelo de gestión de indicadores emitidos por el Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES) del Ecuador, que permita obtener de manera oportuna los diferentes parámetros necesarios para el proceso de acreditación universitaria. Para el desarrollo del modelo se utilizó la metodología de Business Intelligence Roadmap (BIR), tomando como caso de estudio a la Escuela Politécnica del Chimborazo ESPOCH. La validación del modelo se realizó mediante la aplicación de un test de usabilidad y técnicas de minería de datos. Los resultados indican que existe un elevado porcentaje de aceptación del modelo por parte de los usuarios, lo que también fue confirmado mediante la técnica de clustering ya que se pudo encontrar diferentes patrones de comportamiento en diferentes criterios de evaluación. Se espera que el modelo propuesto contribuya con información de calidad y conocimiento para la toma de decisiones.

PALABRAS CLAVE:

- **INDICADORES DE ACREDITACIÓN**
- **MODELO DE GESTIÓN**
- **BUSINESS INTELLIGENCE ROADMAP**
- **CLUSTERING**

ABSTRACT

The Ecuadorian universities are subjected to certification indicators compliance that allow the quality assurance of higher education. However, it was generally evidenced that the time to generate the indicators is excessively demanding since those responsible for obtaining this information do not have enough knowledge about the accreditation inherent processes. Thus, this project implements a business intelligence system based on the indicator management model issued by Consejo de Aseguramiento de la Calidad de la Educación Superior del Ecuador (CACES), which enables to get different parameters necessary for the university accreditation process in a convenient manner. For the development of the model, the Business Intelligence Roadmap (BIR) methodology was used taking Escuela Superior Politécnica de Chimborazo (ESPOCH) as a case study. The application of a usability test and data mining techniques were developed for the system approval. The results indicated that the system had a high level of acceptance by users and this was also proved through the clustering technique since different behavior patterns were found in various evaluation criteria. Therefore, it is expected that proposed model contributes with quality information and knowledge for decision making.

KEY WORDS:

- **ACCREDITATION INDICATORS**
- **MANAGEMENT MODEL**
- **BUSINESS INTELLIGENCE ROADMAP**
- **CLUSTERING**

CAPÍTULO I

INTRODUCCIÓN

1.1 Antecedentes

Los procesos de aseguramiento de la calidad a nivel mundial se vienen consolidando a partir de los años 80, pese a que estos se han instaurado muchos años atrás como en el caso de Estados Unidos, quien dio inicio al proceso de acreditación educativa en el siglo XIX como una forma de proteger la salud, la seguridad y los servicios de interés público. Cada país posee un sistema de acreditación que asegura la calidad de la educación superior, por ejemplo en España el organismo encargado de la acreditación de la educación superior es la Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA), en Chile se tiene la Comisión Nacional de Acreditación (CNA), en Finlandia se tiene el Consejo de Evaluación para la Educación y Capacitación y en Estados Unidos se tiene el Consejo de Educación Superior (Council for Higher Education Accreditation - CHEA).

En el Ecuador, las IES son reguladas por el Consejo de Aseguramiento de la Calidad de la Educación Superior (CACES), organismo que cuenta con modelos de evaluación institucional y de carreras, que establecen criterios e indicadores que determinan la acreditación y el aseguramiento de la calidad de las IES.

El primer proceso de evaluación institucional se realizó en el año 2009, categorizándose las instituciones de educación superior en cinco grupos: 11 en categoría A, 9 en categoría B, 13 en categoría C, 13 en categoría D y 21 en categoría E. De las últimas, 13 fueron cerradas por no cumplir con los estándares mínimos de acreditación. El segundo proceso de evaluación institucional se realizó en el año 2013, como resultado del total de 54 instituciones de educación superior (IES) evaluadas, 5 se mantuvieron en la categoría A, 23 en la categoría B, 18 en la

categoría C y 8 en la categoría D, observándose de esta manera una mejora en los resultados. Las carreras que al momento han sido evaluadas son: Medicina, Enfermería, Odontología y Derecho.

Las IES acogiendo a las normativas del CACES, y a la búsqueda permanente de la excelencia académica a través de la práctica de la calidad en todas sus actividades han creado departamentos o dependencias encargadas de los procesos de evaluación y aseguramiento de la calidad.

1.2 Planteamiento del problema

Contexto del problema

La Escuela Politécnica del Chimborazo (ESPOCH) cuenta con un plan estratégico institucional 2014 - 2018, en el cual se definieron los subsistemas de Investigación, Vinculación, Academia, Administración y Estrategia. Las dependencias como: Talento Humano, DTIC's, Publicaciones y Evaluación, poseen sistemas de información que almacenan datos del personal docente, de las actividades académicas, de la producción científica generada y del desempeño docente. Estos sistemas de información poseen datos útiles para los procesos de evaluación y autoevaluación institucional y de carreras.

Por otro lado, la dependencia responsable de los procesos de autoevaluación, acreditación y aseguramiento de la calidad de la ESPOCH se denomina Departamento de Evaluación y Aseguramiento de la Calidad (DEAC), la cual se encarga de planificar la autoevaluación institucional y de carreras, para ello aplica los siguientes modelos de evaluación del CACES:

- Modelo de Evaluación Institucional de Universidades y Escuelas Politécnicas 2018, el cual se enfoca a evaluar a la institución y posee 37 indicadores, de los cuales 24 son cualitativos y 13 son cuantitativos
- Modelo Genérico de Evaluación del Entorno de Aprendizaje de Carreras Presenciales y Semipresenciales de las Universidades y Escuelas Politécnicas del Ecuador 2015, el

mismo que posee 37 indicadores, de los cuales son 20 cualitativos y 17 cuantitativos, y se enfoca a evaluar cada carrera.

La autoevaluación institucional se realiza cada año y la autoevaluación de carreras se realiza cada periodo académico y para estos procesos se mantiene un repositorio (nube) en el cual se almacena información de los indicadores de evaluación de la IES del caso práctico. El proceso de evaluación institucional lo realiza la DEAC en conjunto con el personal docente que conforman las comisiones de evaluación y aseguramiento de la calidad, de cada carrera que posee la IES. Las comisiones con el apoyo del personal docente levantan información de los indicadores de evaluación por criterio y generan una ficha por cada indicador, las fichas manejan medidas cualitativas y cuantitativas. Ambas medidas son obtenidas después de verificar manualmente la información.

El problema de investigación

Al momento, los procesos de autoevaluación que se realizan en la IES del caso práctico incluyen actividades como: levantar información de las evidencias, verificar evidencias, generar o calcular indicadores, entre otras. Sin embargo, estas actividades presentan un bajo rendimiento debido a que el tiempo empleado para los procesos demanda de varias semanas y se incrementa considerablemente, cuando los responsables de obtener las evidencias y los indicadores no tienen conocimiento de los procesos. Cabe mencionar que, en algunas actividades se realiza una manipulación manual de los datos y de las evidencias. Tampoco existe una articulación entre los procesos de evaluación y de gestión, lo que afecta a la consecución de los objetivos estratégicos institucionales. El bajo rendimiento de los procesos de autoevaluación afecta directamente a la acreditación institucional o a la acreditación de una determinada carrera, provocando una baja satisfacción en el cliente interno y una disminución

en la demanda de la oferta académica, lo que representa limitaciones en el acceso de los graduados al mundo laboral y debilidades en la imagen institucional.

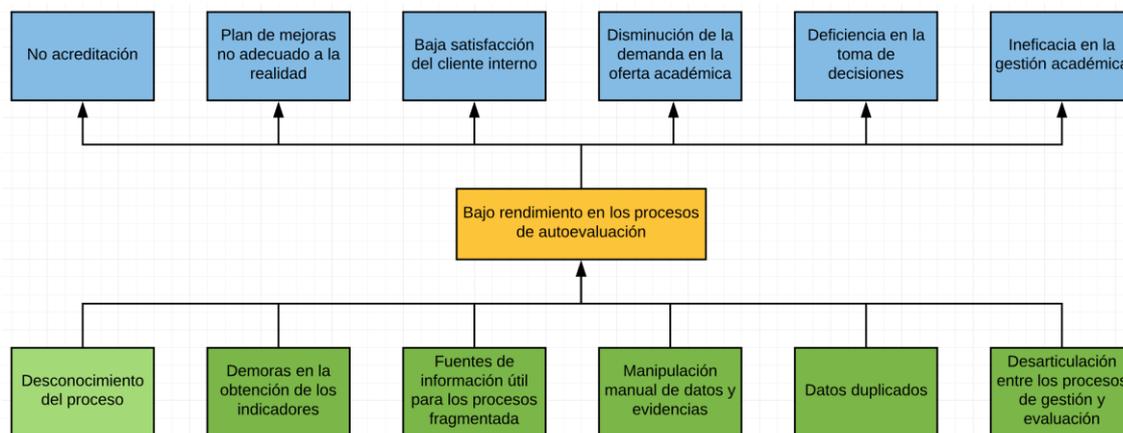


Figura 1. Árbol de Problemas del Proyecto de Investigación

Las fuentes de información útil que poseen las IES para los procesos de evaluación se encuentran en bases de datos independientes, esta fragmentación, amplía las posibilidades de encontrarse con datos e información duplicada, disminuyendo de esta manera, la calidad de la información almacenada. El no disponer de datos en tiempo real y reportes que generen valor en los procesos de autoevaluación, incide en que la IES no disponga de un plan de mejoras adecuado a la realidad, disminuyendo de esta manera la oportuna toma de decisiones y la efectividad de la gestión académica provocando el incumplimiento del mínimo de los indicadores de evaluación y acreditación según los modelos del CACES.

Preguntas de investigación

En el estudio planteado, se han definido las siguientes preguntas de investigación:

OE1. P1: ¿En qué consiste el Data Warehousing y la minería de datos aplicada a la educación?

OE1.P2: ¿Cuáles son las técnicas de minería de datos y las técnicas de validación de datos que contribuyen a mejorar los resultados en la evaluación de la IES?

OE2.P1: ¿El modelo de gestión de indicadores es adecuado para la evaluación institucional y de carreras?

OE2.P2: ¿Qué indicadores de evaluación se puede obtener con el modelo de gestión?

OE3.P1: ¿Cuál es el nivel de aceptación que tiene el modelo de gestión de indicadores?

OE3.P2: ¿Cuál es el resultado óptimo que se debe considerar al implementar el modelo de gestión de indicadores?

1.3 Justificación, Importancia y Alcance del Proyecto (Preguntas de Investigación)

Justificación

El presente trabajo se justifica por la necesidad que tiene el departamento de evaluación de la IES para contar con reportes de los indicadores de evaluación, que son medidas que determinan la calidad de la IES y garantizan su acreditación y funcionamiento. Si no se considera una solución al problema antes planteado, el departamento y las carreras seguirán demandando de más esfuerzos y tiempo empleado en la revisión de la información y la generación de los indicadores, provocando retrasos en la entrega de resultados, generando una baja satisfacción del cliente interno de la IES, debilitando la toma de decisiones oportuna frente a indicadores de altos niveles de importancia.

Importancia

El no disponer de datos en tiempo real y reportes que generen valor, incide en que no se disponga de un plan de mejoras adecuado a la realidad, disminuyendo de esta manera la efectividad de las decisiones tomadas y la calidad de la educación superior.

Alcance

Se realizará un análisis del estado actual de los procesos de la autoevaluación para la selección de los indicadores de evaluación institucional y de carreras, luego se realizará un mapeo sistemático de literatura para determinar las plataformas de la inteligencia de negocios, la

técnica de validación de datos y las técnicas de minería de datos que requiere el caso práctico utilizando las bases científicas. Se continuará con el diseño del modelo de gestión de indicadores de evaluación basado en el análisis descriptivo y predictivo de los datos que permita optimizar la toma de decisiones de la IES. Finaliza con el análisis de la evaluación de la usabilidad del sistema de inteligencia de negocios y la validación del modelo de minería de datos.

El caso práctico se aplica al modelo de evaluación institucional y de carreras, específicamente de la carrera de Finanzas de la ESPOCH.

1.4 Objetivos

General

Diseñar un modelo de gestión de indicadores de evaluación basado en un sistema de Inteligencia de Negocios (BI) para fortalecer la eficiencia de los procesos de autoevaluación de la IES.

Objetivos específicos

OE1: Determinar los fundamentos teóricos de la inteligencia de negocios, la minería de datos aplicada a la educación y el marco teórico referencial mediante una revisión de literatura.

OE2: Diseñar el modelo de gestión de indicadores de evaluación e implementar un sistema de BI basado en Data Warehousing y minería de datos educativos.

OE3: Validar el modelo de gestión de los indicadores de evaluación mediante un test de usabilidad y técnicas de validación implementadas en la minería de datos.

1.5 Hipótesis

El modelo de gestión de indicadores de evaluación permitirá mejorar la eficiencia de los procesos de autoevaluación de la IES.

Identificación de Variables:

Variable dependiente: Procesos de autoevaluación de la IES

Variable independiente: Modelo de Gestión de Indicadores de Evaluación

Para la demostración de la hipótesis se aplicará una prueba de usabilidad a los docentes miembros de la comisión de evaluación y aseguramiento de la calidad de la carrera de Finanzas de la ESPOCH, y además se emplearán las técnicas de validación al modelo de minería de datos generado.

CAPÍTULO II

FUNDAMENTACIÓN TEÓRICA

2.1 Marco Teórico Referencial (Legal, Epistemológico y Conceptual)

La fundamentación teórica busca articular la teoría con la hipótesis, para esto se realiza un análisis de la teoría usando las variables del problema, con la finalidad de investigar jerárquicamente cada categoría hasta llegar a las variables dependientes e independientes del tema de estudio, para esto se propone la siguiente jerarquía:

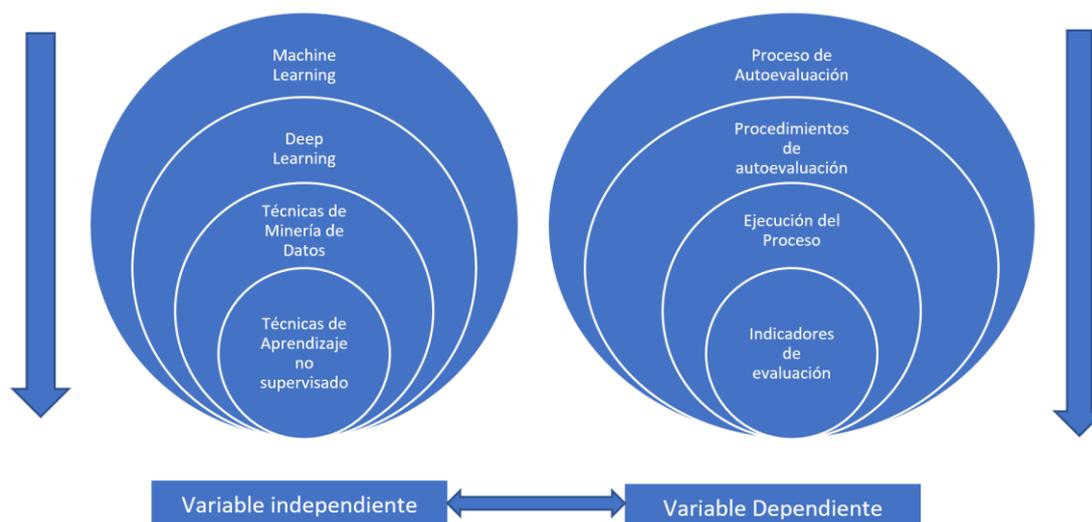


Figura 2. Relación entre las variables dependiente e independiente del marco conceptual

Caracterización de las variables

Variable independiente

Inteligencia de Negocios

Según (Curto, 2016) la Inteligencia de Negocios también conocida como Business Intelligence (BI) es el conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación

y administración de información que permite tomar mejores decisiones a los usuarios de una organización.

Por otro lado, (Gartner, 2017) menciona que la Inteligencia de Negocios (BI) es un término general que incluye las aplicaciones, la infraestructura, las herramientas, y las mejores prácticas que permiten el acceso y el análisis de la información para mejorar y optimizar las decisiones y el rendimiento de las organizaciones.

Data Warehousing

El Data Warehousing hace referencia al proceso de construcción y uso de un almacén de datos, este término fue introducido en 1988 por los investigadores de IBM Barry Devlin y Paul Murphy (Frankenfield, 2019). Un almacén de datos se construye integrando datos de múltiples fuentes heterogéneas que admiten informes analíticos, consultas estructuradas y/o ad-hoc, y el análisis de datos históricos para la toma de decisiones. El Data Warehousing implica la limpieza, integración y consolidación de datos (TutorialsPoint, 2019).

Minería de Datos

Según (Zucco, 2018), la minería de datos es una disciplina que combina la estadística clásica y algoritmos informáticos con la finalidad de obtener conocimiento e información útil para una organización. La minería de datos forma parte del conjunto de tecnologías de la Inteligencia de Negocios, que comprende una serie de técnicas, algoritmos y métodos cuyo fin es la explotación de grandes volúmenes de datos con vistas al descubrimiento de información previamente desconocida y que pueda servir de ayuda en el proceso de toma de decisiones. Para (Beltrán, 2018) la minería de datos combina técnicas semiautomáticas de inteligencia artificial, análisis estadístico, bases de datos y visualización gráfica, para la obtención de información que no esté representada explícitamente en los datos, logrando que la toma de decisiones se realice con un mayor conocimiento, debido a que se descubren relaciones, tendencias, desviaciones, comportamientos atípicos, patrones y trayectorias ocultas en los datos.

Técnicas de Minería de Datos

La minería de datos comprende una serie de técnicas de análisis de datos que permiten extraer conocimiento. Para decidir cuál es la técnica más adecuada en una determinada situación, es necesario distinguir el tipo de información que se desea extraer. Según su nivel de abstracción, el conocimiento contenido en los datos puede clasificarse en distintas categorías y requerirá una técnica más o menos avanzada para su recuperación (Román, García, & Rueda, 2012). Los tipos de información que se pueden obtener de la minería de datos son: asociaciones, secuencias, clasificaciones, agrupamientos y pronósticos (Laudon & Laudon, 2016).

Asociaciones: Son ocurrencias vinculadas a un solo evento. Por ejemplo, un estudio de los patrones de compra en supermercados podría revelar que, cuando se compran frituras de maíz, el 65 por ciento del tiempo se compra un refresco de cola, pero cuando hay una promoción, el 85 por ciento se compra un refresco de cola. Esta información ayuda a los gerentes a tomar mejores decisiones debido a que descubren la rentabilidad de una promoción.

Secuencias: los eventos se vinculan en el transcurso del tiempo. Por ejemplo, podríamos descubrir que, si se compra una casa, el 65 por ciento del tiempo se compra un nuevo refrigerador dentro de las siguientes dos semanas, y el 45 por ciento se compra un horno dentro del mes posterior a la compra de la casa.

Clasificación: Reconoce los patrones que describen el grupo al que pertenece un elemento, para lo cual se examinan los elementos existentes que hayan sido clasificados y se infiere un conjunto de reglas. Por ejemplo, las empresas como las compañías de tarjetas de crédito o las telefónicas se preocupan por la pérdida de clientes estables. La clasificación ayuda a descubrir las características de los clientes con probabilidades de dejar de serlo y puede proveer un modelo para ayudar a los gerentes a predecir quiénes son esos clientes, de modo que puedan idear campañas especiales para retenerlos.

Agrupamiento: Funciona de una manera similar a la clasificación cuando aún no se han definido grupos. Una herramienta de minería de datos puede descubrir distintas agrupaciones dentro de los datos, como el hecho de encontrar grupos de afinidad para tarjetas bancarias o particionar una base de datos en grupos de clientes con base en la demografía y los tipos de inversiones personales.

Pronóstico: Utiliza las predicciones de una manera distinta. Se basa en una serie de valores existentes para pronosticar cuáles serán los otros valores. Por ejemplo, el pronóstico podría encontrar patrones en los datos para ayudar a los gerentes a estimar el futuro valor de variables continuas, como las cifras de ventas.

(Laudon & Laudon, 2016)

Las técnicas de minería de datos crean modelos que son predictivos y descriptivos. Un modelo predictivo responde preguntas sobre datos futuros. Un modelo descriptivo proporciona información sobre las relaciones entre los datos y sus características.

A continuación, se tiene un resumen de las diferentes técnicas de minería de datos:

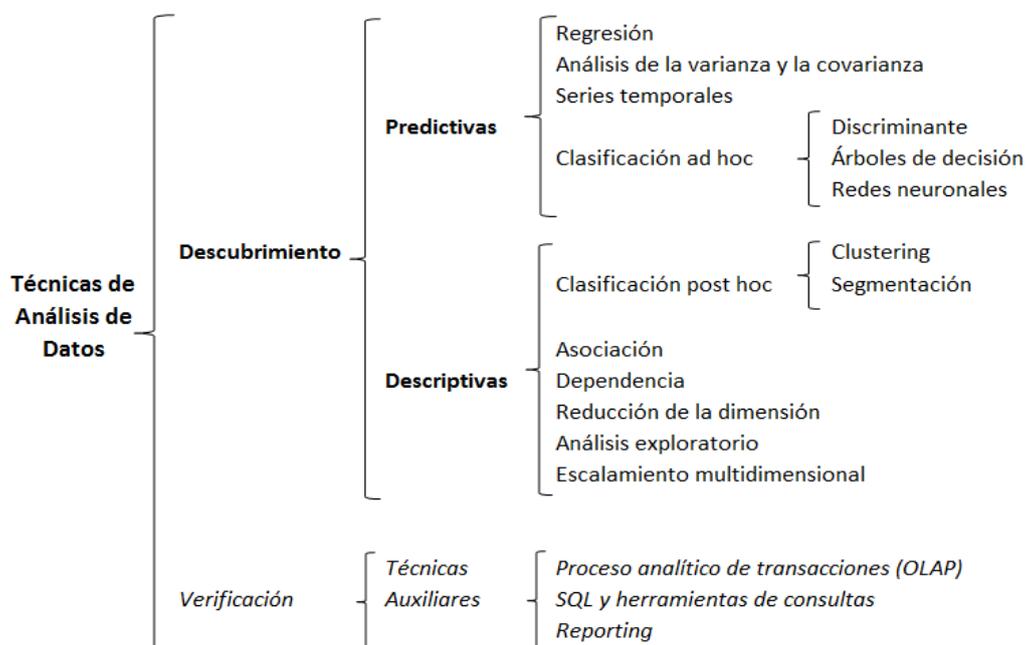


Figura 3. Resumen de las Técnicas de Análisis de Datos

Clustering

Es el proceso de agrupar datos en clases o clusters de tal forma que los objetos de un cluster tengan una similitud alta entre ellos, y baja (sean muy diferentes) con objetos de otros clusters (Morales & Jair, 2018). El objetivo del clustering es que los objetos dentro de un grupo sean similares (o relacionados) entre sí y diferentes de (o no relacionados con) los objetos en otros grupos. A mayor similitud (u homogeneidad) dentro de un grupo y a mayor diferencia entre grupos, mejor o más distinto es el clustering.

El término *clustering* hace referencia a un amplio abanico de técnicas no supervisadas cuya finalidad es encontrar patrones o grupos (*clusters*) dentro de un conjunto de observaciones. Las particiones se establecen de forma que, las observaciones que están dentro de un mismo grupo, son similares entre ellas y distintas a las observaciones de otros grupos. Se trata de un método no supervisado, ya que el proceso ignora la variable respuesta que indica a que grupo pertenece realmente cada observación (si es que existe tal variable). Esta característica diferencia al *clustering* de las técnicas estadísticas conocidas como análisis discriminante, que emplean un set de entrenamiento en el que se conoce la verdadera clasificación.

Dada la utilidad del *clustering* en disciplinas muy distintas (genómica, marketing...), se han desarrollado multitud de variantes y adaptaciones de sus métodos y algoritmos. Pueden diferenciarse tres grupos principales:

- *Partitioning Clustering*: Este tipo de algoritmos requieren que el usuario especifique de antemano el número de *clusters* que se van a crear (*K-means*, *K-medoids*, *CLARA*).
- *Hierarchical Clustering*: Este tipo de algoritmos no requieren que el usuario especifique de antemano el número de *clusters*. (*agglomerative clustering*, *divisive clustering*).

- Métodos que combinan o modifican los anteriores (*hierarchical K-means, fuzzy clustering, model based clustering* y *density based clustering*). (Amat, 2017)

Por otro lado, según (Dropcore, 2019) en el clustering se tiene dos tipos de métodos: los divisivos y los jerárquicos o aglomerativos.

Los métodos divisivos se caracterizan porque inician formando un grupo único con todas las observaciones y se van dividiendo según lo lejanos que estén.

Los métodos aglomerativos se caracterizan porque inician con los objetos o individuos de modo individual; para luego formar tantos clusters iniciales como objetos. Luego se van agrupando de modo que los primeros en hacerlo son los más similares y al final, todos los subgrupos se unen en un único cluster.

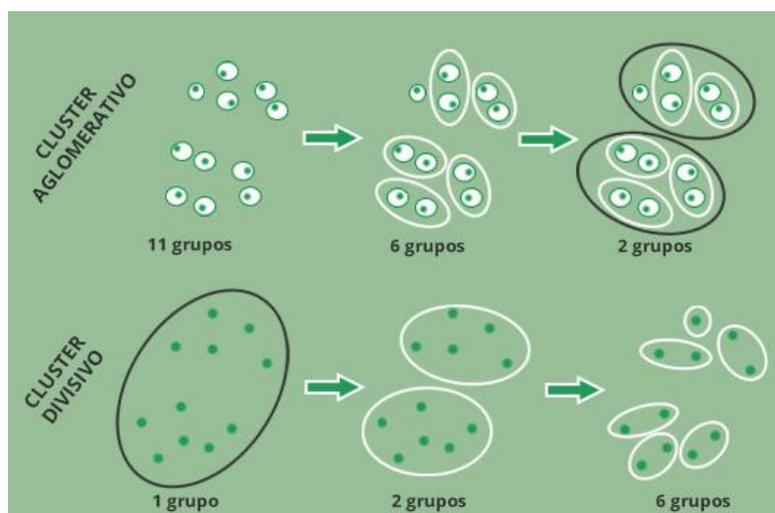


Figura 4. Tipos de Cluster

Fuente: <http://blog.dropcore.com/metodos-de-clustering-mas-utilizados/>

Validación en el clustering

La validación de *clusters* es el proceso por el cual se evalúa la veracidad de los grupos obtenidos. A modo general, este proceso consta de tres partes: estudio de la tendencia

de *clustering*, elección del número óptimo de *clusters* y estudio de la calidad/significancia de los *clusters* generados (Amat, 2017)

Estudio de la tendencia de clustering

Antes de aplicar un método de *clustering* a los datos es conveniente evaluar si hay indicios de que realmente existe algún tipo de agrupación en ellos. A este proceso se le conoce como *assessing cluster tendency* y puede llevarse a cabo mediante test estadísticos (*Hopkins statistic*) o de forma visual (*Visual Assessment of cluster Tendency*).

Elección del número óptimo de clusters

Determinar el número óptimo de *clusters* es uno de los pasos más complicados a la hora de aplicar métodos de *clustering*, sobre todo cuando se trata de *partitioning clustering*, donde el número se tiene que especificar antes de poder ver los resultados. No existe una forma única de averiguar el número adecuado de *clusters*. Es un proceso bastante subjetivo que depende en gran medida del tipo de *clustering* empleado y de si se dispone de información previa sobre los datos con los que se está trabajando, por ejemplo, estudios anteriores pueden sugerir o acotar las posibilidades. A pesar de ello, se han desarrollado varias estrategias que ayudan en el proceso

El método *Elbow* sigue una estrategia comúnmente empleada para encontrar el valor óptimo de un hiperparámetro. La idea general es probar un rango de valores del hiperparámetro en cuestión, representar gráficamente los resultados obtenidos con cada uno e identificar aquel punto de la curva a partir del cual la mejora deja de ser sustancial (principio de verosimilitud). En los casos de *partitioning clustering*, como por ejemplo *K-means*, las observaciones se agrupan de una forma tal que se minimiza la varianza total *intra-cluster*.

Estudio de la calidad/significancia de los clusters generados

Una vez seleccionado el número adecuado de *clusters* y aplicado el algoritmo de *clustering* pertinente se tiene que evaluar la calidad de los de los mismos, de lo contrario, podrían derivarse conclusiones de agrupación que no se corresponden con la realidad. Pueden diferenciarse tres tipos de estadísticos empleados con este fin:

- Validación interna de los *clusters*: Emplean únicamente información interna del proceso de *clustering* para evaluar la bondad de las agrupaciones generadas. Se trata de un proceso totalmente *unsupervised* ya que no se incluye ningún tipo de información que no estuviese ya incluida en el *clustering*.
- Validación externa de los *clusters* (*ground truth*): Combinan los resultados del *clustering*(*unsupervised*) con información externa (*supervised*), como puede ser un set de validación en el que se conoce el verdadero grupo al que pertenece cada observación. Permiten evaluar hasta qué punto el *clustering* es capaz de agrupar correctamente las observaciones. Se emplea principalmente para seleccionar el algoritmo de *clustering* más adecuado, aunque su uso está limitado a escenarios en los que se dispone de un set de datos de validación.
- Significancia de los *clusters*: Calculan la probabilidad (*p-value*) de que los *clusters* generados se deban únicamente al azar.

Según (Hasperué, 2017), la validación del clustering se la realiza con el Índice Silhouette, el cual se calcula para cada cluster

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Figura 5. Fórmula para calcular el Índice de silhouette

Fuente: (Hasperué, 2017)

Donde

- $s(i)$ es la medida de validación para el cluster i
- $a(i)$ es la distancia promedio entre todos los vectores del cluster i
- $b(i)$ es la distancia promedio entre todos los vectores del cluster i y todos los vectores del cluster más cercano
- $-1 \leq s(i) \leq 1$, donde 1 indica una muy buena separación de los clusters

(Hasperué, 2017)

Variable dependiente

IES

Institución de Educación Superior (IES), puede ser una universidad o escuela politécnica que emite títulos de tercer y cuarto nivel.

Modelo de evaluación

Los modelos de gestión para la mejora continua o modelos de evaluación son el punto de referencia que justifica cualquier proceso de cambio o plan de mejora en la educación (Universidad de Salamanca, 2018). La calidad educativa se mide en función de los resultados obtenidos al aplicar modelos de evaluación o de gestión para la mejora continua.

Los modelos en educación proporcionan una guía o heurística para pensar acerca de cómo puede desarrollarse una evaluación (Hoz, 1985)

Evaluación

La evaluación es un proceso o conjunto de procesos que sirven para cerciorarse del valor que posee la realidad educativa de una institución, según su grado de adecuación (o inadecuación) a una instancia de referencia o criterio. La evaluación es un componente intrínseco del proceso

educativo en su conjunto, en la medida en que la educación sea una actividad sistemática intencional (Hoz, 1985).

Acreditación

La acreditación es un medio para reconocer si la educación superior está respondiendo a las exigencias planteadas en los planes de desarrollo propuestos por el Estado y a las necesidades sociales. Así, ésta se define como un proceso de control y garantía de la calidad en la educación superior, resultado de una autoevaluación y de una evaluación externa, con la finalidad de reconocer que una institución, carrera o programa satisfacen los estándares mínimos de calidad y excelencia educativa (Dias, 2008).

Por un lado, la acreditación es el camino para que el Estado reconozca la calidad de las universidades y de sus carreras; una instancia para comparar la formación que se imparte y que la comunidad académica reconoce como válida y deseable (cna, 2006b, 2013). Por otro lado, es un instrumento para promover y reconocer la cultura de la calidad en las universidades, estableciendo sistemas de evaluación permanentes y de mejoramiento continuo (Acosta & Acosta, 2016).

Proceso de autoevaluación

Según el artículo 99 de la LOES, la autoevaluación es el riguroso proceso de análisis que una institución realiza sobre la totalidad de sus actividades institucionales o de una carrera, programa o posgrado específico, con amplia participación de sus integrantes, a través de un análisis crítico y un diálogo reflexivo, a fin de superar los obstáculos existentes y considerar los logros alcanzados para mejorar la eficiencia institucional y mejorar la calidad académica (Ecuador, 2010)

Fases de la autoevaluación

Según la (Puce, 2005) el proceso de autoevaluación tiene 3 fases:

- Planificación de la autoevaluación
- Desarrollo de la autoevaluación
- Plan de mejoramiento

A continuación, se detalla cada una de las fases:

Planificación de la autoevaluación

Esta fase inicia con la capacitación para el diseño, la cual permite difundir, sensibilizar, socializar y legitimar entre los participantes del proceso, la orientación metodológica y estructura de la guía de autoevaluación. La organización para la autoevaluación parte de la conformación de la comisión de Evaluación Interna y la integración de grupos de trabajo a nivel de unidades académicas y carreras. El Diseño, consiste en formular el proyecto de autoevaluación, en donde se define el marco conceptual, se realiza el análisis de la situación institucional inicial, se concretan los objetivos del proyecto, se ajusta el modelo de autoevaluación, se fija el objeto a evaluar, la metodología, la organización y el producto e impacto esperados. Finalmente debe socializarse y aprobarse el diseño del proyecto de autoevaluación, ante las instancias que la Comisión de Evaluación Interna y la normativa lo estipulen.

Desarrollo de la autoevaluación

Es la fase de ejecución del proceso que se inicia con la capacitación para el desarrollo de los procesos de Autoevaluación que consolida la fundamentación teórico-metodológica y afina la utilización de técnicas, instrumentos para la recolección de información y su correspondiente análisis valorativo que conduce a la elaboración de informes. Luego tenemos la aplicación de instrumentos en base a procedimientos preestablecidos que permitan recoger datos, procesarlos, analizar la información recogida, valorar y ponderar la información. La preparación de Informes comprende la presentación de informes parciales, borrador de informe final e informe final. En

el caso de los informes parciales y borrador de informe final éstos deben ser sometidos a análisis y socialización de toda la comunidad universitaria.

Plan de mejoramiento

La implementación de los planes de mejora se deriva de las fortalezas y debilidades encontradas en la autoevaluación y que están reflejadas en el informe final. Estas acciones se articularán al plan estratégico institucional, a través de un proceso de autorregulación.

Actividades de autoevaluación

Según la guía de la (Puce, 2005) y de la (Dirección de Evaluación y Aseguramiento de la Calidad, 2017) las fases de la autoevaluación poseen las siguientes actividades:

Planificación para la autoevaluación

- Capacitación para el diseño
- Organización
- Diseño del proyecto
- Socialización del diseño

Desarrollo de la autoevaluación

- Capacitación para el desarrollo
- Ejecución
- Análisis y valoración de la información
- Borrador de Informe
- Socialización del borrador de informe
- Informe final

Planes de mejoramiento

- Articulación con la planificación institucional

- Acciones

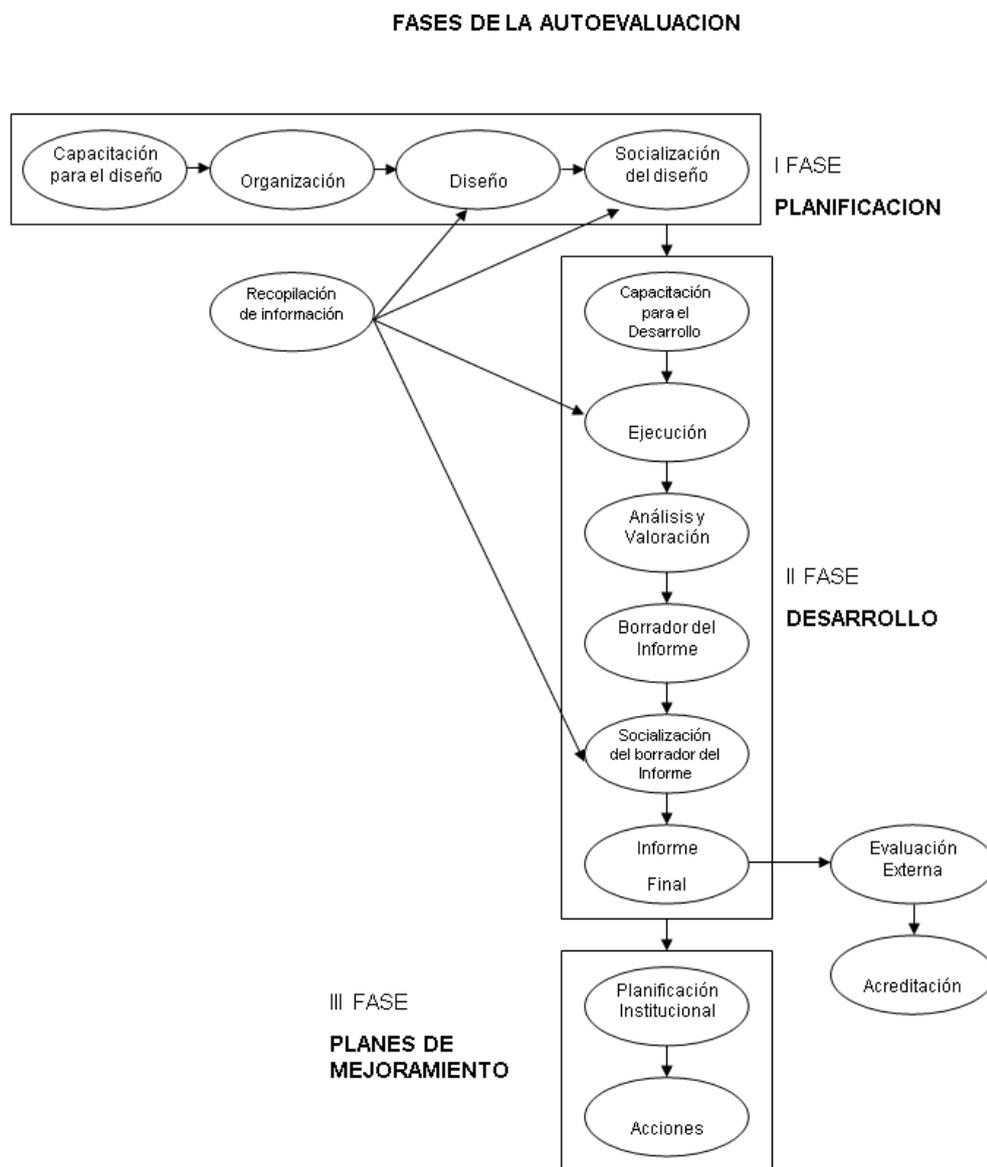


Figura 6. Fases del proceso de autoevaluación de la IES

Criterio de evaluación

Los criterios de evaluación son el principal referente para evaluar los aprendizajes del alumnado. Describen aquello que se quiere valorar (Departamento de Educación del Gobierno de Aragón, 2013).

En evaluación, el criterio es un parámetro en función del que se juzgará el o los atributos de un objeto de estudio que, en este caso, es el aprendizaje de los estudiantes (Ministerio de Educación Chile, 2016)

Los criterios de evaluación son los principios, normas o ideas de valoración en relación a los cuales se emite un juicio valorativo sobre el objeto evaluado. Deben permitir entender qué conoce, comprende y sabe hacer el alumno, lo que exige una evaluación de sus conocimientos teóricos, su capacidad de resolución de problemas, sus habilidades orales y sociales, entre otros aspectos.

Los criterios de evaluación deben concretarse en distintas dimensiones, sub dimensiones y atributos que permitan medir de manera más precisa la evolución en el aprendizaje del alumno, su nivel y calidad.

Las dimensiones en la evaluación hacen referencia a los diferentes aspectos que componen el objeto de la evaluación. Las subdimensiones, son diferentes facetas de una dimensión. Los atributos son cada uno de los elementos o ítems que globalmente constituye una dimensión (Hoz, 1985).

Indicador de evaluación

Los indicadores de evaluación detallan un desempeño observable (y por lo tanto evaluable) del estudiante en relación al objetivo de aprendizaje al cual está asociado, y que permite al docente evaluar el logro del objetivo. Son de carácter sugerido por los programas de estudio, por lo que el docente puede complementarlos. Cada objetivo de aprendizaje cuenta con varios indicadores, y la totalidad de los indicadores dan cuenta del aprendizaje (Ministerio de Educación Chile, 2016)

Indicadores de evaluación institucional

- Doctores
- Doctores tiempo completo
- Doctores en formación

- Estudiantes por profesora
- Titularidad
- Titularidad Tiempo Completo
- Profesoras mujeres
- Dirección Mujeres
- Producción científica Impacto Mundial
- Producción científica Impacto Regional
- Tasa de Retención
- Tasa de Graduación Grado
- Tasa de Graduación Postgrado

Indicadores de evaluación de carrera

- Afinidad Formación Postgrado
- Actualización Científica y/o pedagógica
- Titularidad
- Profesor TC, MT,TP
- Estudiante por profesor
- Producción Regional
- Libros y capítulos de libros

2.2 Antecedentes del Estado del Arte

El estado del arte tiene como objetivo responder a las preguntas de investigación planteadas en el capítulo 1, para su desarrollo se empleó un Mapeo Sistemático de Literatura (SMS), en el cual se tienen criterios de inclusión y exclusión, además de una estrategia de búsqueda. Las

bases científicas consultadas corresponden a: ScienceDirect, IEEEExplore, ACM Digital Library y Springer

Las características idóneas de los artículos a ser tomados en cuenta responden a los siguientes criterios de inclusión y exclusión:

Criterios de Inclusión

- Con el fin de analizar técnicas y metodologías utilizadas en la actualidad, se tomaron artículos a partir de los últimos 5 años, es decir desde el año 2014.
- Se tomaron en cuenta solamente artículos científicos, documentos de conferencias publicados en el idioma inglés.
- Se consideraron los artículos que han utilizado la inteligencia de negocios o la minería de datos en el contexto educativo.

Criterios de Exclusión

- Artículos que tengan temas de inteligencia de negocios no relacionados con el contexto educativo.
- Artículos que no estén en el idioma inglés

Definición de la estrategia de búsqueda

Revisión Inicial: Se realiza una búsqueda inicial en los distintos repositorios académicos para buscar estudios relacionados con las preguntas de investigación.

Validación cruzada de estudios: En esta fase se procede a verificar que los estudios cumplan con los criterios de inclusión y exclusión, con lo cual finalmente se obtienen el listado inicial de documentos académicos con los cuales se va a trabajar en las siguientes fases del estudio.

Integración del Grupo de Control: El grupo de control está conformado por los estudios que cumplen con los criterios de inclusión y exclusión, para lo cual se realizó un análisis del título de los estudios, introducción, conclusiones y palabras claves. Los estudios seleccionados para el grupo de control son los siguientes:

Tabla 1

Estudios por grupo de control

Grupo de control	Título	Palabras Clave
EC1	Clustering for improving Educational Process Mining	Clustering, process mining, educational data mining, learning analytics
EC2	Educational data mining applications and tasks: A survey of the last 10 years	Educational data mining · Surveys · Taxonomy of applications
EC3	Predictive Analytics Using Data Mining Technique	Weka, EDM, Data Mining, Prediction, Classification, Decision Trees
EC4	A New Data Mining Model Adopted for Higher Institutions	Data Models, Data Preprocessing, Higher Education Institution (HEI), Data Mining, Admission System, Educational Data Mining
EC5	Using data mining to predict instructor performance	Data mining; decision tree; multilayer perception; naïve bayes; sequential minimal optimization
EC6	Mining Educational Data for Academic Accreditation: Aligning Assessment with Outcomes	Big Data Higher education Learning analytics Educational data mining
EC7	An Approachable Analytical Study on Big Educational Data Mining	Big Data, Educational Data, Educational Data Mining, Data Mining, Analytical Study
EC8	Fuzzy K- Means Cluster Validation for Institutional Quality Assessment	Data Mining, Educational institution, Cluster, decision tree

Construcción de la cadena de búsqueda

Para la construcción de la cadena de búsqueda se utilizaron las palabras que más se repiten en cada contexto definido en los estudios del grupo de control: algoritmos de minería de datos, descripción del proceso, objetivo del análisis.

Tabla 2
Construcción de la cadena de búsqueda

Contexto	Palabra clave	EC1	EC2	EC3	EC4	EC5	EC6	EC7	EC8	Número de Repeticiones
Algoritmos de minería de datos	Clustering	X					X		X	3
	data mining techniques		X		X					2
	classification algorithms			X		X				2
	Association						X			1
	Prediction						X			1
Descripción del proceso	Educational process mining	X								1
	Educational data mining		X		X		x			3
	process of data mining			X				X		2
	<i>Data mining technology</i>								X	1
	<i>Data mining in education</i>					X				1
Objetivo de la minería	Performance learning quality	X								1
	The quality of higher education institutions		X						X	1
	instructor performance					X				1
	education system quality				X					1
	process of accreditation			X			X			2
	present the right offers							X		1

La cadena de búsqueda está formada por la unión de las palabras claves que más se repiten en cada contexto, los conectores usados son OR para las palabras que están dentro del mismo contexto y el conector AND para las palabras que están en contextos distintos, de esta manera se establece la siguiente cadena de búsqueda:

“Clustering and Educational data mining and process of accreditation”

Al aplicar la cadena de búsqueda en las bases científicas seleccionadas para el estudio se obtienen los siguientes resultados:

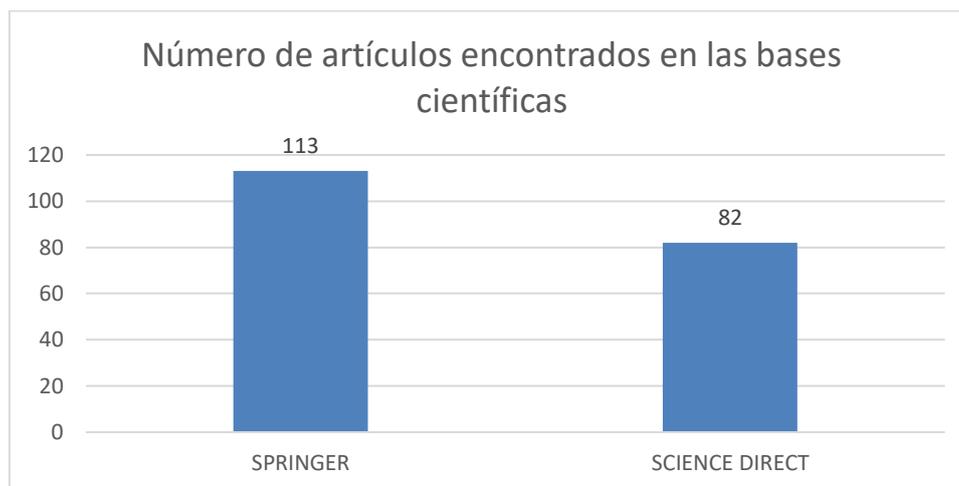


Figura 7. Número de artículos encontrados en las bases científicas

Para seleccionar los estudios se aplicaron los criterios de inclusión y exclusión, es decir se seleccionaron solo artículos científicos o documentos de conferencias publicados en el idioma inglés, cuya fecha de publicación fue mayor al 2014.

El proceso de búsqueda de literatura relacionada con el tema de investigación arrojó resultados que se observan en la Figura 7, la cual indica que se tiene 113 artículos en la base científica de Springer y la 82 en la base de Science Direct. Cabe mencionar que la búsqueda también se realizó en las bases científicas de la IEEE y ACM, las cuales arrojaron como resultado 1 artículo, sin embargo, considerando los criterios de inclusión y exclusión, estos se descartaron.

Una vez obtenidos los resultados, se seleccionaron aquellos estudios relacionados con el tema de investigación, los cuales se listan a continuación:

Predictive Analytics Using Data Mining Technique (Gulati, 2015)

En este artículo se expone sobre el uso de las técnicas de minería de datos para identificar los factores que provocan el incremento de la tasa de deserción escolar con el fin de mejorar la calidad del sistema educativo superior.

Using data mining to predict instructor performance (Ahmed, Rizaner, & Ulusoy, 2016)

En este artículo los autores exponen sobre la aplicación de la minería de datos en la educación, con el objetivo de mejorar la calidad del sistema educativo, a través de la predicción del rendimiento del instructor e investiga los factores que afectan los logros de los estudiantes. Aplican un sistema de clasificación de datos a través de los algoritmos de árbol de decisiones J48, Percepción multicapa, Naïve Bayes y Optimización mínima secuencial.

Mining Educational Data for Academic Accreditation: Aligning Assessment with Outcomes (M. Hussain, Al-Mourad, Mathew, & Hussein, 2017)

En este artículo se expone sobre la aplicación de la minería de datos educacional para evaluar a una institución de educación superior en comparación con el estándar establecido por el organismo de acreditación. El artículo propone la alineación de la evaluación con los resultados de aprendizaje de los estudiantes.

Fuzzy K- Means Cluster Validation for Institutional Quality Assessment (Kumar & Ramaswami, 2010)

En este artículo los autores muestran el resultado de la aplicación de la minería de datos para evaluar la calidad de los servicios brindados por una institución de educación superior. El modelo predictivo que se utilizó es el árbol de decisión y el algoritmo K-means. La gestión de las decisiones utilizando la minería se logra a través del análisis de datos de los estudiantes, docentes y administrativos.

Classification, Clustering and Association Rule Mining in Educational Datasets Using Data Mining Tools: A Case Study (S. Hussain, Atallah, & Kamsin, 2019)

En este artículo, los autores exponen los resultados de aplicar la minería de datos educacional para descubrir la relación estudiante, aprendizaje y comportamiento, con el fin de

mejorar la calidad de la educación y disminuir la tasa de deserción. Para el análisis los autores aplicaron algoritmos de agrupación y clasificación.

Academic Advising System Using Data Mining Method for Decision Making Support

(Elcullada-Encarnacion, 2018)

En esta investigación se ha considerado las altas tasas de abandono de estudiantes universitarios, emprendiendo un análisis continuo de la población estudiantil y evaluar la efectividad de los servicios de asesoramiento académico. El sistema facilita el asesoramiento de actividades académicas y realiza un seguimiento del rendimiento académico del estudiante, siendo capaz de generar el conjunto de datos requerido que se puede usar para, análisis y modelado de datos obteniendo conocimiento útil con respecto al perfil y registro de los estudiantes. Los resultados de la minería de datos en gráficos y tablas son muy esenciales para tomar decisiones oportunas e implementar acciones apropiadas.

Conclusión

La minería de datos aplicada a la educación se denomina minería de datos educacional (EDM), la cual es un campo de tecnología emergente. La mayoría de las aplicaciones de la EDM se enfocan en el análisis de datos de los estudiantes con el fin de mejorar la calidad de la educación, el rendimiento académico y el desempeño de los docentes.

De todas las técnicas de minería de datos analizadas en los artículos científicos, en el presente proyecto se empleará el clustering debido a los resultados obtenidos en el mapeo de literatura y la necesidad de mejorar el rendimiento académico de los estudiantes.

CAPÍTULO III

MEMORIA TÉCNICA METODOLÓGICA

3.1 Metodología de la Investigación

Para el desarrollo del presente trabajo de investigación se realizó una combinación de metodologías, cada metodología está orientada a cumplir con los objetivos de la investigación, es así como para la consecución del primer objetivo se realizó una revisión de literatura para definir los fundamentos teóricos de la inteligencia de negocios, para determinar la técnica de minería de datos que se aplicará en el presente caso de estudio y el marco teórico referencial.

Para la consecución del segundo objetivo se empleó la metodología de Business Intelligence Roadmap (BIR), la cual es una guía para desarrollar proyectos y aplicaciones de inteligencia de negocios. Para el desarrollo del Data warehouse se empleó el paradigma de Kimball.

Para la consecución del tercer objetivo se aplicó un test de usabilidad basado en el modelo de (Poropat, 2014) , el cual aborda criterios de evaluación como: facilidad de aprendizaje del sistema, la eficiencia del sistema, la flexibilidad del sistema, la participación de los usuarios del sistema y la usabilidad del sistema.

A la vez, se emplearon técnicas de validación del clustering para evaluar el modelo de data mining aplicado al rendimiento académico de los estudiantes.

Tabla 3

Relación entre los objetivos, las preguntas de investigación y las metodologías

Objetivos	Preguntas de Investigación	Metodologías Investigación
OE1: Determinar los fundamentos teóricos de la inteligencia de negocios, la minería de datos aplicada a la educación y el marco teórico referencial mediante una revisión de literatura	OE1.P1: ¿En qué consiste el Data warehousing y la minería de datos aplicada a la educación? OE1.P2: ¿Cuáles son las técnicas de minería de datos y las técnicas de validación de datos que contribuyen a mejorar los resultados en la evaluación de la IES?	Revisión de Literatura
OE2: Diseñar el modelo de gestión de indicadores de evaluación e implementar un sistema de BI basado en data	OE2.P1: ¿El modelo de gestión es adecuado para la evaluación institucional y de carreras? OE2.P2: ¿Cuál es la periodicidad con que se debe aplicar el modelo de gestión a los indicadores de evaluación?	Business Intelligence Roadmap (BIR) Ralph Kimball

CONTINÚA



warehousing y minería de datos educativos.

OE3. Validar el modelo de gestión de los indicadores de evaluación a través de un test de usabilidad y técnicas de validación implementadas en la minería de datos

OE3.P1: ¿Cuál es el nivel de aceptación que tiene el modelo de gestión de indicadores?

OE3.P2: ¿Cuál es el resultado óptimo que se debe considerar al implementar el modelo de gestión de indicadores?

Test de Usabilidad (Método Descriptivo de la Investigación)
Técnicas de Validación del Data Mining

CAPÍTULO IV

DESARROLLO DEL MODELO DE GESTIÓN DE INDICADORES DE ACREDITACIÓN

El modelo de gestión de indicadores de acreditación se desarrollará utilizando la metodología de BUSINESS INTELLIGENCE ROADMAP(BIR), la cual es una guía del ciclo de vida de un proyecto para desarrollar aplicaciones de soporte a la toma de decisiones utilizando datos estructurados (Moss & Atre, 2003)

La metodología BIR posee 16 etapas, las cuales se encuentran organizadas en 6 secciones como son: Justificación, Planeación, Análisis del Negocio, Diseño, Construcción y Desarrollo.



Figura 8. Etapas de la metodología BIR

4.1 Justificación

Paso 1: Evaluación del Negocio

En esta fase se realizaron las siguientes actividades:

1. Determinar las necesidades de negocio.
2. Evaluar las soluciones de sistemas de apoyo a la decisión actuales.
3. Valorar la fuente operacional y los procedimientos.

4. Evaluar a los competidores de iniciativas BI.
5. Determinar los objetivos de la aplicación BI.
6. Proponer una solución BI.
7. Realizar un análisis de costo-beneficios.
8. Hacer una evaluación de riesgos.
9. Escribir el resultado de la evaluación

1. DETERMINAR LAS NECESIDADES DEL NEGOCIO

Como parte de las actividades para determinar las necesidades del negocio, se aplicó la observación in situ, la revisión de informes de autoevaluación de periodos académicos anteriores al periodo académico: Octubre 2017 – Marzo 2018, la revisión de los modelos de Evaluación Institucional de Universidades y Escuelas Politécnicas 2018 y el Modelo Genérico de Evaluación del Entorno de Aprendizaje de Carreras Presenciales y Semipresenciales de las Universidades y Escuelas Politécnicas del Ecuador 2015 del Consejo de Aseguramiento de la Calidad de la Educación Superior CACES, anteriormente denominado CEAACES¹, así como la inmersión y aprendizaje como responsable del proceso de autoevaluación de la carrera de Finanzas de la ESPOCH.

A partir de ello, como ya se indicó en el capítulo 1, el principal problema que se busca solucionar es el bajo rendimiento en los procesos de autoevaluación, provocado por el desconocimiento de los procesos, las demoras en la obtención de indicadores, la existencia de información de los procesos fragmentada, la manipulación manual de datos y evidencias, la duplicidad en los datos y la desarticulación entre los procesos de gestión y evaluación. Esto

¹ La Ley Orgánica Reformatoria a la Ley Orgánica de Educación Superior -LOES- aprobada por el Pleno de la Asamblea Nacional y que entró en vigencia el jueves 2 de agosto de 2018 al publicarse su texto en el Registro Oficial, establece varios cambios en el CEAACES, comenzando por su nombre o denominación, ahora se llama CACES que significa Consejo de Aseguramiento de la Calidad de la Educación Superior . Más información: <http://ecuadoruniversitario.com/de-instituciones-del-estado/caces/por-mandato-de-la-ley-el-ceaaces-cambia-su-denominacion-por-caces/>

genera la no acreditación de carreras, un plan de mejoras no adecuado a la realidad, la baja satisfacción del cliente interno, la disminución en la demanda de la oferta académica, la deficiencia en la toma de decisiones y la ineficacia en la gestión académica.

El desarrollo del modelo de gestión de indicadores de evaluación y por ende el desarrollo del sistema de gestión para la toma de decisiones, se justifica en razón de los problemas detectados.

2. EVALUAR LAS SOLUCIONES DE SISTEMAS DE APOYO A LA DECISIÓN ACTUALES

De las soluciones al sistema de apoyo a la decisión como es el caso que se desarrollará en el presente proyecto, no existen plataformas comerciales, ya que el desarrollo es a medida de la información que posee la entidad que es parte del caso de estudio.

3. VALORAR LA FUENTE OPERACIONAL Y LOS PROCEDIMIENTOS

Para el desarrollo de los procesos de autoevaluación institucional y de carrera, se han generado una serie de actividades, planificadas por la Dirección de Evaluación y Aseguramiento de la Calidad. Es así como los procesos de autoevaluación de las carreras se realizan cada semestre, mientras que los procesos de autoevaluación institucional se realizan cada dos años.

Según el Reglamento para los Procesos de Autoevaluación para la Escuela Superior Politécnica de Chimborazo, sus carreras y programas, las fases del proceso de autoevaluación son:

1. Planificación del proceso de autoevaluación
2. Ejecución
3. Generación del informe y entrega de los resultados



Figura 9. Fases del proceso de autoevaluación

Fuente: Reglamento para los Procesos de Autoevaluación para la Escuela Superior Politécnica de Chimborazo, Carreras y Programas

A continuación, se presenta la planificación de actividades para el proceso de autoevaluación de carreras:

1	Diseño de instrumentos y guía de presentación de informes de autoevaluación	DEAC
2	Diseño del plan de capacitación	DEAC
3	Notificación a la DEAC de designación de CEACC	AUTORIDADES FACULTADES Y EXTENSIONES
4	Implementación del repositorio para carga de evidencias	DEAC
5	Solicitud y verificación de carga de evidencias sobre indicadores institucionales	DEAC
6	Carga de evidencias sobre indicadores institucionales	DEPENDENCIAS ESPOCH
7	Ejecución plan de capacitación	DEAC
8	Carga de evidencias sobre indicadores de carrera	CEACC
9	Entrevistas actores (autoridades, profesores, estudiantes, servidores) y visitas in situ a dependencias de las carreras	CEACC
10	Realización de pruebas de resultado de aprendizaje	FACULTAD
11	Autoevaluación y valoración de indicadores	CEACC
12	Seguimiento al proceso de autoevaluación de carreras	CEACC/ DEAC
13	Elaboración de informe preliminar de autoevaluación	CEACC
14	Socialización de resultados y del informe preliminar de autoevaluación	CEACC / AUTORIDADES DE FACULTADES Y EXTENSIONES / INTEGRANTES CARRERA (DOCENTES, ESTUDIANTES Y SERVIDORES)
15	Diseño del plan de mejoras de carrera	CEACC / AUTORIDADES DE FACULTADES Y EXTENSIONES / INTEGRANTES CARRERA (DOCENTES, ESTUDIANTES Y SERVIDORES)
16	Elaboración del Informe Final de Autoevaluación y entrega a la DEAC	CEACC
17	Análisis y validación de informes de autoevaluación	DEAC
18	Alcance a informes de autoevaluación de carreras período abril 2018 - agosto 2018	CEACC
19	Elaboración y entrega del Informe Final del Proceso de Autoevaluación de carreras del periodo abril 2018 - agosto 2018	DEAC
Realizado por: DEAC		
<input type="text"/>		
<input type="text"/>		
Nota: La DEAC realizará un monitoreo constante en todo el proceso de autoevaluación.		
DEAC: Dirección de Evaluación y Aseguramiento de la Calidad		
CEACC: Comisión de Evaluación y Aseguramiento de la Calidad de Carrera		

Figura 10. Actividades del proceso de autoevaluación

Fuente: Dirección de Evaluación y Aseguramiento de la Calidad

La planificación del proceso de autoevaluación institucional incluye las siguientes actividades:

Componente - Tarea
Elaboración y aprobación del plan de autoevaluación institucional
Aprobación de Plan de autoevaluación institucional
Capacitación responsables de Recolección
Recolección de evidencias
Validación de Evidencias
Cálculo de indicadores
Análisis de resultados
Redacción del informe preliminar de autoevaluación institucional
Socialización y recolección de aportes del informe preliminar de autoevaluación
Redacción del informe final de autoevaluación
<i>Inclusión de aportes al borrador de informe de autoevaluación</i>
<i>Revisión del informe por la CGEI</i>
<i>Recomendación de aprobación de informe de autoevaluación</i>
Aprobación del informe de autoevaluación institucional
Difusión de resultados

Figura 11. Actividades del proceso de autoevaluación institucional

Fuente: Dirección de Evaluación y Aseguramiento de la Calidad

Los actores involucrados en los procesos de autoevaluación institucional son:

- Dirección de Evaluación y Aseguramiento de la Calidad (DEAC)
- Consejo Politécnico
- Unidades administrativas y académicas
- Comisión General de Evaluación Interna

4. EVALUAR A LOS COMPETIDORES DE INICIATIVAS BI.

No se tienen definidos competidores

5. DETERMINAR LOS OBJETIVOS DE LA APLICACIÓN BI

- Incrementar el rendimiento en los procesos de autoevaluación de la ESPOCH

- Obtener un plan de mejoras adecuado a la realidad de la institución y de cada una de las carreras de la ESPOCH
- Incrementar la satisfacción del cliente interno,
- Incrementar la demanda de la oferta académica,
- Incrementar la eficiencia en la toma de decisiones
- Incrementar la eficacia en la gestión académica

6. PROPONER UNA SOLUCIÓN BI

Frente a los objetivos planteados se propone generar una plataforma de soporte a la toma de decisiones que permita:

- Generar indicadores cuantitativos de evaluación institucional y de carreras
- Obtener medidas complementarias

7. REALIZAR UN ANÁLISIS DE COSTO-BENEFICIOS

El análisis de costo – beneficio se realizará mediante el retorno de la inversión (ROI), según la siguiente fórmula:

$$ROI = \frac{\text{Valor para el negocio}}{\text{Costo del proyecto}}$$

Para el costo total del proyecto se ha considerado un costo inicial y costos fijos.

El costo inicial comprende los gastos en cuanto a licencias del software, hardware y capacitación en Inteligencia de Negocios.

Tabla 4
Costos iniciales del proyecto

COSTO INICIAL					
Ítem	Descripción	Cantidad	Precio Unitario	Subtotal	Total + IVA
Equipos	Computador (PC i7; RAM 16GB; Disco 1TB)	1	1600	1600	1792
	Impresora	1	300	300	336
	Scanner	1	1000	1000	1120
Materiales	Resmas de papel	2	4	8	8,96

CONTINÚA



Herramientas	Licencia de software de visualización de datos	1	600	600	672
Movilización	Movilización a la ciudad de Quito	10	100	1000	1120
Estadía	Estadía en la ciudad de Quito	10	30	300	336
Capacitación	Talleres de especialización	2	225	450	504
				Total	5888,96

Los costos fijos mostrados en la tabla 4, corresponden a los gastos anuales que se realizarán en el alojamiento de la base de datos en la nube, actualizaciones y mantenimiento de software y hardware.

Tabla 5
Costos fijos del proyecto

COSTOS FIJOS	
Descripción	Valor
Actualizaciones y mantenimiento Tableau 2018.2	800
Mantenimiento PC	100
Total	900

La proyección de gastos al 2020 se resume en la tabla 6.

Tabla 6
Proyección de gastos

	Inicial	Año 1 (2019)	Año 2 (2020)
Costo	5888,96	900	900

BENEFICIOS

Para el cálculo de beneficios cuantificables, se considera el Modelo de Análisis Costo Beneficio propuesto por la Agencia de Gobierno Electrónico y Sociedad de la Información y del Conocimiento de Uruguay, el cual propone como beneficio la reducción del tiempo (expresado en horas) empleado en este caso, por los docentes de la carrera de Finanzas, para generar los indicadores y actividades de evaluación así como la reducción de gastos asociados a la interacción entre los docentes y los sistemas de información de la ESPOCH.

REDUCCIÓN DEL TIEMPO

Tabla 7

Análisis de la reducción en horas del proceso

	Horas Actuales	Horas Post Implementación	Reducción de horas
Elaboración de informe preliminar de autoevaluación	40	20	20

Valor por hora de los docentes= \$13,21 por hora

Reducción = $20 * 13,21 = \$ 264,20$ mensual

GASTOS

Valor por hora de los docentes= \$13,21 por hora

Reducción = $40 * 3,21 = \$ 528,4$ mensual

Beneficios = $528,4 - 264,20 = \$264,20$

ROI = $264,2 / 5888,96 = 0,045$

El valor del ROI superior a 0 quiere decir que el retorno de la inversión es igual no se gana ni se pierde.

4.2 Planeación

Paso 2: Evaluación de la Infraestructura

Para la evaluación de la infraestructura que requiere el proyecto de Inteligencia de Negocios, se considerará la evaluación de la infraestructura técnica y no técnica, la cual se detalla a continuación:

Infraestructura Técnica

Para evaluar la infraestructura técnica se describirá brevemente la infraestructura tecnológica que posee la institución para luego hacer una evaluación de las plataformas a desarrollar en el caso de estudio.

Al momento, la ESPOCH cuenta con un centro de datos el cual se caracteriza por disponer de las siguientes tecnologías:

- Centro de datos
- Sistema de seguridad perimetral
- Backbone de fibra óptica
- Servidores
- Entre otros

Bajo esta infraestructura se encuentran los sistemas de información como:

- Sistema Académico
- Plataforma de Enseñanza Virtual Moodle
- Correo Institucional
- Sistema para la evaluación al desempeño docente
- Sistemas para la gestión del talento humano, entre otros

Además, dispone de software para la gestión de diagramas empresariales, para la gestión de proyectos empresariales, para la gestión de base de datos básico, para la virtualización de sistemas operativos.

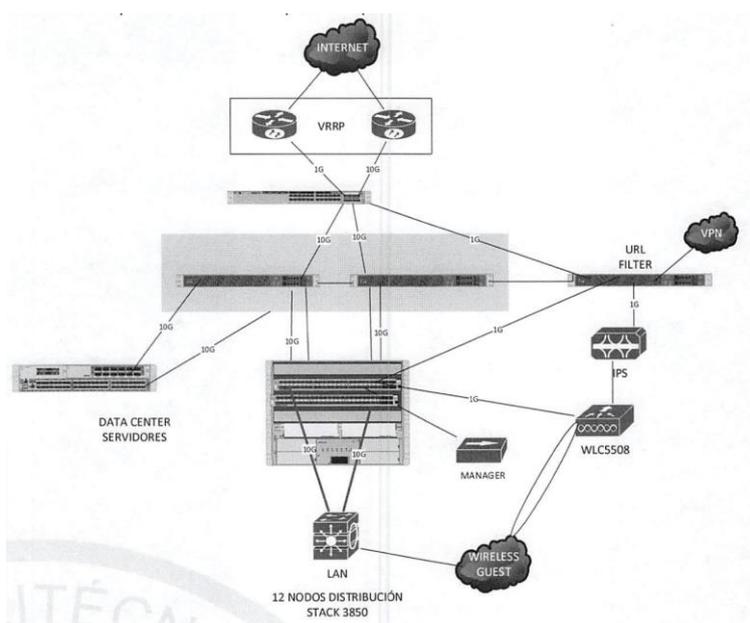


Figura 12. Diagrama de la Infraestructura Tecnológica de la ESPOCH

Por otro lado, la arquitectura del flujo de datos a ser utilizada en el presente proyecto tiene la siguiente estructura:

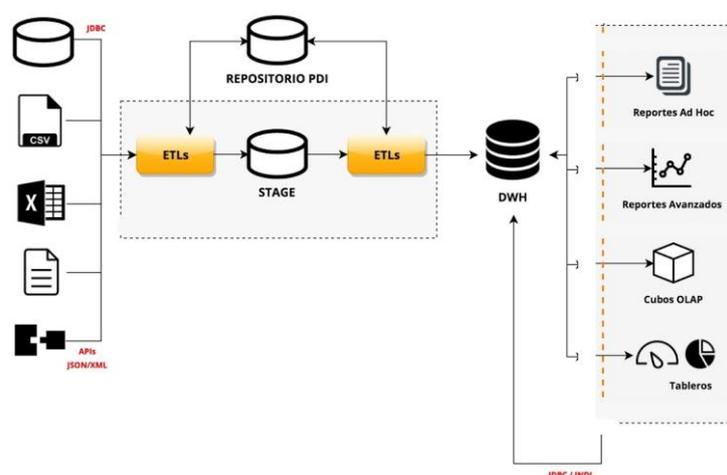


Figura 13. Arquitectura del sistema de Inteligencia de Negocios propuesto

Para el proceso de extracción, transformación y carga de los datos (ETL) al Data warehouse, se utilizará una plataforma de integración de datos (PDI). En el mercado hay un sin número de plataformas de este tipo, sin embargo, para su selección se utilizó la plataforma de

Gartner, la misma que sirvió para comparar las plataformas de integración de Pentaho Hitachi Vantara, Tibco Software y Microsoft, obteniendo como resultado lo siguiente:

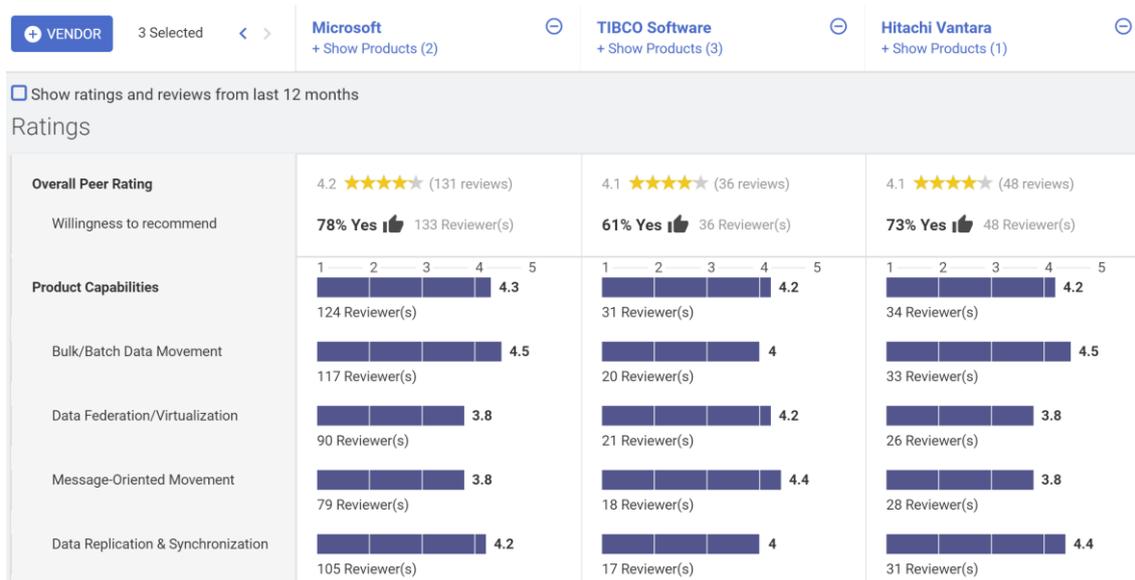


Figura 14. Comparativa de las plataformas de integración

Fuente: <https://www.gartner.com/reviews/market/data-integration-tools/compare/microsoft-vs-tibco-vs-hitachi-vantara>

Los criterios que determinaron la selección de Pentaho Data Integration fueron: el costo, la sincronización y replicación de datos, así como la capacidad de integración con bases de datos estructuradas y no estructuradas (Hitachi Vantara, 2019).

Para la visualización de los datos y la generación de reportes AD-HOC u otros reportes avanzados se utilizará la plataforma de Tableau, la misma que posee un formato self-service, que permite prescindir de los técnicos del departamento TIC, el criterio de Gartner y la facilidad de uso que tiene.



Figura 15. Comparativa de las plataformas de visualización de datos

Fuente: Gartner (Febrero 2017)

Para el desarrollo de la minería de datos se empleará la plataforma de R, debido a que este lenguaje posee todas las librerías de validación que requiere el clustering.

PENTAHO

Pentaho es un conjunto de programas libres, que abordan las distintas fases de la inteligencia de negocios. En la actualidad se perfila como la suite de inteligencia de negocio de código abierto por excelencia, gracias a su versatilidad, facilidad de uso e integración con prácticamente cualquier origen y base de datos. Ofrece dos versiones, la versión pagada llamada Pentaho Enterprise y la versión gratuita conocida como Pentaho Community. (Pentaho Community, 2011)

Los requisitos de hardware y sistema operativo que requiere Pentaho Data Integration para el servidor se muestran en la siguiente tabla:

Tabla 8*Requisitos de Hardware y Sistema Operativo para el Servidor*

Hardware — 64 bit	Sistema Operativo — 64 bit
Procesador: Intel EM64T o AMD64 Dual-Core RAM: 8 GB con 4 GB dedicados a servidores Pentaho Espacio en disco: 20 GB gratis después de la instalación	<ul style="list-style-type: none"> • Microsoft Windows 2012 Server R2 y Microsoft Windows 2016 Server • CentOS 6 y 7 • Red Hat Enterprise 6 y 7 • Ubuntu Server 14.04 LTS y 16.04 LTS • SUSE Linux SLES 11 (SP3 +)

Tabla 9*Requisitos de Hardware y Sistema Operativo para el Cliente*

Hardware — 64 bit	Sistema operativo — 64 bit
Procesadores <ul style="list-style-type: none"> • Apple Macintosh Dual-Core • Intel EM64T o AMD64 Dual-Core • RAM: 2 GB de RAM para la mayoría de las herramientas de diseño, PDI requiere 2 GB dedicados • Espacio en disco: 2 GB gratis después de la instalación • Tamaño mínimo de pantalla: 1280 x 960 	<ul style="list-style-type: none"> • Microsoft Windows 7, 8 y 10 • Ubuntu Desktop 14.04 LTS y 16.04 • OS X 10.12 y 10.13

Fuente: (Pentaho, 2011)

LENGUAJE R

R es un lenguaje y entorno de programación para análisis estadístico y gráfico, proporciona un amplio abanico de herramientas para realizar cálculos numéricos y minería de datos (Goette, 2014).

Los requisitos de instalación de R son:

- Plataforma con el sistema operativo Windows 9x/ME/NT4.0/2000/XP/2003/Vista/7/8/2012 Server/8.1. compatible con Intel.
- Un mínimo de 32 MB de RAM, un ratón y espacio de disco suficiente para archivos, archivos de imagen u otros elementos recuperados.

- Privilegios de administrador para instalar y ejecutar utilidades de R-Studio en Windows NT/2000/XP/2003/Vista/7/8/2012 Server/8.1

(R Tools Technology, 2019)

TABLEAU

Tableau es un software que permite realizar análisis sobre los datos de un negocio, proyecto, web, etc. de forma rápida y sencilla, a la vez que genera gráficos y visualizaciones de los datos del negocio. Esta herramienta permite conectar con bases datos tanto locales como remotos. (Cetatech, 2015)

Los requerimientos mínimos para instalar Tableau en un equipo son:

- Microsoft Windows 7 o una versión más reciente (32 y 64 bits)
- Microsoft Server 2008 R2 o una versión más reciente
- Procesador Intel Pentium 4, AMD Opteron o una versión más reciente
- Memoria de 2 GB
- Espacio libre mínimo en disco de 1,5 GB
- Resolución de pantalla de 1366×768 o superior

(Devoteam, 2018)

Infraestructura no técnica

Al momento la ESPOCH, posee el siguiente mapa de procesos:

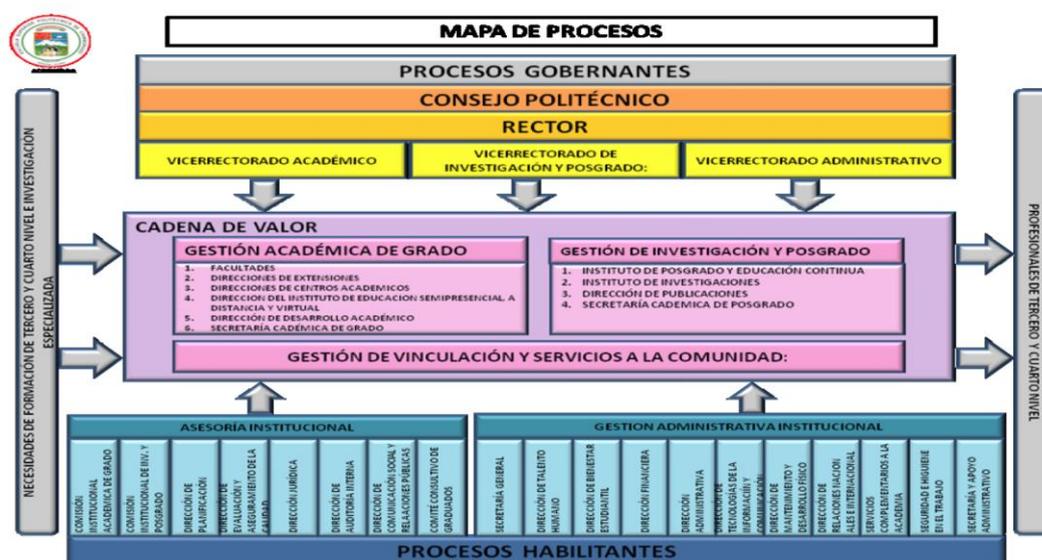


Figura 16. Mapa de Procesos de la ESPOCH

Los sistemas de información que interactúan en los procesos de evaluación son:

- Sistema de Evaluación del Desempeño Docente
- Sistema Académico OASIS
- Sistema de Talento Humano
- Sistema de Publicaciones

Paso 3: Planificación del Proyecto

Para realizar el proyecto se desarrolló una planificación de actividades, las cuales se encuentran distribuidas conforme lo requiere la metodología Business Intelligence Roadmap (BIR), a continuación, en la figura 17 se tiene el detalle:

APLICACIÓN BUSINESS INTELLIGENCE ROADMAP(BIR)	2018 - 2019					
	MES 1	MES 2	MES 3	MES 4	MES 5	MES 5
Fase 1: JUSTIFICACIÓN						
1. Evaluación del negocio	X					
Fase 2: PLANIFICACIÓN						
2. Evaluación de la infraestructura empresarial		X				
3. Planificación del Proyecto		X				
Fase 3: ANÁLISIS DEL NEGOCIO						
4. Definición de requerimientos del proyecto			X			
5. Análisis de Datos			X			
6. Aplicación de prototipos			X			
7. Análisis del repositorio de metadatos			X			
Fase 4: DISEÑO						
8. Diseño de la base de datos				X		
9. Diseño del ETL				X		
10. Diseño del repositorio de Metadatos				X		
Fase 5: CONSTRUCCIÓN						
11. Desarrollo del ETL					X	
12. Desarrollo de la Aplicación					X	
13. Data Mining					X	
14. Desarrollo del Repositorio de Metadatos					X	
Fase 6: DESARROLLO						
15. Implementación						X
16. Evaluación de la versión						X

Figura 17. Planificación del proyecto

4.3 Análisis Del Negocio

Paso 4: Definición de Requisitos del Proyecto

Para definir los indicadores que se obtendrán con las fuentes de información, se realizó el análisis de los modelos de evaluación de carrera e institucional. De estos se obtuvo una tabla resumen que define el modelo, los criterios, indicadores de cada criterio y el tipo de indicador, es decir si este es cualitativo o cuantitativo.

Tabla 10*Análisis del modelo de evaluación de carreras según el tipo de indicador*

MODELO	CRITERIO	SUBCRITERIO	INDICADOR	TIPO DE INDICADOR
MODELO GENÉRICO DE EVALUACIÓN DEL ENTORNO DE APRENDIZAJE DE CARRERAS PRESENCIALES Y SEMIPRESENCIALES DE LAS UNIVERSIDADES Y ESCUELAS POLITÉCNICAS DEL ECUADOR 2015	PERTINENCIA (A)	A.1. CONTEXTO	A.1.1. ESTADO ACTUAL Y PROSPECTIVA	CUALITATIVA
			A.1.2. PROYECTOS-PROGRAMAS DE VINCULACIÓN CON LA SOCIEDAD	CUALITATIVA
		A.2. PROFESIÓN	A.2.1. PERFIL PROFESIONAL	CUALITATIVA
	PLAN CURRICULAR (B)	B.1. MACROCURRÍCULO	B.1.1. PERFIL DE EGRESO	CUALITATIVA
			B.1.2. ESTRUCTURA CURRICULAR	CUALITATIVA
		B.2. MESOCURRÍCULO	B.2.1. PLAN DE ESTUDIOS	CUALITATIVA
		B.3. MICROCURRÍCULO	B.3.1. PROGRAMAS DE LAS ASIGNATURAS	CUALITATIVA
			B.3.2. PRÁCTICAS EN RELACIÓN A LAS ASIGNATURAS	CUALITATIVA
	ACADEMIA (C)	C.1. CALIDAD DOCENTE	C.1.1. AFINIDAD FORMACIÓN- POSGRADO	CUANTITATIVO
			C.1.2. ACTUALIZACIÓN CIENTÍFICA Y-O PEDAGÓGICA	CUANTITATIVO
			C.1.3. TITULARIDAD	CUANTITATIVO
		C.2. DEDICACIÓN	C.2.1. PROFESOR TC, MT, TP	CUANTITATIVO
			C.2.2. ESTUDIANTES POR PROFESOR	CUANTITATIVO
			C.2.3. DISTRIBUCIÓN HORARIA	CUALITATIVA
		C.3. PRODUCCIÓN ACADÉMICA	C.3.1. PRODUCCIÓN ACADÉMICA-CIENTÍFICA	CUANTITATIVO
			C.3.2. PRODUCCIÓN REGIONAL	CUANTITATIVO
			C.3.3. LIBROS O CAPÍTULOS DE LIBROS	CUANTITATIVO
			C.3.4. PONENCIAS	CUANTITATIVO
	AMBIENTE INSTITUCIONAL (D)	D.1. GESTIÓN ACADÉMICA	D.1.1. DIRECCIÓN-COORDINACIÓN ACADÉMICA	CUALITATIVA
			D.1.2. EVALUACIÓN DEL DESEMPEÑO DOCENTE	CUALITATIVA
			D.1.3. SEGUIMIENTO DEL SÍLABO	CUALITATIVA
			D.1.4. SEGUIMIENTO AL PROCESO DE TITULACIÓN	CUALITATIVA
			D.1.5. SEGUIMIENTO A GRADUADOS	CUALITATIVA
D.1.6. SEGUIMIENTO DE PRÁCTICAS PREPROFESIONALES			CUALITATIVA	
D.1.7. CALIDAD DE LA INFORMACIÓN			CUANTITATIVA	

CONTINÚA



MODELO DE EVALUACIÓN INSTITUCIONAL DE UNIVERSIDADES Y ESCUELAS POLITÉCNICAS 2018				
MODELO DE EVALUACIÓN INSTITUCIONAL DE UNIVERSIDADES Y ESCUELAS POLITÉCNICAS 2018	ESTUDIANTES (E)	D.2. FONDO BIBLIOGRÁFICO	D.2.1. BIBLIOGRAFÍA BÁSICA	CUANTITATIVA
			D.2.2. FONDO BIBLIOGRÁFICO	CUANTITATIVA
		D.3. LABORATORIOS-CENTROS DE SIMULACIÓN-TALLERES	D.3.1. FUNCIONALIDAD	CUANTITATIVO
			D.3.2. EQUIPAMIENTO	CUANTITATIVO
			D.3.3. DISPONIBILIDAD	CUANTITATIVO
		E.1.PARTICIPACIÓN ESTUDIANTIL	E.1.1. TUTORÍAS	CUALITATIVO
			E.1.2. ACTIVIDADES COMPLEMENTARIAS	CUALITATIVO
			E.1.3. ACTIVIDADES VINCULADAS CON LA COLECTIVIDAD	CUALITATIVO
			E.1.4. BIENESTAR ESTUDIANTIL	CUALITATIVO
			E.1.5. PARTICIPACIÓN EN LA ACREDITACIÓN	CUALITATIVO
	E.2.EFICIENCIA	E.2.1. TASA DE RETENCIÓN	CUANTITATIVO	
		E.2.2. TASA DE TITULACIÓN	CUANTITATIVO	
	1. ORGANIZACIÓN	1.1 MISION, VISION Y PLANIFICACIÓN INSTITUCIONAL	1.1.1 PLANIFICACIÓN ESTRATÉGICA	CUALITATIVO
			1.1.2 PLANIFICACIÓN OPERATIVA	CUALITATIVO
		1.2 GESTIÓN	1.2.1 POLÍTICAS SOBRE ACCIÓN AFIRMATIVA	CUALITATIVO
			1.2.2 GESTIÓN DE LA OFERTTA ACADÉMICA	CUALITATIVO
			1.2.3 GESTIÓN DEL CLAUSTRO	CUALITATIVO
			1.2.4 SISTEMA DE GESTIÓN DE LA CALIDAD	CUALITATIVO
			1.2.5 POLÍTICAS SOBRE INTERNACIONALIZACIÓN	CUALITATIVO
		1.3 CONTROL ÉTICO Y TRANSPARENCIA	1.3.1 PROMOCIÓN Y CONTROL ÉTICO	CUALITATIVO
		1.3.2 TRANSPARENCIA Y DIFUSIÓN	CUALITATIVO	
2. CLAUSTRO DE PROFESORES		2.1 FORMACIÓN DOCTORAL	2.1.1 DOCTORES	CUANTITATIVO
		2.1.2 DOCTORES A TC	CUANTITATIVO	
		2.1.3 DOCTORES EN FORMACIÓN	CUANTITATIVO	
	2.2 SUFICIENCIA Y DEDICACIÓN	2.2.1 ESTUDIANTES POR PROFESORA TC	CUANTITATIVO	
		2.2.2 TITULARIDAD TC	CUANTITATIVO	
		2.2.3 DISTRIBUCIÓN TEMPORAL DE ACTIVIDADES	CUALITATIVO	
	2.3 CARRERA DEL PROFESOR	2.3.1 TITULARIDAD	CUANTITATIVO	
		2.3.2 EVALUACIÓN DE LOS PROFESORES	CUALITATIVO	
		2.3.3 PROFESORAS MUJERES	CUANTITATIVO	
		2.3.4 DIRECCIÓN MUJERES	CUANTITATIVO	
3. INVERSIÓN	3.1 INSTITUCIONALIZACIÓN	3.1.1 ORGANIZACIÓN Y PLANIFICACIÓN DE LA INVESTIGACIÓN CIENTÍFICA	CUALITATIVO	

CONTINÚA 

		3.1.2 GESTIÓN DE RECURSOS PARA LA INVESTIGACIÓN	CUALITATIVO
	3.2 RESULTADOS	3.2.1 PRODUCCIÓN CIENTÍFICA EN REVISTAS DE IMPACTO MUNDIAL	CUANTITATIVO
		3.2.2 PRODUCCIÓN CIENTÍFICA EN REVISTAS DE IMPACTO REGIONAL	CUANTITATIVO
		3.2.3 PRODUCCIÓN DE OBRAS	CUALITATIVO
4. VINCULACIÓN CON LA SOCIEDAD	4.1 INSTITUCIONALIZACIÓN	4.1.1 ORGANIZACIÓN Y PLANIFICACIÓN DE LA VINCULACIÓN CON LA SOCIEDAD	CUALITATIVO
		4.1.2 EJECUCIÓN DE LAS ACTIVIDADES DE VINCULACIÓN	CUALITATIVO
	4.2 RESULTADOS	4.1.1 PROGRAMAS DE VINCULACIÓN CON LA SOCIEDAD Y SU IMPACTO	CUALITATIVO
5. RECURSOS E INFRAESTRUCTURA	5.1 INFRAESTRUCTURA	5.1.1 ENTORNO DE LAS ACTIVIDADES ACADÉMICAS	CUALITATIVO
		5.1.2 ESPACIOS DE BIENESTAR	CUALITATIVO
		5.1.3 SISTEMAS INFORMÁTICOS	CUALITATIVO
	5.2 BIBLIOTECAS	5.2.1 BIBLIOTECAS	CUALITATIVO
6. ESTUDIANTES	6.1 ADMISIÓN Y SEGUIMIENTO	6.1.1 PROCESOS DE ADMISIÓN Y NIVELACIÓN A LAS CARRERAS DE GRADO	CUALITATIVO
		6.1.2 PROCESOS DE ADMISIÓN A POSTGRADO	CUALITATIVO
		6.1.3 SEGUIMIENTO A GRADUADOS	CUALITATIVO
	6.2 RESULTADOS DE LA ADMISIÓN Y SEGUIMIENTO	6.2.1 TASA DE RETENCIÓN	CUANTITATIVO
		6.2.2 TASA DE GRADUACIÓN DE GRADO	CUANTITATIVO
		6.2.3 TASA DE GRADUACIÓN DE POSTGRADO	CUANTITATIVO

Para el desarrollo del modelo de evaluación se identificaron los indicadores cuantitativos que posee cada modelo y se realizó un análisis que permitió determinar las fuentes de información que posee la ESPOCH, para obtener los datos que permitirán medir los indicadores de evaluación. La siguiente tabla resume el análisis realizado:

Tabla 11

Análisis de indicadores, fuentes de datos y tablas relacionadas

MODELO	CRITERIO	SUBCRITERIO	INDICADOR	FÓRMULA DE CÁLCULO	FUENTE DE DATOS	TABLAS/ENTIDADES RELACIONADAS	
MODELO GENÉRICO DE EVALUACIÓN DEL ENTORNO DE APRENDIZAJE DE CARRERAS PRESENCIALES Y SEMIPRESENCIALES DE LAS UNIVERSIDADES Y ESCUELAS POLITÉCNICAS DEL ECUADOR 2015	ACADEMIA (C)	C.1. CALIDAD DOCENTE	C.1.1. AFINIDAD FORMACIÓN- POSGRADO	ESCALA DEL INDICADOR / FORMA DE CÁLCULO: $AFP = \frac{1,5 \cdot N_{A\text{PhD}} + N_{AMSc}}{N_A}$ Dónde: AFP: Afinidad formación de posgrado. AphD: Número de asignaturas impartidas por profesores con un título de PhD afín a la asignatura. NAMSc: Número de asignaturas impartidas por profesores con un título de MSc o Especialidad afín a la asignatura. NA: Número total de asignaturas impartidas.	Base de Datos Talento Humano Sistema Académico	Docente Asignatura Títulos Periodo académico Campos de conocimiento Carrera	
			C.1.2. ACTUALIZACIÓN CIENTÍFICA Y-O PEDAGÓGICA	ESCALA DEL INDICADOR / FORMA DE CÁLCULO $ACP = 100 \cdot \frac{NDAD}{NTD}$ Dónde: ACP: Actualización científica y/o pedagógica. NDAD: Total de profesores que han asistido a eventos de actualización científica y/o pedagógica afín a la carrera. NTD: Total de profesores que han impartido clases en la carrera durante el periodo de evaluación.	Talento Humano Sistema Académico	Docente Asignatura Periodo académico Carrera Documento reporte asistencia a capacitaciones	
			C.1.3. TITULARIDAD	ESCALA DEL INDICADOR / FORMA DE CÁLCULO $TIT = 100 \cdot \frac{NDT}{NTD}$ TIT: Titularidad. NDT: Total de profesores titulares que han impartido clases en la carrera. NTD: Total de profesores que han impartido clases en la carrera durante el periodo de evaluación.	Talento Humano Sistema Académico	Docentes Periodo académico Carrera Relación laboral	
		C.2. DEDICACIÓN	C.2.1. PROFESOR TC, MT, TP	ESCALA DEL INDICADOR / FORMA DE CÁLCULO $D = \frac{TC + 0,5 \cdot MT + 0,25 \cdot TP}{NTD}$ D: Profesores de TC, MT, TP TC: Profesores a tiempo completo. MT: Profesores a medio tiempo. TP: Profesores a tiempo parcial. NTD: Total de profesores que han impartido clases en la carrera durante el periodo de evaluación.	Talento Humano Sistema Académico	Docentes Tiempo de dedicación Carrera Periodo académico	
				C.2.2. ESTUDIANTES POR PROFESOR	ESCALA DEL INDICADOR / FORMA DE CÁLCULO $ED = \frac{NE}{TC + 0,5 \cdot MT + 0,25 \cdot TP}$ ED: Estudiantes por profesor a tiempo completo, medio tiempo y tiempo parcial NE: Total de estudiantes de la carrera. TC: Total de profesores de tiempo completo. MT: Total de profesores de medio tiempo. TP: Total de profesores de tiempo parcial.	Sistema Académico Publicaciones Talento Humano	Docentes Tiempo de Dedicación Estudiantes Matriculados Carrera Periodo Académico

CONTINÚA



C.3. PRODUCCIÓN ACADÉMICA	C.3.1. PRODUCCIÓN ACADÉMICA- CIENTÍFICA	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $IPAC = \frac{1}{N} \left(\sum_{i=1}^M 1 + 3,61SJR_i \right)$ <p>IPAC: Indicador de producción académica científica. N: Número de profesores de la carrera durante el semestre en el que se efectúa la evaluación. M: Número de artículos académicos publicados por los profesores/investigadores de la carrera durante el período de evaluación. SJR: Índice de impacto calculado por Scimago Journal Rank.</p>	Sistema Académico Talento Humano Publicaciones	Docentes Publicaciones Artículos publicados Impacto de los artículos Carrera
	C.3.2. PRODUCCIÓN REGIONAL	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $IPR = \frac{1}{N} \sum_{i=1}^M PR_i$ <p>Donde:</p> <p>IPR= Indicador de producción académica - científica regional. N= Número de profesores de la carrera. M= Total de artículos o trabajos científicos de los profesores/investigadores de la carrera, publicados o aceptados en revistas que constan en bases de datos, cuyos criterios de indexación contemplen parámetros de calidad reconocidos regionalmente. PR_i: Artículo o trabajo científicos del profesor/investigador i, de la carrera.</p>	Sistema Académico Talento Humano Publicaciones	Docentes Publicaciones Artículos publicados Impacto de los artículos Carrera Periodo académico
	C.3.3. LIBROS O CAPÍTULOS DE LIBROS	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $LCL = \frac{1}{N} (L + 0,5CL)$ <p>Donde:</p> <p>LCL: Libros o capítulos de libros L: Número de libros publicados por profesores o investigadores de la carrera. CL: Capítulos de libros publicados por profesores de la carrera. N: Número de profesores de la carrera.</p>	Académico Publicaciones	Docente Carrera Periodo académico Libros publicados

CONTINÚA



	C.3.4. PONENCIAS	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $P = \frac{1}{N} \sum_{i=1}^M (1 + \beta_i)$ <p>Donde:</p> <p>P: Indicador per cápita de la presentación de artículos o investigaciones de profesores de la carrera en eventos académicos o científicos, nacionales o internacionales.</p> <p>N: Número total de profesores de la carrera al período de evaluación.</p> <p>M: Número total de ponencias realizadas por profesores/ investigadores de la carrera, que hacen filiar a la universidad.</p> <p>β_i Puntuación de 0; 0,5 ó 1 de acuerdo a la relevancia del evento debidamente argumentada por el evaluador, considerando la siguiente escala:</p> <p>Altamente Relevante: (1): Eventos que cumplan con dos condiciones:</p> <ol style="list-style-type: none"> 1.- participan al menos tres expertos/académicos internacionales con trayectoria. 2.- el evento se ha realizado al menos por cinco ocasiones consecutivas. <p>Relevante: (0.5): Eventos en los que participan al menos tres expertos/académicos internacionales con trayectoria.</p> <p>No relevante (0). - El evento no corresponde a las categorías de Altamente relevante o Relevante. El factor β_i tiene un valor de 0.</p>	Sistema Académico Talento Humano Publicaciones	Docentes Ponencias realizadas Periodo académico Carrera	
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">AMBIENTE INSTITUCIONAL (D)</p>	D.1. GESTIÓN ACADÉMICA	D.1.7. CALIDAD DE LA INFORMACIÓN	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $IE = 100 \cdot \frac{NVC}{NTV}$ <p>Donde:</p> <p>NVC: Número de variables de buena calidad.</p> <p>NTV: Número total de variables del modelo.</p>	Sistema de información no disponible	

CONTINÚA



D.2. FONDO BIBLIOGRÁFICO	D.2.1. BIBLIOGRAFÍA BÁSICA	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $BB = \frac{1}{Q} \sum_{i=1}^Q \frac{LIB_i}{NE_i}$ <p>Donde:</p> <p>BB: Bibliografía Básica LIB_i: Número de ejemplares del libro i constatados en la biblioteca. NE_i: Número de estudiantes que tomaron asignaturas donde consta el libro i como bibliografía básica. Q: Número de asignaturas evaluadas (10).</p>	Sistema de información no disponible
	D.2.2. FONDO BIBLIOGRÁFICO	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> <p>CB= 0, 25.PER+0.3.ACE+0, 25.CON+0, 2.PRE</p> <p>Donde:</p> <p>CB: Calidad bibliográfica PER: Pertinencia de la bibliografía ACE: Acervo bibliográfico CON: Conservación del material bibliográfico PRE: Presupuesto ejecutado en la adquisición del material bibliográfico</p> <p>Todos los aspectos calificados en escala de 1 a 10.</p>	Sistema de información no disponible
D.3. LABORATORIO S-CENTROS DE SIMULACIÓN-TALLERES	D.3.1. FUNCIONALIDAD	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $FUN = \frac{1}{Q} \sum_{i=1}^Q FUN_i$ <p>FUN: Funcionalidad de laboratorios y talleres FUN_i: Valoración obtenida por cada laboratorio y taller de acuerdo al aspecto de funcionalidad, evaluado en una escala del 1 al 10. Q: Número de laboratorios y talleres evaluados.</p>	Sistema de información no disponible
	D.3.2. EQUIPAMIENTO	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $EQ = \frac{1}{Q} \sum_{i=1}^Q EQ_i$ <p>Donde:</p> <p>EQ: Equipamiento EQ_i: Valoración obtenida por cada laboratorio y taller, de acuerdo al aspecto de equipamiento, evaluado en una escala del 1 al 10. Q: Número de laboratorios y talleres evaluados.</p>	Sistema de información no disponible
	D.3.3. DISPONIBILIDAD	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $DIS = \frac{1}{Q} \sum_{i=1}^Q DIS_i$ <p>Donde:</p> <p>DIS: Disponibilidad. DIS_i: Valoración obtenida por cada laboratorio o taller de acuerdo al aspecto de disponibilidad, Evaluado escala del 1 al 10. Q: Número de laboratorios y talleres evaluados.</p>	Sistema de información no disponible



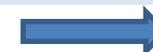
MODELO DE EVALUACIÓN INSTITUCIONAL DE UNIVERSIDADES Y ESCUELAS POLITÉCNICAS 2018	ESTUDIANTES (E)	E.2.EFICIENCIA	E.2.1. TASA DE RETENCIÓN	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $TR = 100 \cdot \frac{NEMA}{NTEA}$ <p>Donde:</p> <p>TR: Tasa de retención NEMA: Número de estudiantes matriculados en la carrera en el periodo de evaluación que fueron admitidos dos años antes. NTEA: Número total de estudiantes que fueron admitidos en la carrera dos años antes del periodo de evaluación.</p>	Indicador calculado por el sistema académico	
			E.2.2. TASA DE TITULACIÓN	<p>ESCALA DEL INDICADOR / FORMA DE CÁLCULO</p> $TG = \frac{\sum_i NEG}{\sum_i NEC}$ <p>Donde:</p> <p>TG: Tasa de titulación. NEG: Número de estudiantes graduados. NEC: Número de estudiantes que ingresaron en la (s) cohorte (s).</p>	Indicador calculado por el sistema académico	
		2.1 FORMACIÓN DOCTORAL	2.1.1 DOCTORES	$PD = 100 \frac{Total\ Doctores}{Total\ Profesores}$ <p>La institución cuenta con una cantidad mínima de doctores que permite desarrollar las actividades académicas de la institución. Se considera un porcentaje adecuado el 40%.</p>	Sistema académico Talento Humano	Docente Título de cuarto nivel
	2.1.2 DOCTORES A TC	$PDTC = 100 \frac{Total\ Doctores\ TC}{Total\ Profesores\ TC}$ <p>Es esencial la participación efectiva de los doctores en las actividades académicas, las cuales demandan dedicación completa para un desarrollo institucional sostenible. Se considera adecuado que al menos el 80% de los doctores de la institución tengan dedicación a tiempo completo; esto es, al menos el 32 % de los profesores, lo cual constituye aproximadamente el 53% de los profesores a tiempo completo (60% de profesores). Por tanto, de acuerdo a la definición de este indicador, el estándar se ubica en 53%.</p>	Sistema académico Talento Humano	Docentes Tiempo dedicación Titulo cuarto nivel		
	2.1.3 DOCTORES EN FORMACIÓN	$FD = 100 \frac{Total\ Doctores\ en\ formación}{0,4 Total\ profesores - Total\ Doctores}$ <p>Según la definición del indicador este se cumple cuando están formándose todos los profesores requeridos para completar el 40% de doctores exigidos; por lo tanto, el estándar para este indicador es el 100%.</p> <p>La estrategia de formación doctoral debe tender a la consecución del porcentaje mínimo de doctores requeridos en el estándar del indicador Doctores; el cumplimiento del estándar de ese indicador implica el cumplimiento también de este indicador (Si $0,4 Total\ profesores - Total\ Doctores \leq 0$, entonces $FD=100\%$).</p>	Sistema académico Talento Humano	Docentes Tiempo dedicación		

CONTINÚA



2.2 SUFICIENCIA Y DEDICACIÓN	2.2.1 ESTUDIANTES POR PROFESORA TC	$ETC = \frac{NEP + 0,5(NESP + NED)}{Total\ Profesores\ a\ TC}$ <p>La IES debe contar con un claustro de profesores suficiente para el desarrollo de las actividades sustantivas de la universidad, en particular, para la formación, acompañamiento y guía de los estudiantes. Se considera adecuada una relación máxima de 30 estudiantes por profesor a tiempo completo</p>	Sistema académico Talento Humano Talento Humano	Docentes Tiempo dedicación Estudiantes Periodo académico
	2.2.2 TITULARIDAD TC	$TTC = 100 \frac{Total\ titulares\ a\ tiempo\ completo}{0,6\ total\ profesores}$ <p>Se espera que al menos el 75% de profesores con dedicación a tiempo completo sea titular.</p>	Sistema académico Talento Humano	Docentes Tiempo dedicación Periodo académico Relación laboral
2.3 CARRERA DEL PROFESOR	2.3.1 TITULARIDAD	$TT = 100 \frac{Total\ titulares}{total\ profesores}$ <p>Para garantizar la estabilidad y efectividad en el desarrollo de los procesos sustantivos se requiere que el 60% de los profesores sean titulares.</p>	Sistema académico Talento Humano	Docentes Periodo académico Relación laboral
	2.3.3 PROFESORAS MUJERES	$DM = 100 \cdot \frac{NMTI}{NTI}$ <p>Donde: DM= Porcentaje de profesoras e investigadoras titulares de la institución NMTI= Número de profesoras e investigadoras titulares de la institución NTI= Número de profesores e investigadores titulares de la institución</p> <p>Se considera adecuada una relación de al menos 2/3, y de máximo 3/2 entre la cantidad de profesoras y profesores titulares; esto es, entre un 40% y 60% de profesoras, y viceversa. Este es un intervalo razonable que toma en cuenta el equilibrio que debe existir en la participación femenina y la masculina dentro del claustro de profesores titulares.</p>	Sistema académico Talento Humano	Docentes Periodo académico Relación laboral
	2.3.4 DIRECCIÓN MUJERES	$DM = 100 \cdot \frac{NMCA}{NCD}$ <p>Donde: DM = Porcentaje de mujeres en cargos de dirección académica considerados en el análisis NMCA = Número de mujeres en cargos de dirección académica considerados en el análisis NCD = Número de cargos de dirección académica considerados en el análisis</p> <p>La participación equitativa de hombres y mujeres evidencia los resultados de las políticas de la institución en esta materia; sin embargo, se considera que debido a su naturaleza el rango de tolerancia para definir el estándar debe ser más amplio, en este caso, al menos el 30% de los puestos directivos deben estar ocupados por mujeres, hasta un máximo de 70%.</p>	Sistema académico Talento Humano	Docentes Género Roles

CONTINÚA



3. INVESTIGACIÓN	3.2 RESULTADOS	3.2.1 PRODUCCIÓN CIENTÍFICA EN REVISTAS DE IMPACTO MUNDIAL	$PRM = \frac{NA_1}{0,36(NP)^2} \sum_{i=1}^{NART_1} (1 + I_i)$ <p> <i>PRM</i> = Producción en revistas de impacto mundial¹ <i>NA₁</i> = Número autores que han publicado en revistas de impacto mundial <i>NART₁</i> = Número de artículos publicados en revistas de impacto mundial <i>NP</i> = Número total de profesores de la IES <i>I_i</i> = Índice de impacto que se basa en la información de los índices de impacto bibliométrico que proveen las indexadoras (<i>SJR</i>, impact factor u otros) $\frac{NA_1}{0,36NP}$ = Relación entre el número de autores y el número mínimo de profesores con dedicación a tiempo completo (60% de profesores) </p>	Sistema académico Talento Humano	Docentes Publicaciones Artículos publicados Impacto de los artículos Carrera
		3.2.2 PRODUCCIÓN CIENTÍFICA EN REVISTAS DE IMPACTO REGIONAL	$PRR = \frac{NA_2}{(0,6NP - NA_1)} \frac{NART_2}{(0,6NP - NA_1)}$ <p>Se define $F = 0,6NP - NA_1$</p> $PRR = \frac{NA_2 * NART_2}{F^2}$ <p> <i>PRR</i> = Producción en revistas de impacto regional <i>NA₂</i> = Número autores que han publicado en revistas de impacto regional <i>NART₂</i> = Número de artículos publicados en revistas de impacto regional <i>NP</i> = Número total de profesores de la IES <i>NA₁</i> = Número autores que han publicado en revistas de impacto mundial <i>F</i> = Diferencia entre el número mínimo de profesores con dedicación a tiempo completo que debe tener la institución y el número de autores que han publicado en revistas de impacto mundial. </p>	Sistema académico Talento Humano	Docentes Publicaciones Artículos publicados Impacto de los artículos Carrera
6. ESTUDIANTES	6.2 RESULTADOS DE LA ADMISIÓN Y SEGUIMIENTO	6.2.1 TASA DE RETENCIÓN	$TRG = \frac{NEMA}{NTEA}$ <p><i>TRG</i> = Tasa de Retención de grado</p> <p> <i>NEMA</i> = Número de estudiantes matriculados durante el período académico ordinario en el que se efectúa la evaluación de la institución, que fueron admitidos dos años antes <i>NTEA</i> = Número total de estudiantes que fueron admitidos en una carrera de la IES dos años antes </p> <p>La Tasa de retención es un resultado de las buenas prácticas en el proceso formativo, se espera que el 80% de los estudiantes siga cursando sus estudios en la IES.</p>	Sistema académico Talento Humano	Docentes Publicaciones Artículos publicados Impacto de los artículos Carrera

CONTINÚA



<p>6.2.2 TASA DE GRADUACIÓN DE GRADO</p>	$TGG = \frac{NEG}{NEC}$ <p> <i>TGG</i> = Tasa de graduación de grado <i>NEG</i> = Número de estudiantes de grado que ingresaron en las cohortes definidas y se graduaron hasta el final del último periodo académico regular concluido antes de la evaluación. <i>NEC</i> = Número de estudiantes de grado que ingresaron en las cohortes definidas </p> <p>La adecuada formación producto del seguimiento e involucramiento de los estudiantes en los proyectos académicos contribuye a su graduación en el tiempo previsto. En este sentido, la eficiencia es un parámetro de la calidad de la formación impartida por la IES. Se espera que al menos el 80% de los estudiantes de las cohortes analizadas se hayan graduado en el tiempo previsto.</p>	<p>Sistema académico Talento Humano</p>	<p>Docentes Publicaciones Artículos publicados Impacto de los artículos Carrera</p>
<p>6.2.3 TASA DE GRADUACIÓN DE POSTGRADO</p>	$TGP = \frac{NEGP}{NECP}$ <p> <i>TGP</i> = Tasa de graduación de posgrado <i>NEGP</i> = Número de estudiantes de posgrado que ingresaron en las cohortes definidas y se graduaron hasta el final del último periodo académico regular concluido antes de la evaluación. <i>NECP</i> = Número de estudiantes de posgrado que ingresaron en las cohortes definidas </p> <p>La adecuada formación producto del seguimiento e involucramiento de los estudiantes de posgrado en los proyectos académicos, contribuye a su graduación en el tiempo previsto. En este sentido, la eficiencia es un parámetro de la calidad de la formación impartida por la IES. Se espera que al menos el 80% de los estudiantes de las cohortes analizadas se hayan graduado en el tiempo previsto.</p>	<p>Sistema académico Talento Humano</p>	<p>Docentes Publicaciones Artículos publicados Impacto de los artículos Carrera</p>

Fuente: Modelos de Evaluación CACES, elaboración propia

La tabla contiene la información del modelo de evaluación, el criterio, el indicador, la fórmula de cálculo, la fuente de información y las tablas o entidades de las cuales se obtendrán los datos para el Data Warehouse.

Paso 5: Análisis de Datos

En esta fase del proyecto, se realizó el análisis de datos del sistema académico de la ESPOCH y de los archivos que almacenan la información sobre la producción científica. A continuación, se tiene un ejemplo del esquema de la base de datos que se analizó.

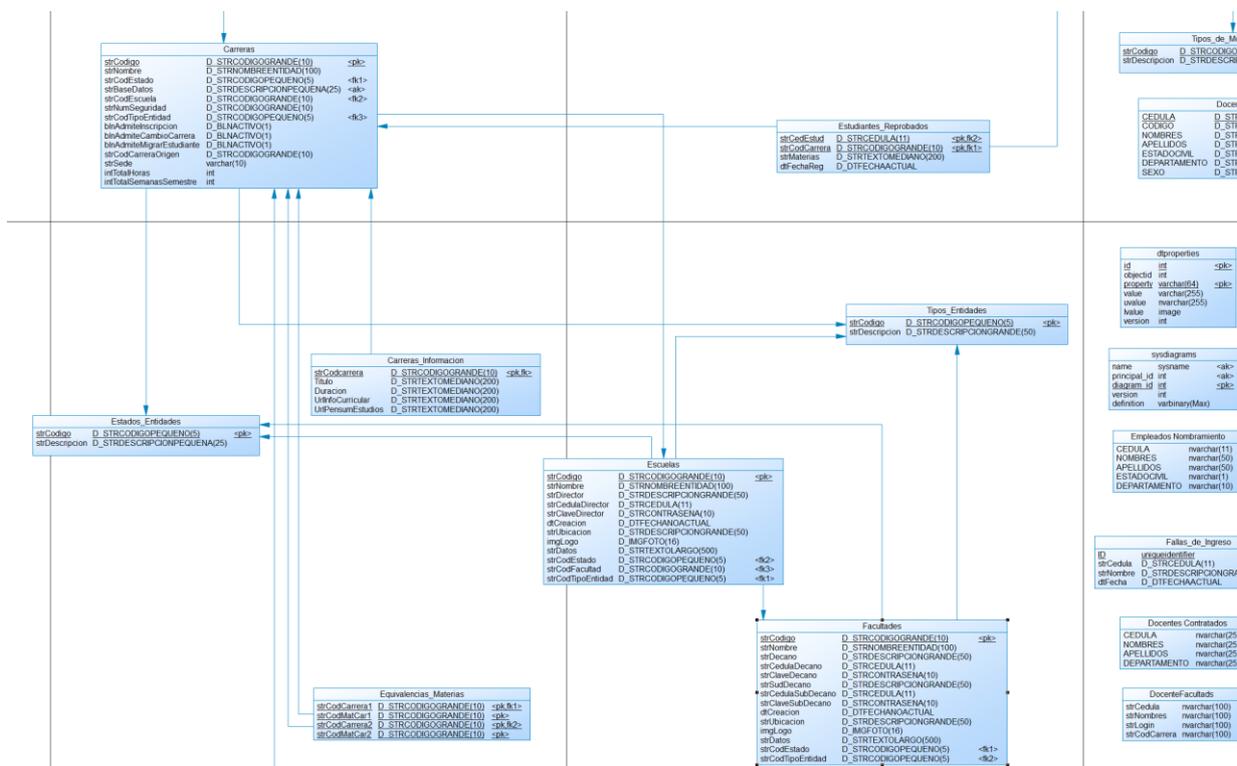


Figura 18. Modelo de Entidad Relación del Sistema Académico

Paso 6: Prototipo de Aplicación

La arquitectura del sistema de gestión de indicadores tendrá la estructura que muestra la figura 19, se creará una base de datos stage(dsa_evaluacion), la misma que es un espacio orientado a almacenar información proveniente de los sistemas de información operacionales y de otras fuentes, que será el punto de partida de los procesos de depuración, transformación y carga en el Data Warehouse (dwh_evaluacion).

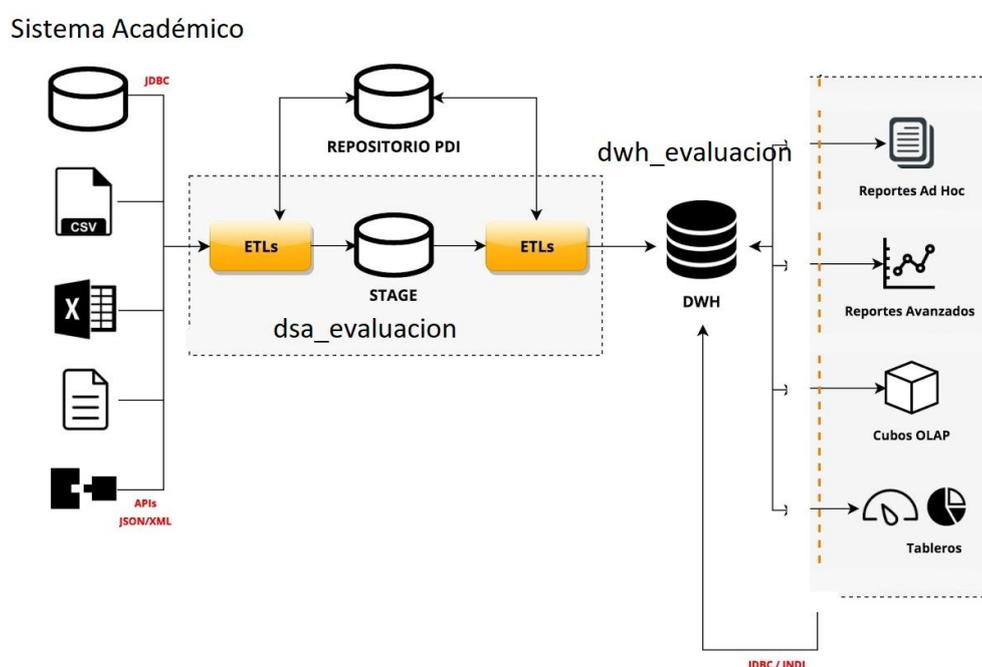


Figura 19. Arquitectura del sistema de gestión de indicadores

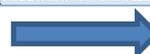
Paso 7: Análisis del Repositorio de Metadatos

En esta fase se realizó el análisis de los metadatos de los diferentes repositorios que posee la ESPOCH, a continuación, se detallan los datos que se obtendrán de cada una de las tablas o entidades a considerar para el modelo de gestión de indicadores:

Tabla 12
Análisis del repositorio de metadata

Tabla	Campos	Descripción
Docentes	TIPO_IDENTIFICACIÓN	Tipo de identificación del profesor
	IDENTIFICACION	Número de documento de identificación del profesor
	PRIMER_APELLIDO	Primer apellido del profesor
	SEGUNDO_APELLIDO	Segundo apellido del profesor
	PRIMER_NOMBRE	Primer nombre del profesor
	SEGUNDO_NOMBRE	Segundo nombre del profesor
	SEXO	Sexo del profesor
	FECHA_NACIMIENTO	Fecha de nacimiento del profesor
	PAIS_NACIONALIDAD	País de la nacionalidad del profesor
	DISCAPACIDAD	Discapacidad del profesor
	NUMERO_CONADIS	Número de carnet conadis
	DIRECCION_DOMICILIO	Dirección del profesor
	EMAIL_PERSONAL	E-mail personal del profesor
	EMAIL_INSTITUCIONAL	E-mail institucional del profesor
	TIPO_DOCUMENTO	Tipo de documento de relación laboral con la ESPOCH asignado al profesor
	NUMERO_DOCUMENTO	Número del documento de relación laboral con la ESPOCH del profesor.
	INGRESO_POR_CONCURSO	Ingresar si el profesor ingresó a la Institución por concurso público
	RELACION_IES	Ingresar la relación laboral de profesor con la IES
	TIEMPO_DEDICACION	Ingresar el tiempo de dedicación del profesor con la ESPOCH, definido en su documento contractual
	CATEGORIA	Ingresar la categoría del profesor TITULAR
	REMUNERACION_MENSUAL	Ingresar la remuneración mensual del profesor
	NIVEL_CLASES	Nivel al que dicta clase
	FECHA_INGRESO	Fecha de ingreso a la a la IES, para los profesores TITULARES
	FECHA_INICIO	Fecha de inicio del contrato vigente, en el período de evaluación, del profesor no titular
	FECHA_FINALIZACION	Fecha de finalización del contrato vigente, en el período de evaluación, del profesor no titular
	SEDE	Sede donde labora el profesor
	FACULTAD	Nombre de la facultad donde labora el profesor (en caso de laborar en facultades)
	CARRERA_1	Nombre de la carrera 1 donde labora el profesor (en caso de laborar en facultades)
	CARRERA_2	Nombre de la carrera 2 donde labora el profesor (en caso de laborar en facultades)
	CARRERA_3	Nombre de la carrera 3 donde labora el profesor (en caso de laborar en facultades)
	CENTRO_ACADEMICO	Nombre del centro académico donde labora el profesor (en caso de laborar en algún centro académico y no en facultades)
	HORA_CLASE	Número de horas de trabajo a la semana, que el profesor destina al desarrollo de clases teóricas y prácticas correspondientes a las asignaturas asignadas, según el distributivo de trabajo, en el periodo académico.
	HORA_TUTORIA	Número de horas de trabajo a la semana, que el profesor destina a los estudiantes en la orientación y acompañamiento de tipo personal y académico, según el distributivo de trabajo, en el periodo académico.
HORA_OTRAS_DOCENCIA	Número de horas de trabajo a la semana, que el profesor destina a las otras actividades de docencia diferentes a las horas clase y horas de tutoría a estudiantes, según el distributivo de trabajo, en el periodo académico.	
HORAS_INVESTIGACION	Número de horas de trabajo a la semana, que el profesor destina al desarrollo de actividades de investigación, según el distributivo de trabajo, en el periodo académico.	

CONTINÚA



	HORAS_VINCULACION	Número de horas de trabajo a la semana, que el profesor destina a las actividades de vinculación, según el distributivo de trabajo, en el periodo académico.
	HORAS_GESTIÓN	Número de horas de trabajo a la semana, que el profesor destina al desarrollo de actividades de dirección o gestión académica, según el distributivo de trabajo, en el periodo académico.
Títulos	Identificacion	Número de documento de identificación del profesor. Se realizará la comprobación del dígito validador.
	Nombres	Nombres completos del profesor
	Nombre_titulo	Título del profesor
	Numero_registro	Número de registro del título otorgado por la Senescyt
	Tipo_Titulo	Procedencia del título obtenido: nacional o extranjera
	Fecha_Grado	Fecha en la cual el profesor obtuvo el título
	Fecha_Registro	Fecha en la cual el título se registró en la Senescyt
	IES	Universidad donde el profesor obtuvo el título
	Nivel	Nivel del título obtenido
	Tipo_Nivel	Tipo de nivel
	NombreClasificacion	Clasificación del título según la UNESCO
	Observacion	Información adicional sobre el título
	Area	Campo Amplio del título
	Codigo_area	Código del Campo Amplio
	Subarea	Campo Específico del título
	Codigo_subarea	Código del Campo Específico
Periodos	[strCodigo]	Código del periodo académico
	[strDescripcion]	Descripción del periodo académico
	[dtFechaInic]	Fecha de inicio del periodo académico
	[dtFechaFin]	Fecha fin del periodo académico
Carreras	[strCodigo]	Código de la carrera
	[strNombre]	Descripción de la carrera
Materias	[strCodigo]	Código de las asignaturas
	[strNombre]	Nombre de la asignatura
Dictado_Materias	[strCodParalelo]	Código del Paralelo
	[strCodPeriodo]	Código del Periodo Académico
	[strCodNivel]	Código del Nivel
	[strCodMateria]	Código de la Materia
	[strCodDocente]	Código del Docente
Facultades	[strCodigo]	Código de la Facultad
	[strNombre]	Nombre de la Facultad
Producción Científica – Artículos de relevancia mundial	CODIGO	Código que utiliza la ESPOCH para identificar el artículo
	NOMBRE_ARTICULO	Nombre o título del artículo
	DOI_ARTICULO	Ingresar el identificador único y permanente para las publicaciones electrónicas
	BASE_DATOS	Especificar Base Indexada: Ingresar la base indexada.
	NOMBRE_REVISTA	Nombre de la revista donde se publicó el artículo.
	ISSN	Código ISSN (International Standard Serial Number, Número Internacional Normalizado de Publicaciones Seriadadas).
	VOLUMEN	Volumen de la revista
	NÚMERO	Número de la revista
	PÁGINAS_ARTÍCULO	Número de páginas del artículo publicado
	FECHA_PUBLICACION	Fecha de publicación del artículo
	ÍNDICE_SJR	Índice de impacto calculado por Scimago Journal Rank
	PUNTUACIÓN	Puntuación en el baremo de tipo de artículo, según el Modelo de Evaluación de Medicina.
	CAMPO_AMPLIO	Ingrese el código del campo amplio del conocimiento a la que corresponde el artículo de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA

CONTINÚA



	NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
CAMPO_ESPECÍFICO	Ingrese el código del campo específico del conocimiento a la que corresponde el artículo de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
CAMPO_DETALLADO	Ingrese el código del campo detallado del conocimiento a la que corresponde el artículo de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
ESTADO	Seleccionar si el artículo ha sido publicado o si está aceptado para publicación en alguna revista científica
AUTOR_1	Ingrese los apellidos y nombres del autor 1 del artículo
IDENTIFICACIÓN_1	Número de documento de identificación del autor 1.
AUTOR_2	Ingrese los apellidos y nombres del autor 2 del artículo
IDENTIFICACIÓN_2	Número de documento de identificación del autor 2
AUTOR_3	Ingrese los apellidos y nombres del autor 3 del artículo
IDENTIFICACIÓN_3	Número de documento de identificación del autor 3
AUTOR_4	Ingrese los apellidos y nombres del autor 4 del artículo
IDENTIFICACIÓN_4	Número de documento de identificación del autor 4
AUTOR_5	Ingrese los apellidos y nombres del autor 5 del artículo
IDENTIFICACIÓN_5	Número de documento de identificación del autor 5
CARRERA_1	Carrera a la cual pertenecen los profesores autores del artículo
CARRERA_2	Carrera a la cual pertenecen los profesores autores del artículo
CARRERA_3	Carrera a la cual pertenecen los profesores autores del artículo
CARRERA_4	Carrera a la cual pertenecen los profesores autores del artículo
CARRERA_5	Carrera a la cual pertenecen los profesores autores del artículo
Producción Científica – Artículos regionales	CODIGO Código que utiliza la ESPOCH para identificar el artículo
	NOMBRE_ARTICULO Nombre o título del artículo
	DOI_ARTICULO Ingresar el identificador único y permanente para las publicaciones electrónicas
	NOMBRE_REVISTA Nombre de la revista donde se publicó el artículo.
	ISSN Código ISSN (International Standard Serial Number, Número Internacional Normalizado de Publicaciones Seriadas).
	VOLUMEN Volumen de la revista
	NÚMERO Número de la revista
	PÁGINAS_ARTÍCULO Número de páginas del artículo publicado
	FECHA_PUBLICACION Fecha de publicación del artículo
	PUNTUACIÓN Puntuación en el baremo de tipo de artículo, según el Modelo de Evaluación de Medicina.
	CAMPO_AMPLIO Ingrese el código del campo amplio del conocimiento a la que corresponde el artículo de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y

CONTINÚA



	GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
CAMPO_ESPECÍFICO	Ingrese el código del campo específico del conocimiento a la que corresponde el artículo de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
CAMPO_DETALLADO	Ingrese el código del campo detallado del conocimiento a la que corresponde el artículo de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
ESTADO	Seleccionar si el artículo ha sido publicado o si está aceptado para publicación en alguna revista científica
AUTOR_1	Ingrese los apellidos y nombres del autor 1 del artículo
IDENTIFICACIÓN_1	Número de documento de identificación del autor 1.
AUTOR_2	Ingrese los apellidos y nombres del autor 2 del artículo
IDENTIFICACIÓN_2	Número de documento de identificación del autor 2
AUTOR_3	Ingrese los apellidos y nombres del autor 3 del artículo
IDENTIFICACIÓN_3	Número de documento de identificación del autor 3
AUTOR_4	Ingrese los apellidos y nombres del autor 4 del artículo
IDENTIFICACIÓN_4	Número de documento de identificación del autor 4
AUTOR_5	Ingrese los apellidos y nombres del autor 5 del artículo
IDENTIFICACIÓN_5	Número de documento de identificación del autor 5
CARRERA_1	Carrera a la cual pertenecen los profesores autores del artículo
CARRERA_2	Carrera a la cual pertenecen los profesores autores del artículo
CARRERA_3	Carrera a la cual pertenecen los profesores autores del artículo
CARRERA_4	Carrera a la cual pertenecen los profesores autores del artículo
CARRERA_5	Carrera a la cual pertenecen los profesores autores del artículo
Producción Científica – Ponencias	CODIGO Código que utiliza la ESPOCH para identificar la ponencia
	NOMBRE_PONENCIA Nombre o título de la ponencia
	NOMBRE_CONGRESO Nombre del congreso o evento académico-científico
	PAIS_CONGRESO País donde se realizó el congreso o evento académico.
	CIUDAD_CONGRESO Ciudad en donde se realizó el congreso
	ISBN Número Estándar Internacional de Libros o Número Internacional Normalizado del Libro (de ser el caso)
	ISSN Ingresar el número internacional normalizado de publicaciones seriadas (de ser el caso)
	FECHA_PUBLICACION Fecha de publicación de la ponencia
	FECHA_INICIO Fecha de inicio del evento
	FECHA_FIN Fecha de finalización del evento
	RELEVANCIA Relevancia del evento, según el Modelo Genérico de Evaluación del Entorno de Aprendizaje de carreras.
	CAMPO_AMPLIO Ingrese el código del campo amplio del conocimiento a la que corresponde la ponencia de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS



	INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
CAMPO_ESPECÍFICO	Ingrese el código del campo específico del conocimiento a la que corresponde la ponencia de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
CAMPO_DETALLADO	Ingrese el código del campo detallado del conocimiento a la que corresponde la ponencia de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
AUTOR_1	Ingrese los apellidos y nombres del autor 1 de la ponencia
IDENTIFICACIÓN_1	Número de documento de identificación del autor 1.
AUTOR_2	Ingrese los apellidos y nombres del autor 2 de la ponencia
IDENTIFICACIÓN_2	Número de documento de identificación del autor 2
AUTOR_3	Ingrese los apellidos y nombres del autor 3 de la ponencia
IDENTIFICACIÓN_3	Número de documento de identificación del autor 3
AUTOR_4	Ingrese los apellidos y nombres del autor 4 de la ponencia
IDENTIFICACIÓN_4	Número de documento de identificación del autor 4
AUTOR_5	Ingrese los apellidos y nombres del autor 5 de la ponencia
IDENTIFICACIÓN_5	Número de documento de identificación del autor 5
CARRERA_1	Carrera a la cual pertenecen los profesores autores de la ponencia
CARRERA_2	Carrera a la cual pertenecen los profesores autores de la ponencia
CARRERA_3	Carrera a la cual pertenecen los profesores autores de la ponencia
CARRERA_4	Carrera a la cual pertenecen los profesores autores de la ponencia
CARRERA_5	Carrera a la cual pertenecen los profesores autores de la ponencia
Producción Científica – Libros	
CODIGO	Código que utiliza la ESPOCH para identificar el libro
NOMBRE_LIBRO	Nombre o título del libro
NOMBRE_CAPÍTULO	Nombre o título del capítulo del libro (de ser el caso)
ISBN	Número Estándar Internacional de Libros o Número Internacional Normalizado del Libro
FECHA_PUBLICACION	Fecha de publicación del libro
REVISION_PARES	Especificar si el libro o capítulo del libro ha sido revisado por pares (dos pares académicos)
CAMPO_AMPLIO	Ingrese el código del campo amplio del conocimiento a la que corresponde el libro de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
CAMPO_ESPECÍFICO	Ingrese el código del campo específico del conocimiento a la que corresponde el libro de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)



CAMPO_DETALLADO	Ingrese el código del campo detallado del conocimiento a la que corresponde el libro de acuerdo a su contenido (según la codificación de los campos de conocimiento del REGLAMENTO DE ARMONIZACIÓN DE LA NOMENCLATURA DE TÍTULOS PROFESIONALES Y GRADOS ACADÉMICOS QUE CONFIEREN LAS INSTITUCIONES DE EDUCACIÓN SUPERIOR DEL ECUADOR)
AUTOR_1	Ingrese los apellidos y nombres del autor 1 del libro
IDENTIFICACIÓN_1	Número de documento de identificación del autor 1.
AUTOR_2	Ingrese los apellidos y nombres del autor 2 del libro
IDENTIFICACIÓN_2	Número de documento de identificación del autor 2
AUTOR_3	Ingrese los apellidos y nombres del autor 3 del libro
IDENTIFICACIÓN_3	Número de documento de identificación del autor 3
AUTOR_4	Ingrese los apellidos y nombres del autor 4 del libro
IDENTIFICACIÓN_4	Número de documento de identificación del autor 4
AUTOR_5	Ingrese los apellidos y nombres del autor 5 del libro
IDENTIFICACIÓN_5	Número de documento de identificación del autor 5
CARRERA_1	Carrera a la cual pertenecen los profesores autores del libro
CARRERA_2	Carrera a la cual pertenecen los profesores autores del libro
CARRERA_3	Carrera a la cual pertenecen los profesores autores del libro
CARRERA_4	Carrera a la cual pertenecen los profesores autores del libro
CARRERA_5	Carrera a la cual pertenecen los profesores autores del libro

4.4 Diseño

MODELO DE GESTIÓN DE INDICADORES

El modelo de gestión de indicadores posee la siguiente estructura:



Figura 20. Modelo de Gestión de Indicadores de Evaluación y de Gestión

ANÁLISIS DE INDICADORES DE EVALUACIÓN Y DE GESTIÓN

En esta fase se realizó un análisis de los criterios e indicadores cuantitativos de evaluación y de gestión, conforme los modelos de evaluación institucional y de carreras, propuestos por los organismos que regulan la educación superior en el Ecuador.

ANÁLISIS DE LAS FUENTES DE DATOS

En función de los indicadores definidos en la fase anterior se evaluó las diferentes fuentes de datos que posee la IES y que permitirán la generación de los indicadores cuantitativos de evaluación y otros indicadores de gestión.

DISEÑO DEL ALMACÉN DE DATOS

Una vez analizadas las fuentes de datos se realizará el diseño del almacén de datos, para el cual se deberá considerar los componentes de Vinculación, Investigación, Academia y de Gestión, es recomendable se diseñe un Datamart por cada componente como mínimo.

IMPLEMENTACIÓN DEL DATA WAREHOUSING

En esta etapa del modelo se utilizará alguna metodología para el desarrollo del Data Warehouse de Evaluación, en el caso práctico se utilizó la metodología Kimball.

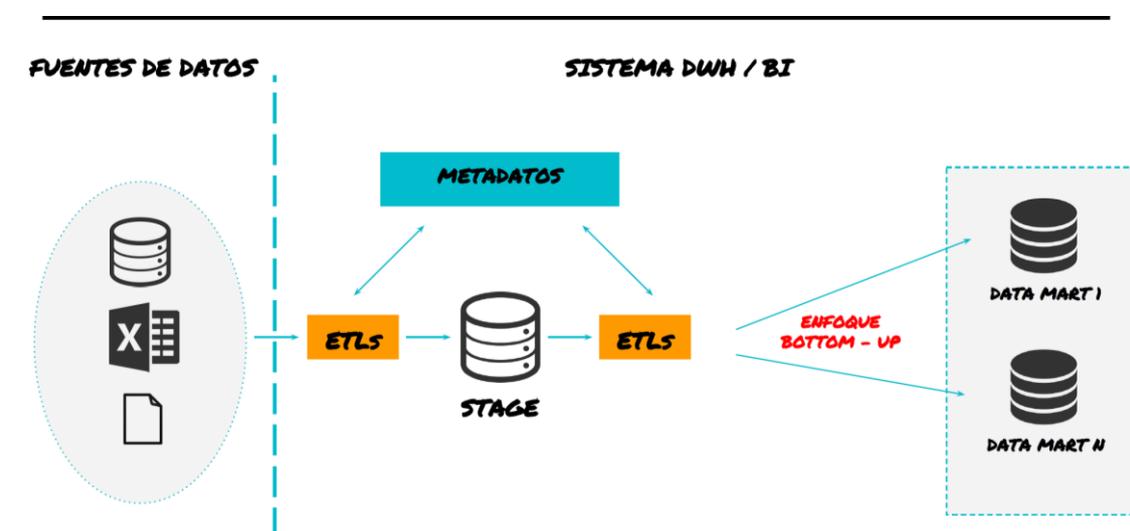


Figura 21. Esquema de la arquitectura del sistema propuesto

Una vez implementado el Data Warehouse se aplicarán los algoritmos de Minería de Datos, en este caso se desarrollará el clustering.

VALIDACIÓN DEL MODELO DE GESTIÓN

Una vez desarrollada la minería de datos, se validará el resultado del clustering con sus propios métodos de validación. Adicionalmente se aplicó una prueba de usabilidad a los usuarios del sistema de información desarrollado.

Paso 8: Diseño del Data Warehouse

La arquitectura del sistema desarrollado para el presente trabajo de investigación, se basa en el paradigma de Ralph Kimball, en el que se generará una base de datos STAGE, la misma que servirá como un área de almacenamiento temporal para hacer una réplica de los datos requeridos (NO Intrusivo) la cual no afecta el rendimiento de las bases de datos fuente y permite indexar y trabajar en un subconjunto de datos con la finalidad de optimizar la solución (Murillo, 2019).

Luego de desarrollar los procesos ETL se creará un repositorio de datos corporativo (Data Warehouse), el mismo que se caracteriza por integrar y depurar información de una o más fuentes de datos, para luego procesarlas en una infinidad de perspectivas que permiten el análisis de los datos y la toma de decisiones (Sinnexus, 2019). A continuación, se dispone del siguiente esquema, bajo el cual se desarrollará el sistema propuesto:

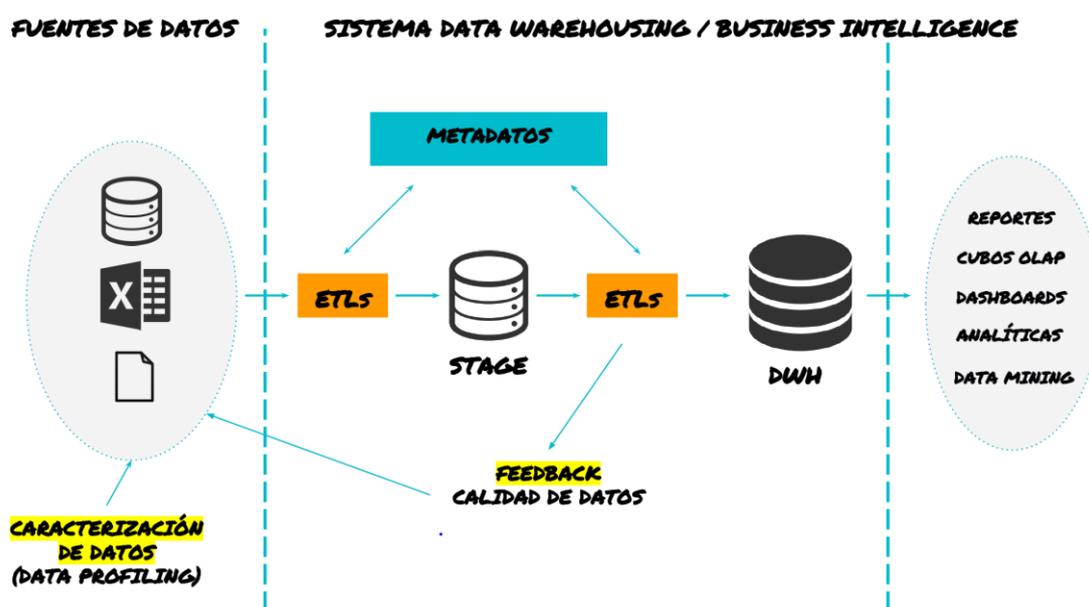


Figura 22. Arquitectura del Modelo de Gestión

Los elementos del Data Warehouse son:

Hechos: Permiten representar los procesos de negocio de la organización.

Dimensiones: Permiten representar las distintas vistas para un cierto proceso de negocio.

Métricas: Son aquellos conceptos cuantificables que permiten medir un proceso de negocio y están asociados a las tablas de hecho.

Para estructurar los datos en un almacén de datos se representará mediante un esquema en estrella; en el cual la tabla de hechos se relaciona con las dimensiones y representan a entidades de la base de datos fuente, esto se puede observar en el siguiente esquema:

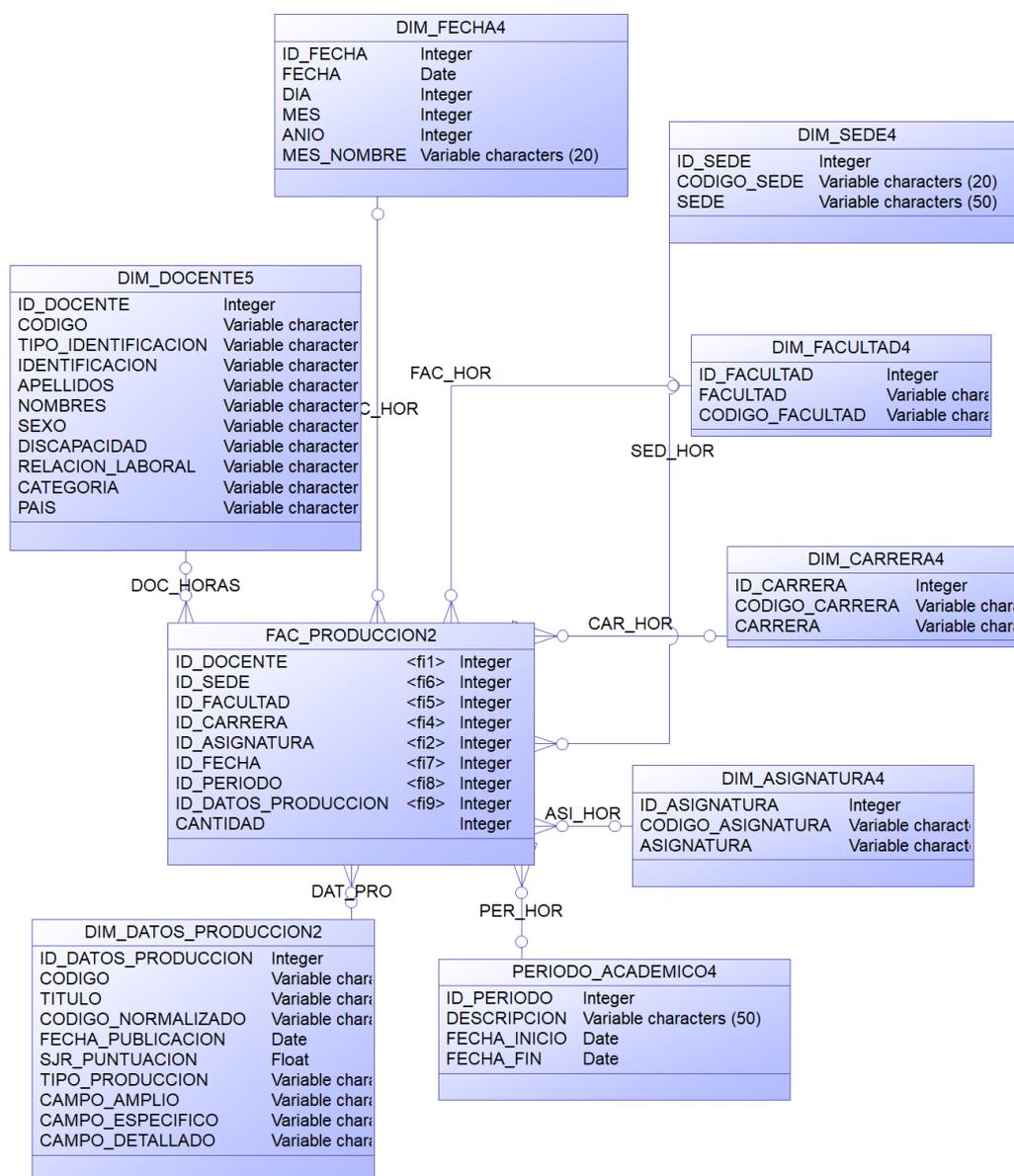


Figura 23. Diseño lógico del Data Warehouse

Paso 9: Diseño del ETL

Los procesos ETL son las tecnologías de integración de datos en el Data Warehouse basados en la definición de fuentes de datos, procesos de extracción, transformación y adecuación a la estructura del repositorio de datos destino.

Un ETL es un proceso de Extracción, Transformación y Carga de Datos, la Extracción consiste en la obtención de información de las distintas fuentes tanto internas como externas, la Transformación se encarga del filtrado, limpieza, depuración, homogeneización y agrupación de la información. Finalmente, la carga es la organización y actualización de los datos y los metadatos en la base de datos (Sinnexus, 2019).

La plataforma que se utilizó para la implementación de los ETLs es Pentaho Versión 8.1. A continuación, se tiene una muestra de los diseños implementados para la carga del Data Warehouse:

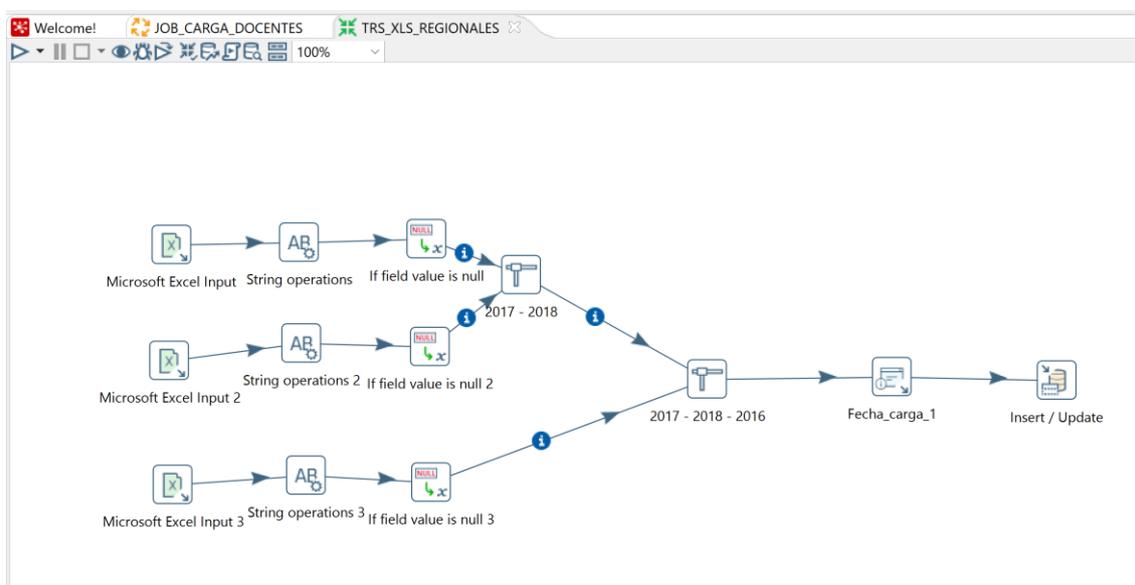


Figura 24. Esquema para la extracción de datos de la producción científica de tipo regional

El job diseñado para la carga de datos, en este caso de los docentes de la Facultad de Administración de Empresas es el siguiente:



Figura 25: Job de carga de los datos de docentes

Paso 10: Diseño del Repositorio de Metadatos

Para el desarrollo del presente proyecto se implementó un repositorio de metadatos al que se lo ha llamado ETLs_EVALUACION, el cual se ubicó en el siguiente directorio físico: C:\Tesis_BI\etls, en el repositorio se almacenarán los ETLs desarrollados y la conexión a las diferentes bases de datos y archivos que posee el sistema académico y el archivo de publicaciones y talento humano, este caso información que corresponde a la Facultad de Administración de Empresas de la ESPOCH.

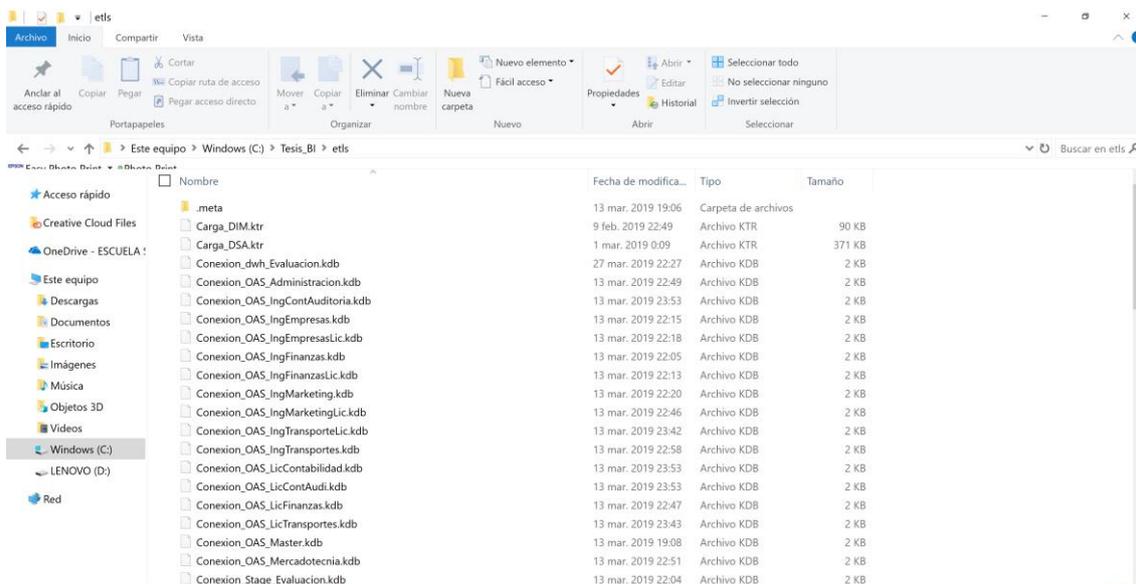


Figura 26. Repositorio físico de metadatos

4.5 Construcción

Paso 11: Desarrollo del ETL

En armonía con las necesidades del negocio, se creó un repositorio de datos de tipo stage al que le hemos llamado `dsa_evaluación`, en el cual se almacenaron los datos de las fuentes de información de publicaciones y del sistema académico, cabe mencionar que previamente se realizó una limpieza y depuración.

The screenshot shows the phpMyAdmin interface for a database named 'dsa_evaluacion'. The left sidebar displays a tree view of the database structure, including tables like 'dsa_articulos', 'dsa_campos_conocimiento', 'dsa_carreras', 'dsa_categoria', 'dsa_dedicacion', 'dsa_dictado_materias', 'dsa_docentes', 'dsa_docentes_16_17', 'dsa_docentes_art_cientificos', 'dsa_docentes_art_regionales', 'dsa_docentes_congresos', 'dsa_docentes_investigacion', 'dsa_docentes_libros', 'dsa_docentes_periodo', 'dsa_escuelas', 'dsa_facultades', 'dsa_investigacion', 'dsa_libros', 'dsa_paises', 'dsa_periodos', and 'dsa_ponencias'. The main area shows a table structure view with columns: Tabla, Acción, Filas, Tipo, Cotejamiento, Tamaño, and Residuo a depu. The table list includes the following data:

Tabla	Acción	Filas	Tipo	Cotejamiento	Tamaño	Residuo a depu
dsa_articulos	Examinar	56	InnoDB	latin1_swedish_ci	64 KB	
dsa_campos_conocimiento	Examinar	95	InnoDB	latin1_swedish_ci	16 KB	
dsa_carreras	Examinar	113	InnoDB	latin1_swedish_ci	32 KB	
dsa_categoria	Examinar	6	InnoDB	latin1_swedish_ci	16 KB	
dsa_dedicacion	Examinar	3	InnoDB	latin1_swedish_ci	16 KB	
dsa_dictado_materias	Examinar	24,070	InnoDB	latin1_swedish_ci	3.5 MB	
dsa_docentes	Examinar	1,393	InnoDB	latin1_swedish_ci	416 KB	
dsa_docentes_16_17	Examinar	1,023	InnoDB	latin1_swedish_ci	432 KB	
dsa_docentes_art_cientificos	Examinar	163	InnoDB	latin1_swedish_ci	48 KB	
dsa_docentes_art_regionales	Examinar	551	InnoDB	latin1_swedish_ci	112 KB	
dsa_docentes_congresos	Examinar	881	InnoDB	latin1_swedish_ci	128 KB	
dsa_docentes_investigacion	Examinar	641	InnoDB	latin1_swedish_ci	128 KB	
dsa_docentes_libros	Examinar	58	InnoDB	latin1_swedish_ci	16 KB	
dsa_docentes_periodo	Examinar	590	InnoDB	latin1_swedish_ci	112 KB	
dsa_escuelas	Examinar	52	InnoDB	latin1_swedish_ci	32 KB	
dsa_facultades	Examinar	10	InnoDB	latin1_swedish_ci	16 KB	
dsa_investigacion	Examinar	994	InnoDB	latin1_swedish_ci	368 KB	
dsa_libros	Examinar	30	InnoDB	latin1_swedish_ci	32 KB	
dsa_paises	Examinar					
dsa_periodos	Examinar					
dsa_ponencias	Examinar					

Figura 27. Base de Datos Stage `dsa_evaluacion`

En el diseño de los ETLs, se emplearon componentes para la depuración de los datos provenientes de los registros de las ponencias desarrolladas por el personal académico.

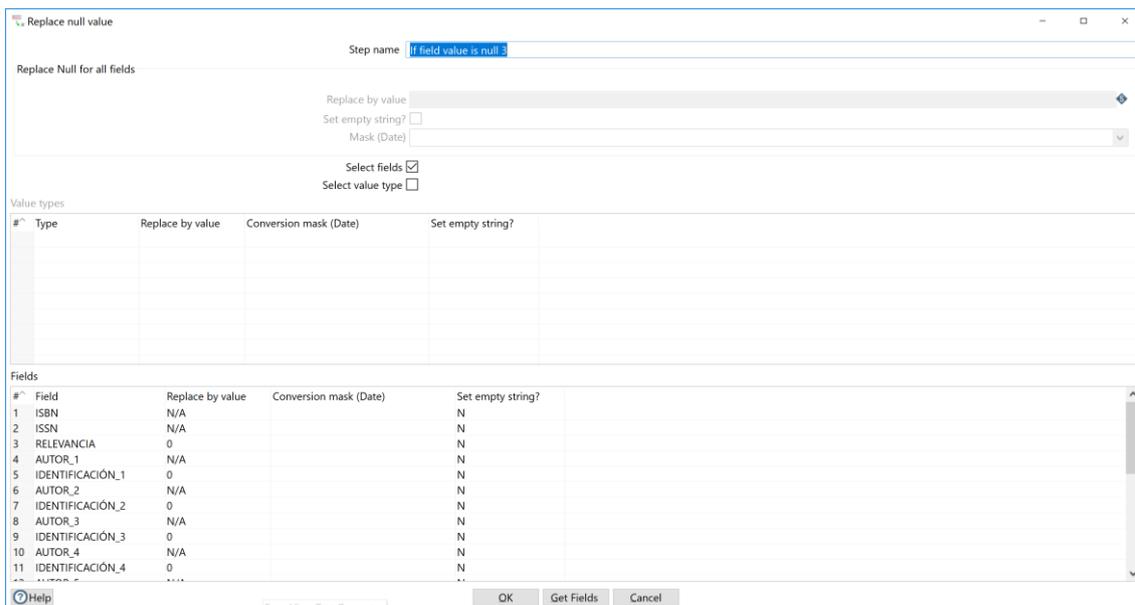


Figura 28. Componente de Pentaho utilizado para la depuración de datos

Luego se cargan los datos en el repositorio del stage, se implementó la base de datos del Data Warehouse a la que se le llamó `dwh_evaluacion`; la misma posee las tablas que corresponden a las dimensiones y hechos. Esto lo puede visualizar en la siguiente figura:

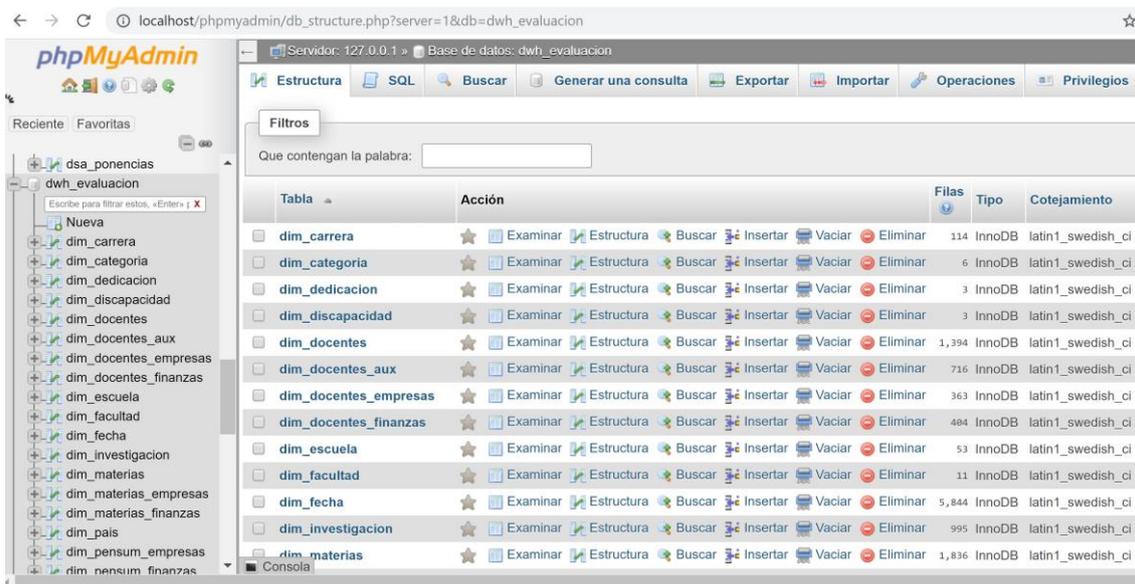


Figura 29. Data Warehouse `dwh_evaluacion`

De la misma manera, se diseñaron ETLs para realizar la carga de datos en el Data Warehouse

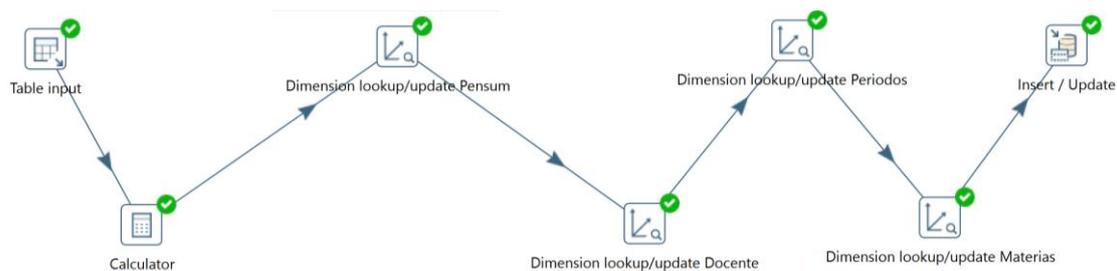


Figura 30. ETL para cargar la fact_rendimiento_finanzas

Paso 12: Desarrollo Aplicación

En esta etapa de desarrollo de la aplicación se utilizó la plataforma de TABLEAU, con la cual se configuró la conexión a la base de datos del Data Warehouse:

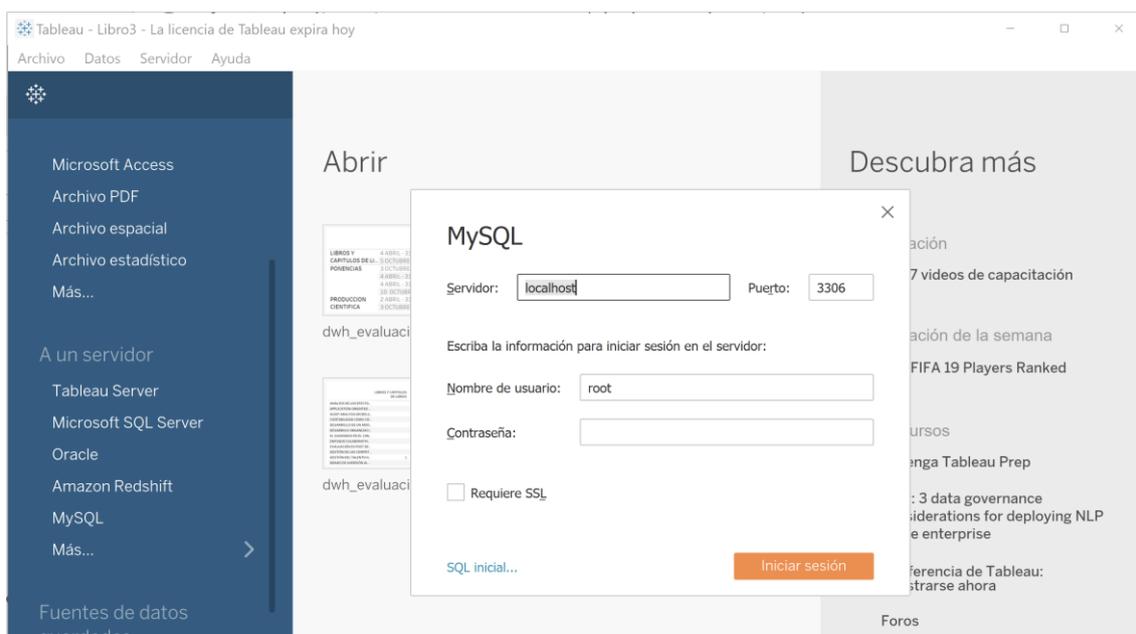


Figura 31. Configuración del Tableau para acceder al Data Warehouse

Paso 13: Minería de Datos

Para desarrollar Minería de Datos se utilizará el lenguaje R, debido a que contiene todas las librerías y funciones necesarias para el clustering y su validación.

1. Preparación de los datos

En esta fase cargamos los datos, en este caso de las evaluaciones de los estudiantes del primer, segundo y tercer parcial por periodo académico, además de cargar las librerías necesarias para ejecutar el clustering

```
## 1. Preparación de los datos
setwd("C:/R")
dir()
evaluacion<-read.csv("Datos_Rapidminer_Completos.csv", header =TRUE, dec=".", sep = ",")
str(evaluacion)
summary(evaluacion)
evaluacion.scale <-as.data.frame(scale(evaluacion[,1:3]))
## ahora cargamos las librerias necesarias
library(cluster)
library(factoextra)
```

2. Obtención de las matrices de distancias

A continuación, se obtendrán las matrices de distancia utilizando el método Pearson

```
res.dist <- get_dist(evaluacion.scale, stand = TRUE, method = "pearson")
memory.size (max = TRUE)
memory.limit(size=8000000000)
fviz_dist(res.dist, gradient = list(low = "#00AFBB", mid = "white", high = "#FC4E07"))
```

3. Determinación óptima del número de cluster

```
## Determinacion óptima del número de cluster
fviz_nbclust(evaluacion.scale, kmeans, method = "gap_stat")
```

En este apartado se utiliza el método `fviz_nbclust()` para obtener el número de cluster óptimo; como resultado se tiene la siguiente figura:

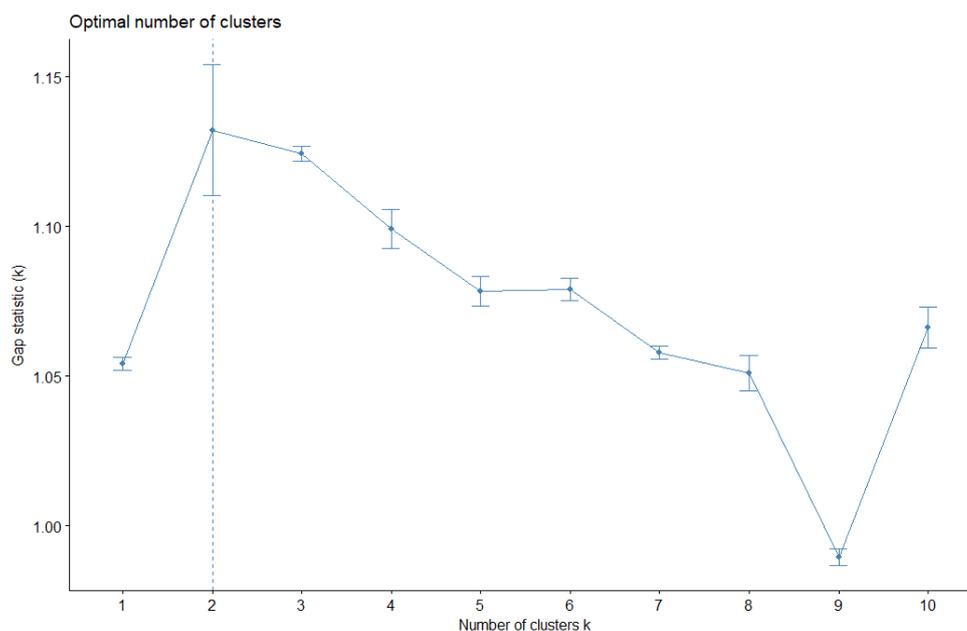


Figura 32. Número de clusters óptimo

Como se puede observar en la figura, el cluster óptimo es el 2.

4. Cluster kmeans

```
## Cluster k-means
km.res <- kmeans(evaluacion.scale, 3, nstart = 25)
fviz_cluster(km.res, data = evaluacion.scale, frame.type = "convex")
```

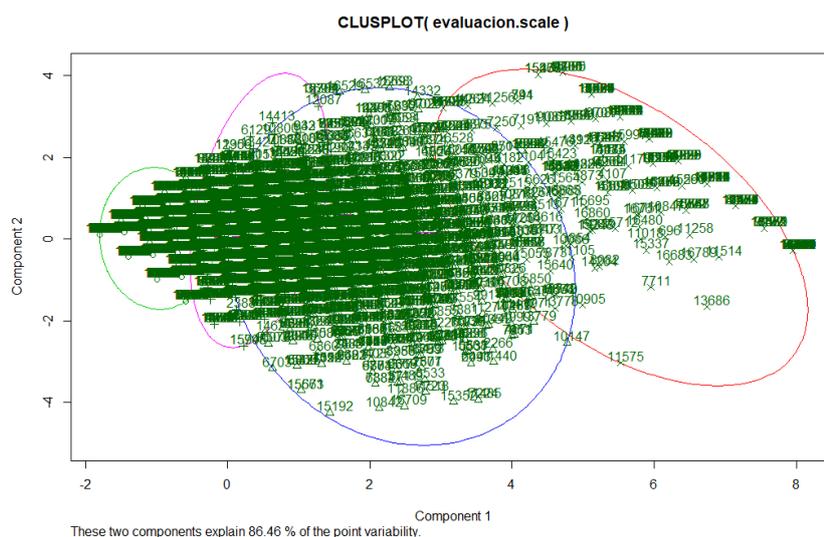


Figura 33. Clusters generados por el algoritmo k-means

5. Cluster PAM

```
## Cluster PAM
library("cluster")
pam.res <- pam(evaluacion.scale, 4)
# Visualize
fviz_cluster(pam.res)
```

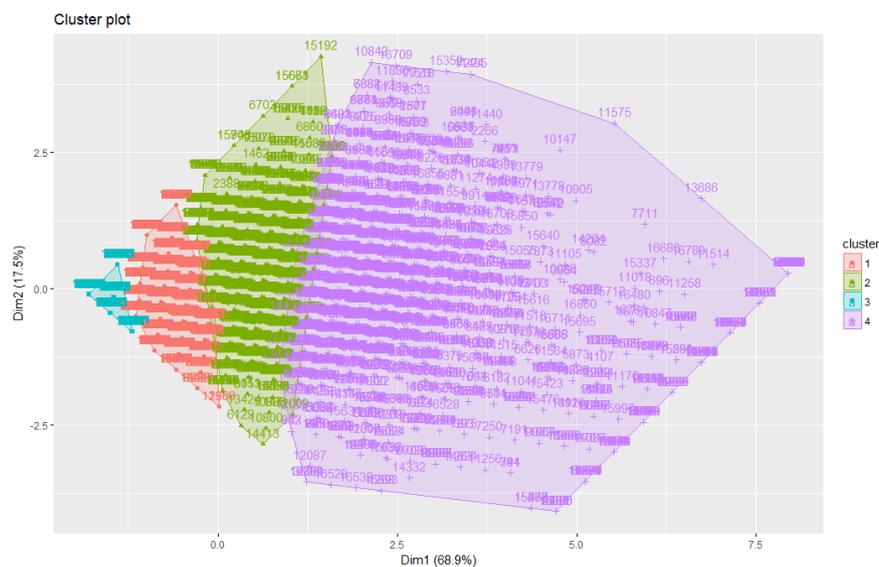


Figura 34. Cluster generado aplicando el algoritmo PAM

6. Cluster jerárquico

```
## Cluster Jerarquico
clarax <- clara(evaluacion.scale, 3)
# Cluster plot
fviz_cluster(clarax, stand = T, geom = "point",
              pointsize = 1)
plot(silhouette(clarax), col = 2:3, main = "Silhouette plot")
```

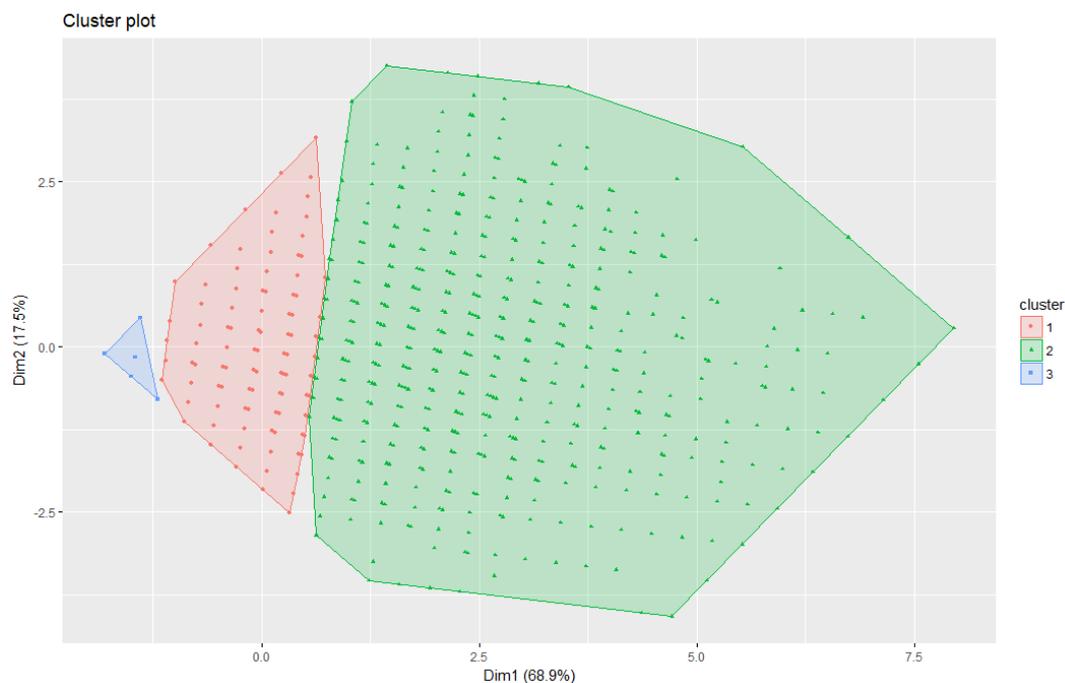


Figura 35. Cluster Generado por el algoritmo jerárquico Clara

7. Validación de los cluster

Para la validación del clustering se realiza 3 actividades: el estudio de la tendencia de *clustering*, la elección del número óptimo de *clusters* y el estudio de la calidad/significancia de los *clusters* generados. El código de R implementado se muestra a continuación:

```
##Generación del Dendograma
evaluacion<-read.csv("Datos_Rapidminer_Completos.csv", header =TRUE, dec="," , sep = "," )
evaluacion.scale <-as.data.frame(scale(evaluacion[,1:3]))
# 2. Compute dissimilarity matrix
d <- dist(evaluacion.scale, method = "euclidean")
# Hierarchical clustering using ward's method
res.hc <- hclust(d, method = "ward.D2" )
# Cut tree into 4 groups
grp <- cutree(res.hc, k = 4)
# Visualize
plot(res.hc, cex = 0.6) # plot tree
rect.hclust(res.hc, k = 4, border = 2:5) # add rectangle
library("factoextra")
```

```
# Compute hierarchical clustering and cut into 4 clusters
res <- hcut(evaluacion, k = 4, stand = TRUE)
# Visualize
fviz_dend(res, rect = TRUE, cex = 0.5,
          k_colors = c("#00AFBB", "#2E9FDF", "#E7B800", "#FC4E07"))
```

8. Evaluación de la tendencia de agrupamiento

```
## Evaluacion de la tendencia de agrupamiento
get_clust_tendency(evaluacion.scale, n = 50,
                  gradient = list(low = "steelblue", high = "white"))
## Determinar número de agrupaciones
evaluacion.scale <- as.data.frame(scale(evaluacion[,1:3]))
library("factoextra")
fviz_nbclust(evaluacion.scale, kmeans, method = "gap_stat")
library("NbClust")
set.seed(123)
res.nbclust <- NbClust(evaluacion.scale, distance = "euclidean",
                    min.nc = 2, max.nc = 10,
                    method = "complete", index = "all")
factoextra::fviz_nbclust(res.nbclust)
## Estadísticas de validación
evaluacion.scale <- as.data.frame(scale(evaluacion[,1:3]))
# Enhanced hierarchical clustering, cut in 3 groups
library("factoextra")
res.hc <- eclust(evaluacion.scale, "hclust", k = 3, graph = FALSE)
# Visualize
fviz_dend(res.hc, rect = TRUE, show_labels = FALSE)

fviz_silhouette(res.hc)
sil <- res.hc$silinfo$widths[, 1:3]
# Objects with negative silhouette
neg_sil_index <- which(sil[, 'sil_width'] < 0)
sil[neg_sil_index, , drop = FALSE]
```

9. Determinación del número óptimo de agrupaciones

```
## Evaluacion de la tendencia de agrupamiento
get_clust_tendency(evaluacion.scale, n = 50,
                  gradient = list(low = "steelblue", high = "white"))
```

10. Estadísticas de validación de clústeres

```
## Estadísticas de validación
evaluacion.scale <- as.data.frame(scale(evaluacion[,1:3]))
# Enhanced hierarchical clustering, cut in 3 groups
library("factoextra")
res.hc <- eclust(evaluacion.scale, "hclust", k = 3, graph = FALSE)
# Visualize
fviz_dend(res.hc, rect = TRUE, show_labels = FALSE)

fviz_silhouette(res.hc)
sil <- res.hc$silinfo$widths[, 1:3]
# Objects with negative silhouette
neg_sil_index <- which(sil[, 'sil_width'] < 0)
sil[neg_sil_index, , drop = FALSE]

library("factoextra")
res.hc <- eclust(evaluacion.scale, "hclust", k = 3, graph = FALSE)
# Visualize
fviz_dend(res.hc, rect = TRUE, show_labels = FALSE)

fviz_silhouette(res.hc)
sil <- res.hc$silinfo$widths[, 1:3]
# Objects with negative silhouette
neg_sil_index <- which(sil[, 'sil_width'] < 0)
sil[neg_sil_index, , drop = FALSE]
```

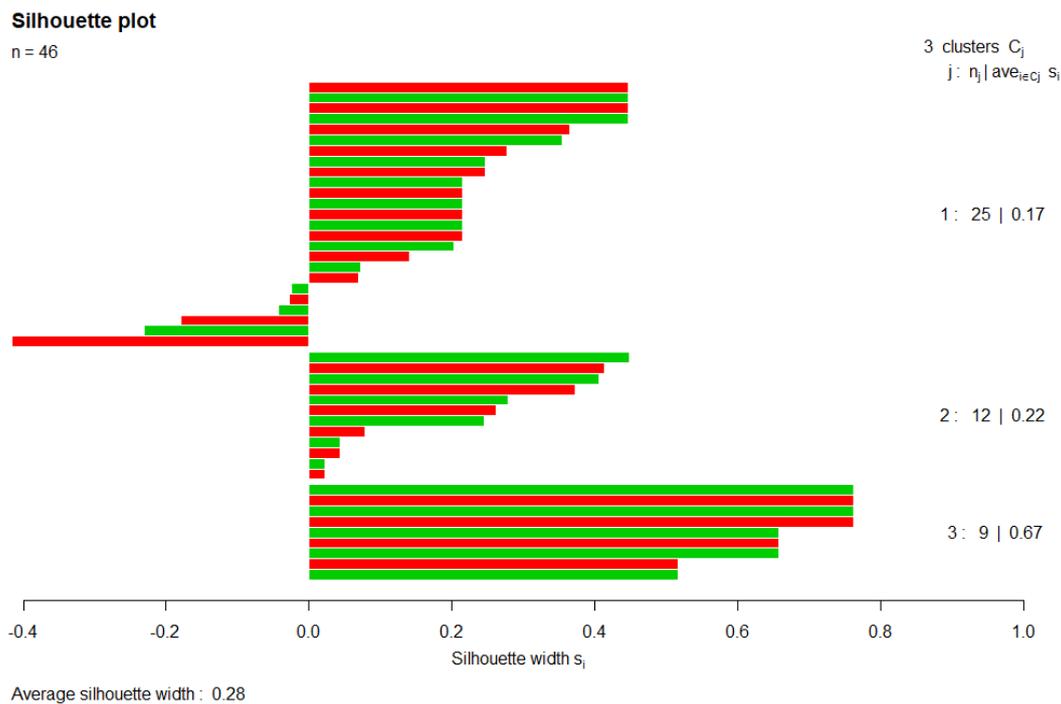


Figura 36. Análisis del índice de silhouette aplicado al cluster jerárquico

En la figura anterior, una vez aplicado el algoritmo jerárquico e-clust() se puede observar como el tercer cluster, de acuerdo al índice de silhouette se acerca más a 1, esto quiere decir que a la tercera iteración se tiene el modelo más óptimo.

11. Selección del algoritmo más apropiado

Si deseamos conocer el algoritmo de Clustering, más apropiado podemos ejecutar el siguiente script:

```
## Selección del Mejor algoritmo
evaluacion.scale <- as.data.frame(scale(evaluacion[,1:3]))
# Compute clValid
library("clValid")
intern <- clValid(evaluacion.scale, nClust = 2:6,
                 clMethods = c("hierarchical", "kmeans", "pam", "clara"),
                 validation = "internal")

# Summary
summary(intern)
plot(intern)
```

4.6 Desarrollo

Paso 15: Implementación

Para la implementación del sistema se utilizó Tableau y en este se configuraron enlaces con las diferentes tablas de hechos que se implementaron, con la finalidad de obtener los reportes requeridos por el sistema, por ejemplo en la siguiente figura se puede observar la configuración de la fact_produccion con las dimensiones de docentes, investigacion y periodos académicos.

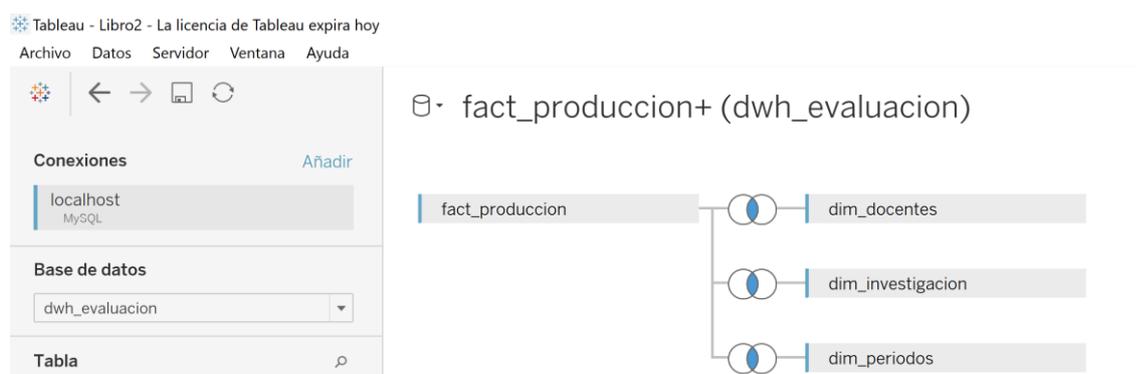


Figura 37. Configuración que vincula la tabla de hechos con las dimensiones

A continuación, se tiene una muestra de los reportes generados en el Tableau

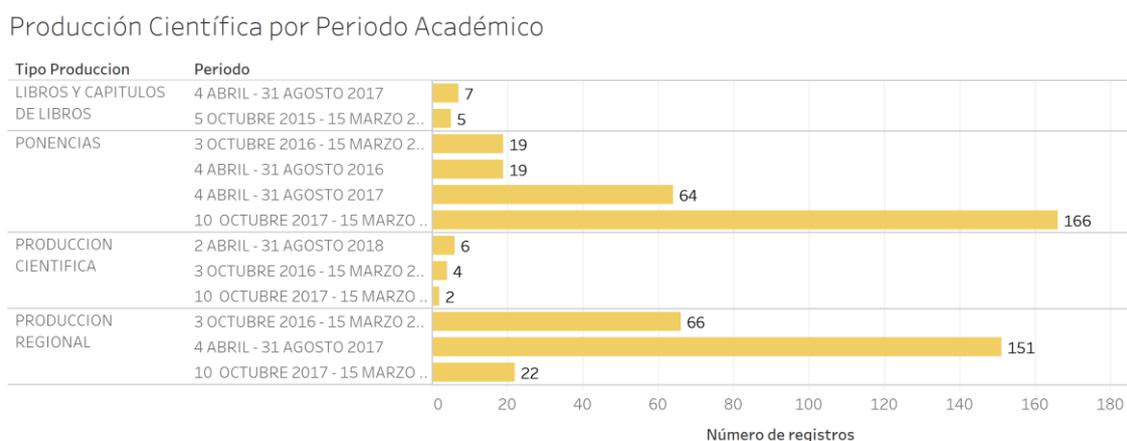


Figura 38. Reporte de la producción científica por periodo académico

Producción científica por Campo Amplio

Campo Amplio	Tipo Produccion			
	LIBROS Y CAPITULOS DE LIBROS	PONENCIAS	PRODUCCION CIENTIFICA	PRODUCCION REGIONAL
ADMINISTRACIÓN	10	216		106
AGRICULTURA, SILVICULT..		1		11
ARTES Y HUMANIDADES		1		8
CIENCIAS NATURALES, M..		6		9
CIENCIAS SOCIALES, PERI..		5		73
EDUCACIÓN	1	10	1	11
INGENIERÍA, INDUSTRIA ..	1	1	6	
SALUD Y BIENESTAR		3	1	2
SERVICIOS		10	4	
TECNOLOGÍAS DE LA INFO..		15		19
Total general	12	268	12	239

Figura 39. Producción científica por campo amplio

Producción Científica por Dedicación

Tipo Produccion	Dedicacion		
	MEDIO TIEMPO	TIEMPO COMPL..	TIEMPO PARCIAL
LIBROS Y CAPITULOS DE L..		6	1
PONENCIAS	7	83	19
PRODUCCION CIENTIFICA		7	
PRODUCCION REGIONAL	10	61	15

Figura 40. Producción científica por dedicación del docente

Paso 16: Evaluación de la Solución

Para evaluar la solución, se realizaron pruebas aplicando diferentes consultas SQL a la base de datos fuente, al stage y al Data Warehouse, para ello se utilizó ACCESS 2016, cotejando el número de registro almacenados en la base de datos, en el stage y en el Data Warehouse.

The screenshot shows a Microsoft Access window with a query named 'Consulta1_DEDICACION_FINANZAS'. The query result is displayed in a table with the following columns and data:

dsa_facultad	dsa_escuelas.strCodigo	dsa_carreras.strCodigo	dsa_periodo	dsa_docente	dsa_mi
FADE	EIFCE	IFIN	P0024	39746	207AB
FADE	EIFCE	IFIN	P0027	39746	305AB
FADE	EIFCE	IFIN	P0026	39746	207AB
FADE	EIFCE	IFIN	P0014	39746	IF0107
FADE	EIFCE	IFIN	P0026	39746	106AB
FADE	EIFCE	IFIN	P0013	39746	IF0305
FADE	EIFCE	IFIN	P0028	39746	305AB
FADE	EIFCE	IFIN	P0026	39746	305AB
FADE	EIFCE	IFIN	P0027	39746	207AB
FADE	EIFCE	IFIN	P0012	39746	IF0207
FADE	EIFCE	IFIN	P0025	39746	106AB
FADE	EIFCE	IFIN	P0013	39746	IF0207
FADE	EIFCE	IFIN	P0025	39746	106AB

Figura 41. Evaluación de los resultados mediante sentencias SQL

Se realizó una inducción a los docentes responsables de la comisión de evaluación a quienes les interesa los indicadores generados. Finalmente se aplicó una encuesta basándonos en el modelo de evaluación de la usabilidad de las plataformas de Inteligencia de Negocios propuesta por (Poropat, 2014). La encuesta aborda atributos como: facilidad de aprendizaje del sistema, la eficiencia del sistema, la flexibilidad del sistema, la participación de los usuarios del sistema y la usabilidad del sistema.



Figura 42. Dimensiones para evaluar la usabilidad del sistema propuesto

CAPÍTULO V

RESULTADOS

Una vez realizada la capacitación formal a los docentes responsables de la comisión de evaluación y aseguramiento de la calidad, de las diferentes carreras de la Facultad de Administración y de la ESPOCH, se aplicó la encuesta para evaluar la usabilidad del sistema de gestión de indicadores de evaluación, a 20 docentes.

A continuación, se tiene los resultados de la encuesta.

Facilidad de Aprendizaje del Sistema

1. En la pregunta relacionada sobre si el sistema al ser utilizado cubre la mayoría de las respuestas a las preguntas que aparecen, se tiene los siguientes resultados:

20 respuestas

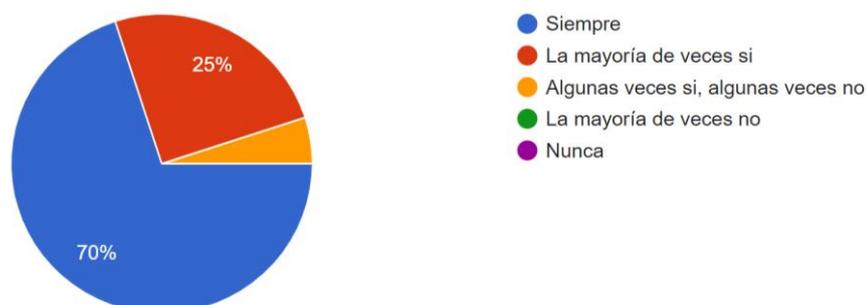


Figura 43. Respuesta que el sistema presenta frente a preguntas del usuario

Esto demuestra que, del total de encuestados, el 70% respondió que el sistema siempre responde a las preguntas que aparecen al momento de ser utilizado, un 25% responde a que el sistema responde la mayoría de veces a las preguntas planteadas por el usuario y un 5% seleccionó, que el sistema algunas veces si responde y algunas veces no lo hace.

2. La capacitación dada sobre el uso del sistema permite la gestión rápida y eficaz de la información, así como la ayuda en el aprendizaje de sus características

20 respuestas



Figura 44. Resultado de la gestión rápida y eficaz del sistema

Esto determina que, del total de encuestados, el 75% respondió que la capacitación permite la gestión rápida y eficaz de la información, mientras que el 10% mencionó que la mayoría de las veces si y el 15% algunas veces si y otras veces no.

Considerando la respuesta de los encuestados en relación a la facilidad de aprendizaje del sistema, se determinó que más del 70% de los encuestados, están de acuerdo con el mismo.

Eficiencia del Sistema

3. El sistema apoya plenamente en el desarrollo de las actividades que usted realiza

20 respuestas



Figura 45. Apoya el sistema al desarrollo de las actividades del usuario

Del total de encuestados, se tiene que el 65% respondió a que el sistema siempre apoya al desarrollo de las actividades que realiza, mientras que el 35% restante respondió que el sistema apoya la mayoría de las veces.

4. A su criterio, el sistema contribuye a generar los indicadores de evaluación necesarios para el proceso en el cual usted se desempeña, por ejemplo: se obtiene el número de publicaciones por docente, por periodo académico

20 respuestas



Figura 46. Contribuye el sistema a generar los indicadores de evaluación

En este apartado, se tiene que el 80% de los encuestados, ratifican que el sistema contribuye al proceso en el cual se desempeñan los miembros de la comisión de evaluación, mientras que el 20% mencionó que la mayoría de las veces si contribuye a generar indicadores de evaluación. En las preguntas planteadas para evaluar la eficiencia del sistema, se tiene que el sistema apoya y contribuye a la generación de indicadores de evaluación, necesarios para los procesos de autoevaluación de las carreras de la ESPOCH

Flexibilidad del Sistema

4. El sistema tiene funcionalidades integradas que pueden ser configuradas de acuerdo a sus necesidades.

Por ejemplo: Permite generar reportes en otros formatos, Permite cambiar los cubos de información, Permite generar gráficas que resumen la información, entre otros

20 respuestas



Figura 47. Se configura el sistema de acuerdo con las necesidades del usuario

En relación a esta pregunta, se tiene que en el 80% de los encuestados se encuentran de acuerdo con las funcionalidades que posee el sistema, mientras que el 15% respondió la mayoría de veces si, mientras que el 5% respondió que algunas veces si, algunas veces no.

5. Usted tiene acceso al sistema en cualquier lugar donde precise de información

20 respuestas



Figura 48. Acceso al sistema en cualquier lugar

Del total de encuestados, se tiene que el 60% mencionó que tiene siempre acceso al sistema, mientras que el 20% respondió que la mayoría de las veces si, un 15% mencionó que algunas veces si, algunas veces no y un 5% mencionó que la mayoría de las veces no. En relación al 5% que mencionó que la mayoría de las veces no tiene acceso al sistema, se puede decir que esto es temporal, en vista de que al momento se tiene un prototipo del sistema.

Para sintetizar los resultados, se puede decir que la flexibilidad del sistema es superior al 60%, hay que considerar que al momento el sistema se desarrolló en una determinada facultad y se dispone al momento de un prototipo puesto a prueba.

Participación de los Usuarios

6. ¿Tengo el conocimiento y las habilidades necesarias para utilizar el sistema?

20 respuestas

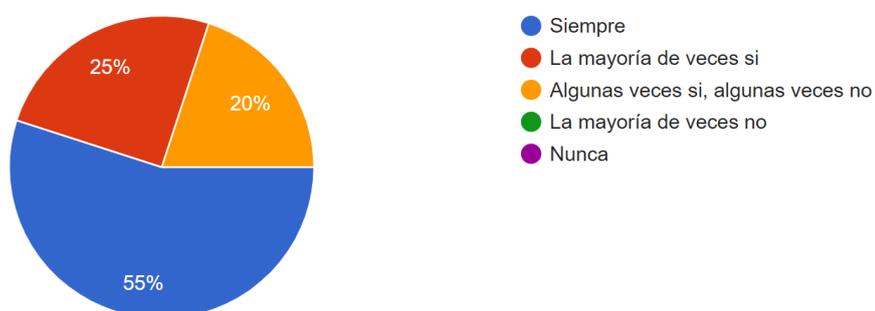


Figura 49. Tiene el usuario el conocimiento y las habilidades para utilizar el sistema

Del total encuestados, se tiene que en un 55% poseen siempre el conocimiento y las habilidades necesarias para utilizar el sistema, mientras que un 25% mencionó que la mayoría de las veces si y un 20% mencionó que algunas veces si, algunas veces no.

7. Conozco todas las funcionalidades que posee el sistema

20 respuestas

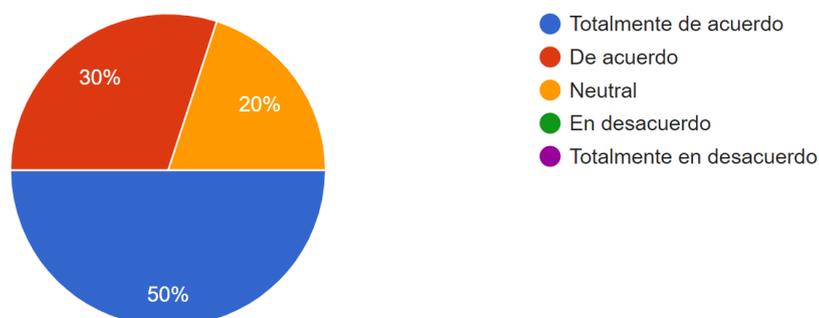


Figura 50. Conoce el usuario todas las funcionalidades del sistema

En relación con el conocimiento de las funcionalidades que posee el sistema, se tiene que el 50% se encuentra totalmente de acuerdo, mientras que el 30% se encuentra de acuerdo y un 20% se muestra neutral al conocimiento de la misma, esto permite determinar que el conocimiento en relación al sistema es progresivo, a medida que se generen más interacciones con el mismo, mejorará su conocimiento.

8. Estoy familiarizado con los procedimientos para alcanzar los objetivos establecidos.

Por ejemplo: Conozco el procedimiento para realizar el cálculo de indicadores

20 respuestas

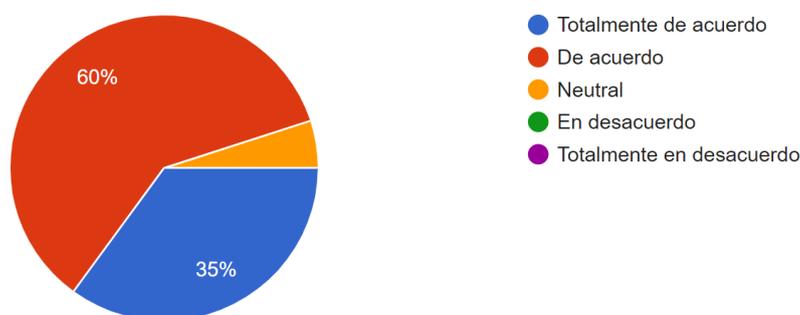


Figura 51. Conoce el usuario los procedimientos para alcanzar los objetivos de evaluación

En relación, al conocimiento del cálculo de indicadores se tiene que, del total de encuestados, el 60% se encuentra totalmente de acuerdo, mientras que el 35% se encuentra de acuerdo y un 5% tiene una posición neutral.

En relación con la participación de los usuarios en los procesos de evaluación y sus objetivos se tiene que más del 80% de los usuarios participan en los procesos y poseen el conocimiento necesario para utilizar el sistema.

Tareas

9. Las tareas que se realizan con el sistema, ¿se utilizan habitualmente?

20 respuestas

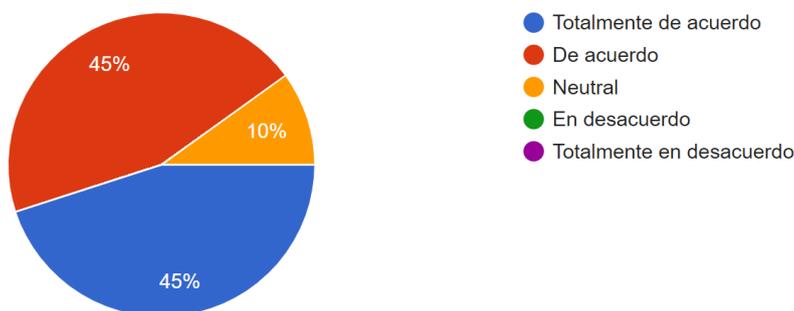


Figura 52. Las tareas que realiza el sistema son habituales para el usuario

Con relación a esta pregunta se tiene que el 45% de los encuestados, está totalmente de acuerdo con el hecho de que las tareas que se realizan con el sistema son habituales, mientras que el 45% se encuentra de acuerdo y el 10% se encuentra en forma neutral.

10. El sistema le permite solucionar alguna tarea complicada

20 respuestas

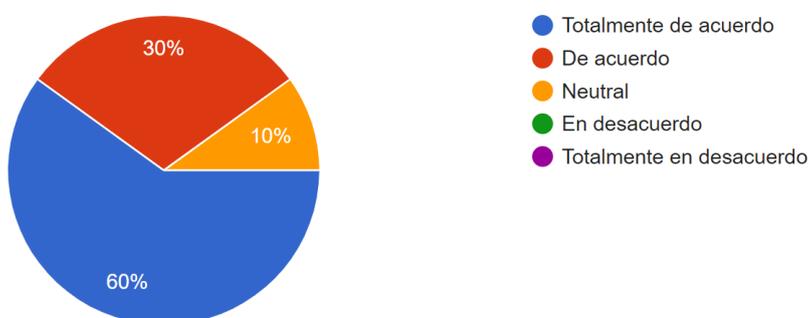


Figura 53. Soluciona el sistema alguna tarea complicada

En relación con esta pregunta se tiene que el 60% de los encuestados está totalmente de acuerdo, mientras que el 30% se encuentra de acuerdo y el 10% se ubica en una posición neutral.

Al evaluar la respuesta del sistema en relación a las tareas que desarrolla los encuestados se tiene que más del 60% está totalmente de acuerdo con el apoyo que tiene el sistema al desarrollo de sus tareas, más si estas son complicadas.

Usabilidad

11. ¿Cree usted que el sistema es amigable con el usuario?

20 respuestas

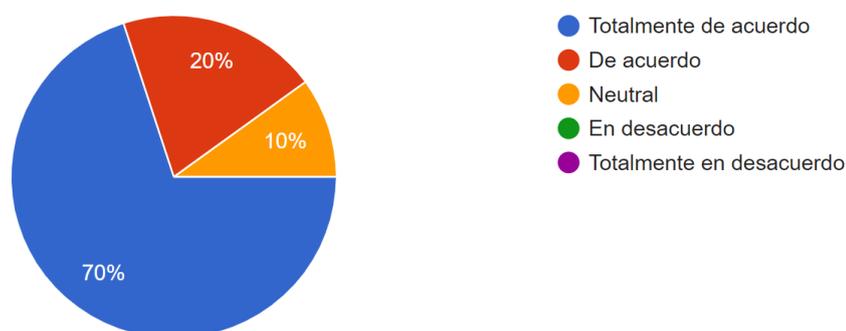


Figura 54. Contribuye el sistema a generar indicadores de evaluación

En esta pregunta, se tiene que el 70% de los encuestados, se encuentra totalmente de acuerdo con que el sistema es amigable al usuario, un 20% se encuentra de acuerdo y un 10% se ubica en una posición neutral.

12. Cree usted que el sistema contribuye al cumplimiento rápido y fácil de tareas y objetivos establecidos

Del total de encuestados se tiene que el 80% ratifica que el sistema a plataforma contribuye al cumplimiento rápido y fácil de tareas y objetivos establecidos en los procesos de autoevaluación.

20 respuestas

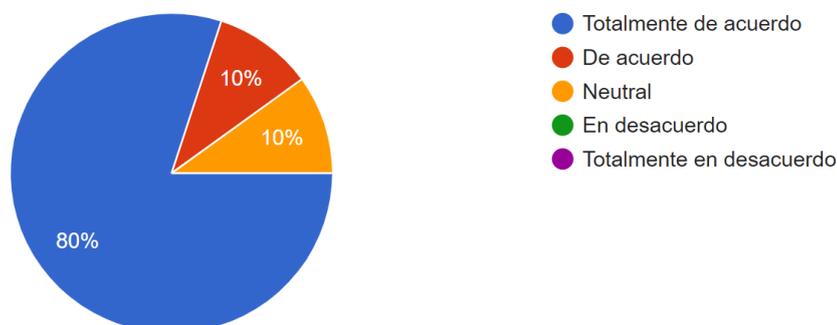


Figura 55. Contribuye el sistema al cumplimiento rápido y fácil de tareas y objetivos

La usabilidad del sistema de gestión de indicadores se vio demostrada al tener más del 70% de criterios, en los cuales los usuarios mencionaron que se encuentran totalmente de acuerdo en que el sistema es amigable y permite el cumplimiento rápido y fácil de tareas y objetivos establecidos en los procesos de evaluación.

Resumen General de los Resultados Obtenidos

La usabilidad del sistema se evaluó mediante la definición de los siguientes criterios: facilidad de aprendizaje del sistema, eficiencia del sistema, flexibilidad del sistema, participación de los usuarios del sistema y la usabilidad del sistema, en función de estos criterios se determinó que:

- Más del 70% de los usuarios encuestados establecieron que el sistema siempre es fácil de aprender.
- En relación a la eficiencia del sistema, se tiene que este apoya y contribuye a la generación de indicadores de evaluación, necesarios para los procesos de autoevaluación de las carreras de la ESPOCH.
- En relación a la flexibilidad del sistema, se tiene que el mismo es superior al 60%, cabe recalcar que al momento el sistema se desarrolló para una determinada facultad y se encuentra en la fase de pruebas para su posterior implementación.

- En relación a la participación de los usuarios en los procesos de evaluación y sus objetivos, se tiene que más del 80% de los usuarios participan en los procesos y poseen el conocimiento necesario para utilizar el sistema.
- Al evaluar la respuesta del sistema en relación a las tareas que desarrolla los encuestados se tiene que más del 60% está totalmente de acuerdo con el apoyo que tiene el sistema al desarrollo de sus tareas, más si estas son complicadas
- La usabilidad del sistema de gestión de indicadores se vio demostrada al tener más del 70% de criterios, en los cuales los usuarios mencionaron que se encuentran totalmente de acuerdo en que el sistema es amigable y permite el cumplimiento rápido y fácil de tareas y objetivos establecidos en los procesos de evaluación

Síntesis de las Preguntas de Investigación Planteadas

Se presenta una síntesis de las preguntas de investigación formuladas al inicio de investigación, con el objetivo de resumir los principales resultados alcanzados.

OE1. P1: ¿En qué consiste el data warehousing y la minería de datos aplicada a la educación?

El Data Warehousing es un proceso de construcción y uso de un almacén de datos. La construcción del almacén de datos incluye la limpieza, integración y consolidación de datos provenientes de diferentes fuentes de almacenamiento que posee una organización.

Por otro lado, la minería de datos aplicada a la educación se denomina minería de datos educacional (EDM), la misma que se constituye en un campo de tecnología emergente y la mayoría de las aplicaciones se enfocan en el análisis de datos de los estudiantes con el fin de mejorar la calidad de la educación, el rendimiento académico y el desempeño de los docentes. La minería de datos educacional se aplica en entornos presenciales y virtuales, tienen como

objetivo mejorar los procesos académicos, así como generar alternativas de solución frente a los hallazgos descubiertos.

OE1.P2: ¿Cuáles son las técnicas de minería de datos y las técnicas de validación de datos que contribuyen a mejorar los resultados en la evaluación de la IES?

Los tipos de información que se pueden obtener de la minería de datos son: asociaciones, secuencias, clasificaciones, agrupamientos (clustering) y pronósticos (Laudon & Laudon, 2016).

Los tipos de información se constituyen en las técnicas de minería de datos, las mismas que crean modelos que son predictivos y/ o descriptivos.

De todas las técnicas de análisis de datos, en el presente proyecto se empleó el clustering debido a los resultados obtenidos en el mapeo sistemático de literatura y la necesidad de mejorar el rendimiento académico de los estudiantes.

El Clustering es el proceso de agrupar datos en clases o clusters dentro de un grupo de observaciones, el objetivo del clustering es que los objetos dentro de un grupo sean similares (o relacionados) entre sí y diferentes de (o no relacionados con) los objetos en otros grupos. A mayor similitud (u homogeneidad) dentro de un grupo y a mayor diferencia entre grupos, mejor o más distinto es el clustering.

Validación en el clustering

La validación de *clusters* es el proceso por el cual se evalúa la veracidad de los grupos obtenidos. A modo general, este proceso consta de tres partes: estudio de la tendencia de *clustering*, elección del número óptimo de *clusters* y estudio de la calidad/significancia de los *clusters* generados (Amat, 2017)

OE2.P1: ¿El modelo de gestión de indicadores es adecuado para la evaluación institucional y de carreras?

Del análisis a la usabilidad del modelo de gestión de indicadores, se tiene que más del 80% de los usuarios encuestados, afirman que el sistema contribuye al cumplimiento rápido y fácil de tareas y objetivos establecidos

OE2.P2: ¿Qué indicadores de evaluación se puede obtener con el modelo de gestión?

Los indicadores de evaluación que se pueden obtener al modelo de gestión son:

Indicadores de evaluación institucional

- Doctores
- Doctores tiempo completo
- Doctores en formación
- Estudiantes por profesora
- Titularidad
- Titularidad Tiempo Completo
- Profesoras mujeres
- Dirección Mujeres
- Producción científica Impacto Mundial
- Producción científica Impacto Regional
- Tasa de Retención
- Tasa de Graduación Grado
- Tasa de Graduación Postgrado

Indicadores de evaluación de carrera

- Afinidad Formación Postgrado
- Actualización Científica y/o pedagógica
- Titularidad

- Profesor TC, MT,TP
- Estudiante por profesor
- Producción Regional
- Libros y capítulos de libros

OE3.P1: ¿Cuál es el nivel de aceptación que tiene el modelo de gestión de indicadores?

Del análisis a la aceptación del modelo de gestión se tiene que más del 80% de los usuarios se encuentran de acuerdo con el sistema y les permite obtener un cumplimiento fácil de tareas y objetivos establecidos.

OE3.P2: ¿Cuál es el resultado óptimo que se debe considerar al implementar el modelo de gestión de indicadores?

El resultado óptimo se obtuvo al realizar la validación del clustering, proceso en el cual se hizo un estudio de la tendencia de clustering, la elección del número óptimo de clusters y el estudio de la calidad/significancia de los clusters generados.

CONCLUSIONES

Del análisis se tiene que mediante el sistema se han generado indicadores de evaluación y de gestión académica, que contribuyen al desarrollo de los procesos de autoevaluación de las carreras, evaluación institucional y la gestión de sus autoridades.

La revisión de literatura permitió determinar que al momento no existen trabajos de investigación que articulan los procesos de evaluación y aseguramiento de la calidad con la inteligencia de negocios, más aún cuando se trata de un campo de tecnología emergente.

El sistema desarrollado a partir del modelo de gestión de indicadores de evaluación fue validado mediante la aplicación de un test de usabilidad, según la propuesta de (Poropat, 2014). Se evaluaron los criterios de facilidad de aprendizaje del sistema, eficiencia del sistema, flexibilidad del sistema, participación de los usuarios del sistema y usabilidad del sistema. Los

resultados en general indican que el sistema cumple con los objetivos propuestos: un 70% indicó que el sistema es fácil de aprender, un 80% expresó que el sistema contribuye a generar indicadores de evaluación, un 80% señaló que el sistema es flexible y tiene funcionalidades integradas que pueden ser configuradas de acuerdo a sus necesidades, un 60% refirió que el sistema apoya al desarrollo sus tareas y un 70% respondió que el sistema es amigable y permite el cumplimiento rápido y fácil de tareas y objetivos planteados en los procesos de evaluación.

De acuerdo a los resultados obtenidos, se puede determinar que la hipótesis planteada se cumple en el caso de estudio. Esto es, que el modelo de gestión de indicadores de evaluación mejora la eficiencia de los procesos de autoevaluación de la IES. Lo cual también fue confirmado mediante el clustering ya que se pudo encontrar diferentes patrones de comportamiento, aplicado a diferentes criterios de evaluación.

RECOMENDACIONES

- Se recomienda hacer uso del modelo de gestión de indicadores, no solo para generar indicadores de evaluación, sino también para obtener indicadores de gestión académica que aporten con información y conocimiento para la toma de decisiones en las IES del Ecuador.
- Se recomienda ampliar el desarrollo de plataformas de inteligencia de negocios en los entornos educativo, tecnologías que hoy en día se constituyen como emergentes.
- Se recomienda emplear R, como software para desarrollar la minería de datos en vista de que posee los algoritmos suficientes para desarrollar modelos de predicción, clasificación, clustering, entre otros.

REFERENCIAS BIBLIOGRÁFICAS

- Acosta, B., & Acosta, M. (2016). Modelos de evaluación para la acreditación de carreras. *Revista Mexicana de Investigación Educativa*, 1249-1274. Obtenido de Modelos de Evaluación para la Acreditación .
- Ahmed, A. M., Rizaner, A., & Ulusoy, A. H. (2016). Using data Mining to Predict Instructor Performance. *Procedia Computer Science*, 137-142.
- Amat, J. (2017). Obtenido de Clustering y heatmaps: aprendizaje no supervisado: https://rpubs.com/Joaquin_AR/310338
- Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *In Computers & Education*, 177-194.
- Beguerí, G., Malberti, A., & Klenzi, R. O. (2017). *Minería de datos en la determinación del perfil de un potencial alumno universitario desertor*. Obtenido de http://sedici.unlp.edu.ar/bitstream/handle/10915/20033/Documento_completo.pdf?Sequence=1
- Beltrán, B. (2018). Obtenido de Minería de Datos: <http://bbeltran.cs.buap.mx/notasmd.pdf>
- Cetatech. (2015). <https://cetatech.ceta-ciemat.es>. Recuperado el 2015, de <https://cetatech.ceta-ciemat.es>: <https://cetatech.ceta-ciemat.es/2015/11/introduccion-a-tableau/>
- Chávez, C. L. (2014). El método científico. En C. L. Chávez, *Epistemología y metodología* (págs. 178 - 205). México: Grupo Editorial Patria.
- Curto, J. (2016). *Introducción al business intelligence*. Editorial UOC.
- Departamento de Educación del Gobierno de Aragón. (2013). Obtenido de Criterios de Evaluación: http://aularagon.catedu.es/materialesaularagon2013/formacion_lomce/bloque_1/Modulo_1_1/3_criterios_de_evalucin.html
- Dev.mysql.com. (2015). <https://dev.mysql.com>. Recuperado el 1 de 1 de 2015, de <https://dev.mysql.com>: <https://dev.mysql.com/downloads/windows/installer/8.0.html>
- Devoteam. (2018). Obtenido de Tableau Desktop: <https://rubiconbi.es/requerimientos-tecnicos-tableau-desktop/>
- Díaz, F. J., Osorio, M. A., Amadeo, A. P., & Romero, D. L. (2013). Aplicando estrategias y tecnologías de Inteligencia de Negocio en sistemas de gestión académica. *XV Workshop de Investigadores en Ciencias de la Computación*, 225 - 229.
- Dirección de Evaluación y Aseguramiento de la Calidad. (2017). *Plan de autoevaluación de carreras* . Riobamba.

- Dropscore. (2019). Obtenido de Clustering: aspectos relevantes y métodos de clustering más utilizados: <http://blog.dropscore.com/metodos-de-clustering-mas-utilizados/>
- Ecuador, P. D. (2010). *Ley Orgánica de Educación Superior*. Quito.
- ESPOCH. (2014). *Plan Estratégico Institucional 2014 - 2018*. Riobamba.
- Frankenfield, J. (2019). *Almacenamiento de Datos*. Obtenido de <https://www.investopedia.com/terms/d/data-warehousing.asp>
- Gartner. (2017). Obtenido de IT Glossary: <https://www.gartner.com/it-glossary/>
- Gironés, J., Casas, J., & Minguillón, J. (2017). *Minería de Datos*. Madrid: Jordi Gironés, Jordi Casas, and Julià Minguillón.
- Goette, P. (2014). Obtenido de R, un lenguaje y entorno de programación para análisis estadístico: <https://www.genbeta.com/desarrollo/r-un-lenguaje-y-entorno-de-programacion-para-analisis-estadistico>
- González, A. (2014). *Cleverdata*. Obtenido de ¿Qué es Machine Learning?: <https://cleverdata.io/que-es-machine-learning-big-data/>
- Guevara, C., Ortega, S., Guevara, V., & J. Q. (2016). Business Intelligence aplicado al proceso de seguimiento de graduados de la Universidad Técnica del Norte. 1 - 17. Obtenido de <http://dspace.redclara.net:8080/bitstream/10786/1064/1/Business%20Intelligence%20aplicado%20al%20proceso%20de%20seguimiento%20de%20graduados%20de%20la%20Universidad%20T%C3%a9cnica.pdf>
- Hasperué, W. (2017). *Clustering*. La Plata .
- Hernandez, R., Fernandez, C., & Baptista, M. D. (2010). *Metodología de la Investigación*. México: mcgrawhill.
- Hitachi Vantara. (2019). Obtenido de Meet Performance and Scalability Needs With Hitachi Vantara's Pentaho: <https://www.hitachivantara.com/en-us/pdf/white-paper/performance-scalability-overview-whitepaper.pdf>
- Hoz, A. D. (1985). Modelos de Evaluación Universitaria. *Revista Española de Pedagogía*, 521 - 537. Obtenido de Modelos de Evaluación Universitaria.
- Laudon, K., & Laudon, J. (2016). *Sistemas de Información Gerencial*. México: Pearson Educación.
- Microsystem. (2019). Obtenido de Rapidminer: <https://www.microsystem.cl/plataforma/rapidminer/>
- Ministerio de Educación Chile. (2016). Obtenido de Taller 3: Criterios de Evaluación : <http://basica.mineduc.cl/wp-content/uploads/sites/25/2016/06/ANEXO5.pdf>

- Morales, E., & Jair, H. (2018). Obtenido de Clustering: <https://ccc.inaoep.mx/~emorales/Cursos/NvoAprend/Acetatos/clustering.pdf>
- Moss, L. T., & Atre, S. (2003). *Business Intelligence Roadmap: The Complete Project Lifecycle for Decision-Support Applications*. Addison-Wesley Information Technology Series.
- Murillo, M. (2019). *Data warehousing con Pentaho*. Quito: Epn.
- Pentaho. (2011). <https://help.pentaho.com>. Recuperado el 2011, de <https://help.pentaho.com>: https://help.pentaho.com/Documentation/8.1/Setup/Components_Reference
- Pentaho Community. (2011). <https://jossjack.wordpress.com>. Recuperado el 2011, de <https://jossjack.wordpress.com>: <https://jossjack.wordpress.com/2016/03/18/pentaho-community-introduccion/>
- Poropat, E. (2014). *Evaluation of Business Intelligence System Usability*. Ljubljana.
- Puce. (2005). *Guía de autoevaluación con fines de acreditación para las universidades y escuelas politécnicas*. Obtenido de <ftp://ftp.puce.edu.ec/Facultades/SinaPuce/pucematriz/conea/.../capitulos.doc>
- R Tools Technology. (2019). *Documentación Técnica de R Studio*. Obtenido de https://www.r-studio.com/es/Unformat_Help/systemrequirements.html
- RapidMiner. (2019). *Requisitos del sistema RapidMiner Studio*. Obtenido de <https://docs.rapidminer.com/latest/studio/installation/system-requirements.html>
- Rodriguez, T. (2017). *Xataka*. Obtenido de Machine Learning y Deep Learning: cómo entender las claves del presente y futuro de la inteligencia artificial: <https://www.xataka.com/robotica-e-ia/machine-learning-y-deep-learning-como-entender-las-claves-del-presente-y-futuro-de-la-inteligencia-artificial>
- Sinnexus. (2019). Obtenido de Datawarehouse: https://www.sinnexus.com/business_intelligence/datawarehouse.aspx
- SM, K., & Belwal, M. (2007). A Service Oriented Architecture for Business Intelligence. *IEEE International Conference*, 1201 - 1207.
- The information lab. (2016). <https://www.theinformationlab.co.uk>. Recuperado el 2016, de <https://www.theinformationlab.co.uk>: <https://www.theinformationlab.co.uk/2017/03/07/recommended-not-minimum-tableau-desktop-system-requirements/>
- Timarán-Pereira, S. R.-A.-Z.-T. (2016). *El proceso de descubrimiento de conocimiento en bases de datos*. Bogota: Ediciones Universidad Cooperativa de Colombia.

TutorialsPoint. (20 de 04 de 2019). *¿Qué es el almacenamiento de datos?* Obtenido de https://www.tutorialspoint.com/dwh/dwh_data_warehousing.htm

Universidad de Salamanca. (2018). Obtenido de Calidad Educativa: <http://ocw.usal.es/ciencias-sociales-1/investigacion-evaluativa-en-educacion/contenidos/Calidad.pdf>

Wikipedia. (2018). *es.wikipedia.org*. Recuperado el 15 de octubre de 2018, de *es.wikipedia.org*: https://es.wikipedia.org/wiki/MySQL_Workbench#Caracter%C3%ADsticas

Zucco, C. (2018). Minería de datos en Bioinformática. *Módulo de referencia en Ciencias de la vida*.