



Universidad Nacional

SAN LUIS GONZAGA



[Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional](https://creativecommons.org/licenses/by-nc-sa/4.0)

Esta licencia permite a otras combinar, retocar, y crear a partir de su obra de forma no comercial, siempre y cuando den crédito y licencia a nuevas creaciones bajo los mismos términos.

<http://creativecommons.org/licenses/by-nc-sa/4.0>



UNIVERSIDAD NACIONAL "SAN LUIS GONZAGA"



ESCUELA DE POSGRADO

EVALUACION DE ORIGINALIDAD

CONSTANCIA

El que suscribe, deja constancia que se ha realizado el análisis con el software de verificación de similitud al **BORRADOR DE TESIS** cuyo título es:

**"SEGUIMIENTO DEL EGRESADO CON TÉCNICAS DE MINERÍA DE DATOS:
CASO FACULTAD DE INGENIERÍA PESQUERA DE LA UNIVERSIDAD
NACIONAL SAN LUIS GONZAGA PERIODO 2017-2021"**

Presentado por:

SANDOVAL PEREZ LOIS ALEXANDER

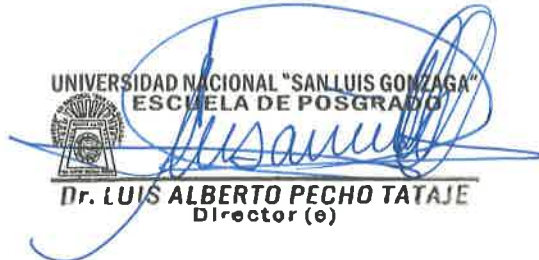
**DE LA MAESTRÍA EN INGENIERÍA DE SISTEMAS mención GESTIÓN DE
TECNOLOGÍAS DE LA INFORMACIÓN.**

Que, se ha recibido del operador del programa informático evaluador de originalidad de la Escuela de Posgrado de la UNICA, el informe automatizado de originalidad, el mismo que concluye de la siguiente manera:

El documento de investigación APRUEBA los criterios de originalidad con un porcentaje de similitud de 2%.

Para dar fe, se adjunta al presente el reporte de similitud de las bases de datos de iThenticate. En Ica 23 de octubre de 2024

Atentamente


UNIVERSIDAD NACIONAL "SAN LUIS GONZAGA"
ESCUELA DE POSGRADO
Dr. LUIS ALBERTO PECHO TATAJE
Director (e)

UNIVERSIDAD NACIONAL “SAN LUIS GONZAGA”
VICERRECTORADO DE INVESTIGACIÓN
ESCUELA DE POSGRADO
MAESTRIA: INGENIERÍA DE SISTEMAS
MENCION: GESTION DE TECNOLOGIAS DE LA INFORMACION



TESIS

SEGUIMIENTO DEL EGRESADO CON TÉCNICAS DE MINERÍA DE DATOS: CASO FACULTAD DE INGENIERÍA PESQUERA DE LA UNIVERSIDAD NACIONAL SAN LUIS GONZAGA PERIODO 2017-2021

Línea de investigación:

Ciencias naturales, ingeniería y tecnologías sostenibles

PRESENTADO POR:

LOIS ALEXANDER SANDOVAL PEREZ

GRADO A OBTENER: MAESTRO

ASESOR:

Dr. PEÑA CASAS, ERWIN PABLO

Ica, Perú

2025

Dedicatoria

A Dios con mucho amor y gratitud, a mi señor padre Juan Sandoval que desde el cielo me guía y mi madre Delia Pérez por su apoyo y consejos para hacer de mí una mejor persona.

Agradecimientos

En primera instancia agradezco a mis formadores, personas de gran sabiduría quienes se han esforzado por ayudarme a llegar al punto en el que me encuentro.

Esta ocasión no ha sido la excepción, y exalto su trabajo, y le agradezco con creces por ayudarme a lograr esta nueva meta. Mi maestría.

ÍNDICE DE CONTENIDOS

Dedicatoria	ii
Agradecimientos	iii
Índice de contenidos	iv
Índice de tablas	v
Índice de figuras	vi
Resumen	viii
Abstract	ix
I. INTRODUCCIÓN	1
II. ESTRATEGIA METODOLOGICA.....	2
2.1 Planteamiento del problema	2
2.1.1 Antecedentes de la investigación	2
2.1.2 Bases Teóricas	9
2.1.3 Realidad problemática.....	10
2.1.4 Formulación del Problema	13
2.1.5 Justificación e importancia de la investigación.....	13
2.2 Objetivos	14
2.2.1 Objetivo general	14
2.2.2 Objetivos específicos	14
2.3 Hipótesis y variables de la investigación	14
2.3.1 Hipótesis General	14
2.3.2 Hipótesis Específicas.....	14
2.3.3 Variables de la investigación	15
2.3.4 Variable Independiente	15
2.3.5 Variable Dependiente.....	15
2.3.6 Operacionalización de las variables.....	15
III. ESTRATEGIA METODOLÓGICA	16
3.1 Estrategia de la investigación	16
3.2 Método de investigación	16
3.3 Aplicación del método de investigación	17

3.3.1 Comprensión del problema	17
3.3.2 Comprensión de los datos	19
3.3.3 Preparación de los datos	22
3.3.4 Modelización	27
3.3.5 Evaluación de los resultados	43
IV. RESULTADOS	45
V. DISCUSIÓN.....	50
VI. CONCLUSIONES	51
VII. RECOMENDACIONES	52
VIII. REFERENCIAS BIBLIOGRÁFICAS	53
IX. ANEXOS	55

Índice de tablas

TABLA I ESTUDIANTES POR CICLO ACADÉMICO PERIODO 2017-2021 – PESQUERÍA	11
TABLA II ESTUDIANTES POR CICLO ACADÉMICO PERIODO 2017-2021 – PESQUERÍA	12
TABLA III EGRESADOS POR CADA CICLO ACADÉMICO Y ESCUELA PROFESIONAL	12
TABLA IV ESTIMACION ESTADISTICA DE LOS DATOS DE INGENIERIA PESQUERA	49
TABLA V ESTIMACION ESTADISTICA DE LOS DATOS DE INGENIERIA DE ALIMENTOS	49

Índice de figuras

Fig. 1. CRISP-DM Methodology Process	16
Fig. 2. Datos de la escuela de pesquería	17
Fig. 3. Datos de la escuela de alimentos	18
Fig. 4. Registro de datos de egresados1	19
Fig. 5. Registro de datos de egresados2	19
Fig. 6. Registro de datos de egresados3	19
Fig. 7. Registro de datos de egresados4	20
Fig. 8. Información de los datos	20
Fig. 9. Información detallada de los datos	20
Fig. 10. Datos faltantes en los registros de los egresados	21
Fig. 11. Calidad de los datos	22
Fig. 12. Datos del Dataset	22
Fig. 13. Datos seleccionados	23
Fig. 14. Datos de las características o atributos seleccionados	24
Fig. 15. Datos atípicos en la Edad	24
Fig. 16. Valores perdidos en la edad	25
Fig. 17. Scatter plot de los datos	25
Fig. 18 Scatter plot de la relación Edad y tiempo del bachiller	26
Fig. 19 Nuevas características generadas	27
Fig. 20. Egresados de las escuelas académicas por ciclo de egreso	28
Fig. 21. Egresados por sexo de cada ciclo académico	28
Fig. 22. Distribución de egresado por escuela académica y sexo	29
Fig. 23. Egresado según su estado civil	29
Fig. 24. Distribución de egresados según escuela académica y estado civil	30
Fig. 25. Distribución de egresados según el sexo y estado civil	30
Fig. 26. Egresados con estado del bachillerato	30
Fig. 27. Estado de los bachilleres a la fecha del procesamiento	31
Fig. 28. Estado de los egresados con relación al bachiller	31
Fig. 29. Estado de los titulados a la fecha del procesamiento	32
Fig. 30. Estado de los egresados en relación con el título profesional	32
Fig. 31. DataSet de bachilleres	33
Fig. 32. DataSet de los titulados	34
Fig. 33. Distribución de los egresados según edad y tiempo de estudio	34

Fig. 34. Distribución de los bachilleres según edad y tiempo de obtención	35
Fig. 35. Distribución de los titulados según la edad y tiempo en obtener el título	35
Fig. 36. Distribución de los bachilleres según su edad con un centroide	36
Fig. 37. Distribución de los Titulados con tres centroides	37
Fig. 38. Edad de los egresados por escuela académica	37
Fig. 39. Tiempo de estudios de los egresados por escuela académica	38
Fig. 40. Edad del Bachiller al obtener el bachiller según la escuela académica	38
Fig. 41. Tiempo en obtener el bachiller según la escuela académica	39
Fig. 42. Edad del titulado al obtener el título profesional por escuela académica	39
Fig. 43. Tiempo en obtener el título por escuela académica	40
Fig. 44. Modelo de minería de datos empleado	41
Fig. 45. Diseño de modelo de prueba	42
Fig. 46. Conversión de datos categóricos en datos numéricos	42
Fig. 47. Resultados de las evaluaciones del modelo	43
Fig. 48. Resultados de las predicciones	43
Fig. 49. Egresados por semestre y escuela académica	45
Fig. 50. Egresados, bachilleres con título profesional	45
Fig. 51. Egresados por escuela y sexo	46
Fig. 52. Egresados por escuela y estado civil	46
Fig. 53. Bachilleres por escuela académica	47
Fig. 54. Relación entre bachilleres por escuela	47
Fig. 55. Titulados por escuela académica	48
Fig. 56. Relación de titulados por escuela académica	48

RESUMEN

El Objetivo: Evaluar la información de los egresados de la facultad de Ingeniería Pesquera de la universidad nacional San Luis Gonzaga utilizando minería de datos para el periodo 2017-2021. **La metodología** utilizada del tipo aplicada, no experimental, de corte transversal retrospectivo, para lo cual se utilizó la Metodología CRISP-DM (Cross-Industry Standard Process for Data Mining), aplicando el algoritmo para aprendizaje no supervisado K-Means, del software Orange Data Mining, siguiendo cada una de sus fases. **Los resultados** el tiempo de duración de la carrera profesional, están entre los tiempos establecidos de 5 años, con una edad media de egreso entre 23.41-23.44 para ambas escuelas pesquería y alimentos. Otros resultados, muestran que han obtenido su bachiller solo el 69.88% de la escuela de alimento y 77.46% de la escuela de pesquería siendo este valor muy reducido, considerando que el bachiller es automático. Más se agudiza con los titulados, solo el 22.41% de la escuela de alimentos y 16.42% de la escuela de pesquería. La información de las variables sobre el trabajo de los egresados es muy escasa por lo cual no se pudo realizar el análisis. **Concluyendo** que el software de minería de datos aporta importantes resultados para la facultad de ingeniería pesquera que mejorar sus procesos de bachillerato y titulación, así como el registro de la condición laboral de cada egresado, bachiller o titulado que es muy escasa.

Palabras clave: Minería de datos, Metodología CRISP-DM, Seguimiento de egresados

ABSTRACT

The Objective: To evaluate the information of the graduates of the Faculty of Fisheries Engineering of the Universidad Nacional San Luis Gonzaga national university using data mining for the period 2017-2021. **The methodology** used was applied, non-experimental, retrospective cross-sectional, for which the CRISP-DM Methodology (Cross-Industry Standard Process for Data Mining) was used, applying the algorithm for unsupervised learning K-Means, of the Orange Data Mining software, following each of its phases. **The results** show that the duration of the professional career is within the established time frame of 5 years, with an average age of graduation between 23.41-23.44 for both the fishery and food schools. Other results show that only 69.88% of the food school and 77.46% of the fishery school have obtained their bachelor's degree, which is very low, considering that the bachelor's degree is automatic. It is even more acute with the graduates, only 22.41% of the food school and 16.42% of the fishery school. The information of the variables on the work of the graduates is very scarce, so the analysis could not be carried out. **In conclusion**, the data mining software provides important results for the faculty of fisheries engineering to improve its baccalaureate and degree processes, as well as the record of the employment status of each graduate, baccalaureate, or degree holder, which is very scarce.

Keywords: Data mining, CRISP-DM Methodology, Graduate tracking

I. INTRODUCCIÓN

En un entorno académico y profesional en constante evolución, la capacidad de las instituciones educativas para adaptar sus programas y estrategias en respuesta a las necesidades del mercado laboral y las trayectorias profesionales de sus egresados se ha convertido en un factor crucial para su éxito y relevancia. En este contexto, la Facultad de Ingeniería Pesquera de la Universidad Nacional San Luis Gonzaga enfrenta el desafío de comprender y mejorar el seguimiento de sus egresados, especialmente en un campo tan dinámico como la ingeniería pesquera.

La gestión efectiva del seguimiento de egresados no solo proporciona una visión clara sobre la empleabilidad y el desarrollo profesional de los graduados, sino que también permite a las instituciones educativas ajustar sus currículos y programas de apoyo para satisfacer mejor las demandas del mercado laboral. Sin embargo, la recolección y análisis de datos sobre la trayectoria profesional de los egresados puede ser un proceso complejo y laborioso, que tradicionalmente ha dependido de métodos manuales y limitados.

En este escenario, la minería de datos surge como una herramienta poderosa para transformar grandes volúmenes de datos en información valiosa y procesable. La minería de datos, que involucra la aplicación de técnicas analíticas avanzadas para descubrir patrones y relaciones ocultas en los datos, ofrece una oportunidad única para mejorar el seguimiento de egresados mediante la identificación de tendencias y la segmentación de la población estudiada.

Este estudio tiene como objetivo principal explorar cómo la aplicación de técnicas de minería de datos puede mejorar el seguimiento y análisis del desempeño laboral y la trayectoria profesional de los egresados de la Facultad de Ingeniería Pesquera durante el período 2017-2021. Al utilizar herramientas y métodos de minería de datos, se busca obtener insights detallados que permitan una comprensión más profunda de la empleabilidad, el tipo de empleo, el nivel de satisfacción laboral y la continuación de estudios de los egresados.

La investigación se centra en varios aspectos clave: en primer lugar, la identificación y evaluación de los datos disponibles sobre los egresados; en segundo lugar, la selección y aplicación de técnicas de minería de datos para extraer patrones y tendencias significativas; en tercer lugar, la implementación y evaluación de modelos de análisis para identificar relaciones relevantes; y, por último, la discusión de los desafíos y limitaciones encontrados en el proceso, así como su impacto en la toma de decisiones institucionales.

A través de esta investigación, se pretende proporcionar recomendaciones prácticas para mejorar los procesos de seguimiento de egresados y adaptar las estrategias formativas y de apoyo a las necesidades emergentes del mercado laboral. El objetivo es ofrecer una guía basada en datos que permita a la Facultad de Ingeniería Pesquera optimizar su enfoque hacia la formación y el apoyo a sus egresados, contribuyendo así a su éxito profesional y a la mejora continua de sus programas educativos.

II. ESTRATEGIA METODOLOGICA

2.1 Planteamiento del problema

2.1.1 Antecedentes de la investigación

2.1.1.1 Antecedentes Internacionales

En la investigación enmarcada con el Seguimiento de Egresados de la facultad de Odontología de la UNLP, en la que se realiza un estudio orientado a establecer un sistema de información con los datos básicos, práctica profesional, experiencias, opiniones y sugerencias de los graduados de esta casa de estudios. El trabajo se localiza dentro del cuadro del proyecto de investigación de Seguimiento de Egresados de la Facultad de Odontología de la UNLP. Es un estudio que tiene por objeto establecer un sistema de información sobre datos básicos, práctica profesional, experiencias, opiniones y sugerencias de graduados. El estudio de nivel descriptivo. Las peticiones relacionadas en el terreno laboral en un contexto cada vez más competitivo, están obligando a las universidades a estar en constante desarrollo con la finalidad de formar graduados con pensamiento autónomo, predispuestos al cambio que es lo que exige el mercado actual, con análisis crítico, propios y predispuestos a desarrollar capacidades de aprendizaje continuo. En el estudio cerca del 46% se desempeña profesionalmente de manera independiente, el 70.4% se pudo dedicar exclusivamente al ejercicio de su profesión, con un 52.7% en el ámbito privado y solo un 13% en el ámbito privado es lo más resaltante [1].

En esta investigación en la misma orientación que la nuestra, cuyo objetivo fue desarrollar un modelo de minería de datos a fin de encontrar patrones de empleo en los egresados, así mismo se contribuya en la mejora del porcentaje de inserción al trabajo con la formación profesional. El trabajo reconoce la preocupación por la inserción de los egresados en empleo. La metodología empleada Knowledge Discovery in Databases (KDD) con etapas claramente definidas pero simples, se utilizaron algoritmos de árboles de decisión, redes neuronales y la regresión con la finalidad de establecer la integración de los egresados al campo laboral. El algoritmo de red neuronal fue el que más se adaptó en la solución del problema, este por ser más exacto, preciso, sensible, tiempo de procesamiento, rendimiento y facilidad para interpretar los datos y sus resultados con el modelo empleado. Las variables predictoras con mayor significancia detectadas fueron: 1) uso de conocimientos y habilidades en su trabajo actual, 2) el año de graduación y si buscó trabajo al término de su carrera, y 3) labora actualmente (variable dependiente) [2].

En este mismo contexto en la cual el estudio se aboca tener una caracterización del desempeño profesional de los egresados de las tres primeras promociones de la maestría en

Fármaco epidemiología comprendida entre los años 2010 y 2018, en el estudio no se ha considerado el seguimiento de los egresados en el contexto del desempeño laboral. La investigación de nivel descriptivo transversal, en la que se estudió a 73 egresados en las que se midieron variables como cargo relacionado con el fármaco epidemiología, antes y después de sus estudios, competencias en el área de conocimiento, actividades relacionadas con el área de conocimiento, promoción del cargo, categoría docente y científica, la encuesta se realizó por medio de correo electrónico. 52 egresados respondieron la encuesta, de los cuales el 57,7 % vivía en La Habana. Preponderaron las mujeres (69,2 %), la mayor proporción de edad entre 41 y 50 años (53,8%) y los médicos (52 %). La influencia de la maestría fue aceptada por la mayoría para informar temas científicos y efectuar investigaciones. El 46,2 % laboraba en fármaco epidemiología posteriormente del egreso; el 42,3 % efectuaba actividades de fármaco epidemiología como expertos del programa nacional de medicamentos, conformaba los comités farmacoterapéuticos en su organización y desarrollaba estudios de uso de medicamentos, además; el 53,8 % fue promovido de cargo; el 46,2 % publicó los resultados de su tesis de maestría; y el 38,5 % logró categoría docente. La investigación concluye que hubo cambios positivos en el desenvolvimiento profesional, sobre todo en actividades de investigación, considerando un impacto importante en la formación de egresados. La mayoría no labora en fármacos epidemiológicos, ni realizaba actividades relacionadas, siendo incoherente con la formación y su actividad laboral basados en su capacitación.[3].

En otro camino el trabajo cuyo objetivo fue el desarrollo de una plataforma web para la obtención del grado académico basando su estudio en el seguimiento a graduados del programa de Ingeniería de Sistemas, con la finalidad de sistematizar las graduaciones, con la incorporación de información personal, laboral y académica con la finalidad de facilitar la evaluación del desarrollo profesional. El estudio utilizó la búsqueda de información de los graduados en las redes sociales por medio de sus datos personales disponibles en la oficina de graduados de la universidad. La búsqueda se realizó en las redes sociales de Facebook, LinkedIn y Twiter para 15 graduados, los resultados muestran que se pudo obtener mayor información de las redes de Facebook y LikendIn para el desarrollo del portal web, TrackGrad (Tracking Graduates), tomándose esta como fuente de datos [4].

El objetivo Acreditar la calidad de los contenidos programáticos en los pensum con la finalidad de ofrecer mejores servicios académicos, por medio de sus estudiantes con ciertas competencias para desarrollarse en el mercado laboral. En este camino, los estudiantes que se titulan no solo dan una imagen del ejercicio profesional, sino que se muestra como un referente

de calidad de la IES de la cual egresa. En tal sentido el área de Bienestar Universitario, la corporación universitaria de Sabaneta Unisabaneta, promulga establecer un contacto constante con sus egresados, a fin de realizar las mediciones de calidad de sus programas académicos y la imagen que proyectan sus profesionales formados por la IES en el mercado laboral. Esto debido a que se encuentran en constante evaluación, repercutiendo en la percepción de las IES en el mercado y de ofertas de profesionales. El estudio de nivel descriptivo, para una muestra de 172 egresados a los que se le aplicó la encuesta. Los resultados más relevantes muestran que el 52% son del sexo femenino y 48% del sexo masculino. El 92% cuenta con trabajo, de los cuales el 48% en el sector privado, pero solo un 8% con contrato fijo, el 85% está relacionado su trabajo con el programa académico. Con esto se debe considerar al estudiante como factor de prueba, evidencia y evaluación con la visión de mejora continua, por lo que la investigación se considera un elemento referencial de la calidad académica de la IES [5].

Igualmente, y en la misma línea la investigación de un estudio descriptivo sobre el seguimiento al egresado se constituye en un elemento de retroalimentación obligatoria para los planes de estudio, a fin de que la formación del estudiante en las universidades cubra las brechas que el desarrollo económico y social del país requiere, con orientación al ámbito local y territorial. Se aplicó un cuestionario a una muestra de 314 egresados y graduados. En el caso de la universidad Metropolitana del Ecuador, esta no se encuentra ajena a esta situación, por lo que el presente artículo muestra el resultado obtenido en la investigación descriptiva sobre el seguimiento al egresado y graduado en la carrera de Gestión empresarial de esta universidad y la inserción de sus egresados en el mercado laboral en Quito. Los resultados muestran que el 62% de egresados y graduados de los últimos tres años son del sexo femenino, 6 de cada 10 egresados y graduados trabajan en el sector privado, mientras que 2 de cada 10 solo trabajan en el sector público, de los que trabajan 75% de sus trabajos tienen relación con su profesión [6].

En la misma línea, la investigación cuyo objetivo es analizar el impacto de los egresados en la sociedad, se realiza un estudio de seguimiento desde la óptica de los graduados a fin de evaluar su desenvolvimiento en el mercado laboral, ver la calidad y efectividad del programa académico. con una metodología KDD (Knowledge Discovery in Databases) con métodos de análisis factorial, agrupamiento y minería de texto se elabora el perfil de los graduados entre el 2013 al 2018 de los graduados de Ingeniería Industrial a fin de identificar algunos aspectos que apoyen en la toma de decisiones de los directivos del programa académico. Los resultados más relevantes muestran que el 59.2% son del sexo femenino. El 87.8 son soltero(a). el 69.39% de los encuestados se encuentran trabajando, el 63.89% en trabajo operativo del nivel cero [7].

La investigación en la que se reconoce que todo el esfuerzo que realiza una institución educativa se ve refleja en sus egresados, por lo que el objetivo es identificar los procesos de gestión de dichos egresados. Procesos que se realizan de manera empírica, apoyados con herramientas como la hoja de cálculo en Excel y datos dispersos de diversas fuentes. El presente artículo indaga en la literatura sobre gestión de egresados, la revisión del estado del arte de las redes sociales, conceptos e información publicadas en estas redes sociales, por medio de esta revisión se caracterizan las fases por medio de la metodología CRISP-DM, planteando un modelo adaptado de la metodología a fin de realizar una gestión integral de los egresados, por medio de la minería de datos, soportados con herramientas en la nube, automatizando la información e informes sobre los egresados [8].

El objetivo fue caracterizar la trayectoria, desempeño laboral de los egresados y relación con la formación recibida. se estudiaron diversos factores como: factores personales, socioeconómico, laborales y académico. se reportó que el género femenino tuvo predominancia en las edades de 25 a 35 años, donde 99% estaban tituladas, con un 90% cuyo rendimiento académico fue mayor a 8. con menos de un año en conseguir empleo se encuentra el 97% a través del servicio social; con ello se logra un mayor prestigio a la Institución y su consecuente aceptación en el mercado; 91% contratadas en el sector público, quienes tenían un nivel técnico el 95%, con ingresos mensuales de 5 a 7 sueldos mínimos, en las que el 76% de estos trabajaban en el 2do nivel de atención, con un nivel de satisfacción del 100% en cuanto a los conocimientos y habilidades obtenidas en la institución, se sugirió que la teoría se lleve a la práctica con un 67%, los docentes son considerados con buena formación el 57%, el 100% de estos no cuentan con estudios de posgrado. El estudio de nivel descriptivo [9].

Objetivo definición del índice de empleabilidad. Una buena calidad en la educación superior requiere de diversas condiciones, pero sin duda el mayor impacto lo da el seguimiento a sus egresados, ya que ello permite comprobar su formación y tener necesidades de retroalimentación en el plan curricular. La investigación propuso un índice de empleabilidad que se asocie con un modelo de seguimiento a egresados; con el que se construyó por medio de un nivel hermenéutico con un análisis de contenido a 55 documentos afines, se concluye con un índice de empleabilidad como un aspecto multidimensional cuyos componentes formación, inserción laboral y percepción de empleadores son importantes ya que pueden realizar la medición en un contexto particular de oferta y demanda, mejorando el proceso de autoevaluación y acreditación de calidad [10].

En la investigación igualmente sobre el seguimiento de los egresados de la maestría en Innovación Educativa de la universidad de Sonora. El objetivo de este trabajo fue valorar la pertinencia del plan de estudio, monitorear la trayectoria académica y laboral de los egresados, e identificar áreas de oportunidad para mejorar el programa mediante adecuaciones curriculares y organizativas. Para la obtención de los datos se realizó mediante una encuesta electrónica a tres generaciones de egresados. Los resultados arrojan satisfacción en los egresados, la importancia de las prácticas, la plana de docentes, los contenidos curriculares, la metodología de enseñanza, tutorías y la infraestructura de la institución. Los resultados indican que 41% menciona que su formación fue excelente, el 90% está totalmente de acuerdo con sus profesores y que dominan sus asignaturas, 88% indican que se les reconoce y respetan como personas. 80% con buena predisposición para asesorar trabajos de tesis, el 71% conoce sobre el tema de investigación. El impacto se valida con la inserción laboral de los egresados y la continuación de sus estudios a nivel doctoral [11].

El objetivo realizar el análisis comparativo de dos planes de estudio para presentar los cambios realizados. Similarmente y en la misma línea la licenciatura en Historia de la UACJ en su oferta de formación para investigadores o historiadores, el diseño curricular y plan de estudios fueron mejorados basados en la inserción laboral de sus egresados, cambios que fueron concretizados una década posterior, siendo pertinente conocer la percepción de los estudiantes. El estudio aplicó un cuestionario a 84 egresados de los cuales solo el 47% respondió al cuestionario. El 88% de los encuestados se encuentran trabajando, y el 12% no trabaja. En su mayoría se insertaron al mercado laboral en menos de tres meses, el promedio del salario es bajo para el 57% entre 5,000 a 10,000 pesos y solo el 7.5% tiene un ingreso entre 20,000 a 30,000 pesos. El 52.5% se dedican a la docencia y el 25% en otros sectores. Estos resultados proponen un análisis de las condiciones que predominan, y sugieren algunas modificaciones que mejoren la formación académica de sus estudiantes a fin de culminen con sus tesis y continúen su formación en posgrado y en el campo laboral. [12].

Apoyando aún más la investigación, el estudio sobre el seguimiento a los egresados como herramienta que proporciona información valiosa en las instituciones académicas, el Centro de investigaciones y estudios en salud de Managua que desde el año 2010 viene formando profesionales de la salud, optó por tener información del estado actual de sus egresados en los diversos periodos. El objetivo Conocer la situación actual de los egresados del programa salud ocupacional del Centro de Investigaciones y Estudios de Salud (CIES) UNAN Managua, periodo 2010-2016, cuya metodología se basó en un estudio descriptivo transversal, mediante una

encuesta virtual con Google Forms de los 15 egresados que respondieron 13 eran del sexo femenino, 8 tenían entre 31-35 años, 5 profesionales de salud y 5 licenciados, 14 de ellos eran de Nicaragua, 9 de los egresados con una experiencia de 1 a 3 años, en cuanto a la maestría, 14 satisfechos en aplicar los conocimientos, 13 consideraron herramientas utilizadas en clases y puntos de obstaculicen, 11 tuvieron practicas limitadas. La encuesta refleja que de los 15 que contestaron, todos laboraban, aunque con poca trayectoria laboral, realizar el posgrado tuvo como motivación la profundización de los conocimientos, las pocas prácticas en los módulos y mucho contenido teórico representaron un obstáculo [13].

2.1.1.2 Antecedentes Nacionales

En la revisión bibliográfica sobre antecedentes nacionales, se han encontrado diversas investigaciones y tesis al respecto.

En la investigación desarrollada para aplicar una metodología relacionada al perfil del egresado, establecidos en los procesos de acreditación, tomando como base los datos históricos de las evaluaciones de los egresados. Para ello se aplicó minería de datos, con una metodología basada en un Framework que aplica a los datos históricos de las evaluaciones de las últimas promociones del programa académico de Ingeniería Mecatrónica en la universidad nacional de Trujillo. Con el Framework se realiza la caracterización del plan de estudios con relación a ciertos atributos del perfil de egreso predefinidos. Los datos son procesados y transformados para poder utilizarlos con Minería de Datos. Se aplicó el algoritmo K-Menas para aprendizaje no supervisado, para ello se utilizó la herramienta de minería de datos RapidMiner para encontrar el modelo que mejor se adapte al estudio. Posterior a la ejecución del modelo, se procede a la interpretación y evaluación de los resultados, para luego tener en síntesis el perfil de egreso para el programa académico, logrando tener coherencia. Igualmente se encontró patrones de información que pueden ser utilizadas para toma de decisiones en la mejora del programa académico. con estos resultados se valida el Framework utilizado obtenido por medio de encuesta a expertos del programa de Ingeniería Mecatrónica lográndose tener los indicadores de efectividad y reusabilidad bueno y aceptable respectivamente [14].

Mejorar el control de información de los egresados del Instituto de Educación Superior Pedagógico Público de Chimbote, con la finalidad de mejorar la empleabilidad. La investigación de Estudio de Caso propuso el proyecto del diseño de un sistema de seguimiento a egresados – SISEI en Santa – Chimbote, la investigación realizada en el 2019 cuyo objetivo identificar los factores que obstaculizan el desarrollo de la información de los egresados del SISEI. Estudio enmarcado dentro del modelo de acreditación de los programas de estudios en Institutos y

Escuelas de educación superior correspondiente al estándar 34 “Seguimiento a egresados”. El tipo de investigación realizado fue mixta cualitativa y cuantitativa. Cualitativamente se dio a conocer la percepción de los actores según los objetivos establecidos, y en el análisis cuantitativo se obtuvo la información estadística resultante del SISEI. El análisis realizado se enmarca en las políticas educativas del estado para el seguimiento de los egresados, se evaluaron los logros y resultados obtenidos por el SISEI, se observaron las dificultades en la gestión con el propósito de construir propuestas de Gerencia social y realizar mejoras al SISEI [15].

Al igual que el estudio anterior y en la misma orientación, la investigación de estudio de campo cuyo objetivo es identificar los factores que dificultan el desarrollo de la información de los egresados del SISEI realizada en la Facultad de Ingeniería de la escuela profesional de Sistemas y Computación, en el desarrollo de un sistema para mejorar el seguimiento al egresado en la universidad nacional Hermilo Valdizán de Huánuco, con la finalidad de determinar la influencia del sistema de información transaccional SITE, los resultados muestran que 79% son hombres y 21% son mujeres de una muestra de 19 egresados, solo 10 de ellos utiliza el sistema SITE, el 68.4% conoce el grado de inserción laboral del egresado y 59.7% que el sistema le permite conocer la evolución profesional del egresado [16].

En la siguiente investigación, debido a que la globalización imperante en los países se caracteriza en grandes transformaciones productivas, económicas y sociales. Con la finalidad de adaptarse a estas transformaciones los países están abocados a elevar su competitividad y preparar el capital humano necesario para liderar su desarrollo. Es aquí donde la educación superior se convierte en un factor clave en la construcción del modelo educativo y la supervisión de la calidad de las instituciones de educación. Desde la década de los 90 existe una tendencia regional para que se concrete un diseño de aseguramiento de la calidad en el nivel superior. Por ello es de importancia el seguimiento de los egresados para afrontar las transformaciones requeridas por medio de la información [17].

Finalmente, el estudio cuyo objetivo es identificar los cambios que se necesitan implementar en el plan de estudios para el profesional de enfermería. El estudio analiza la información de egresados del periodo 2016-2019. El estudio descriptivo transversal, para una muestra de 94 egresadas, seleccionadas por muestreo por conveniencia. El cuestionario aplicado virtual incorpora variables como satisfacción del proceso formativo, empleo y desarrollo personal. Los resultados indican la necesidad de ampliar las horas prácticas 36%, los contenidos de salud ocupacional 19%, salud comunitaria 11% dentro de lo más destacado. La satisfacción relacionada con el docente/estudiante 94%, competencia de los docentes 90%. Las estudiantes

estas dispuestas a realizar una segunda especialización en Centro Quirúrgico, Emergencias y Desastres. La investigación concluye con la necesidad de fortalecer las experiencias en áreas como la asistencial y gestión para mejorar su desenvolvimiento en instituciones de salud de primer nivel de atención [18].

2.1.1.3 Antecedentes locales

La revisión bibliográfica en diversas fuentes y bases de datos no dan cuenta de estudios realizados a nivel local, por lo que se considera único dentro de la región Ica.

Finalmente, y para dar un enfoque sobre el tema de investigación. La minería de datos se define como la exploración y el análisis de datos con el objetivo principal de descubrir reglas significativas, patrones y tendencias futuras [19].

2.1.2 Bases Teóricas

2.1.2.1 Metodología CRISP-DM

Metodología CRISP-DM (Cross-Industry Standard Process for Data Mining), es un enfoque estándar utilizado para la minería de datos. La metodología CRISP-DM es iterativa y cada fase puede requerir volver a la fase anterior para mejorar el análisis o corregir problemas encontrados en las fases posteriores [20], [21]. Se compone de seis fases interconectadas, las mismas que se describen a continuación:

1. **Comprensión del problema:** el objetivo de esta fase es comprender los objetivos del proyecto y los requerimientos de negocio. Los pasos de esta fase son:
 - ✓ Determinar los objetivos empresariales y los criterios de éxito
 - ✓ Evaluar la situación (Identificar los requerimientos del negocio)
 - ✓ Determinar los objetivos de la minería de datos
2. **Comprensión de los datos:** el objetivo de esta fase es entender y familiarizarse con los datos disponibles para el proyecto. Los pasos de esta fase son:
 - ✓ Recolectar los datos
 - ✓ Describir los datos (explorar y visualizar los datos)
 - ✓ Explorar los datos (Identificar problemas de calidad de datos)
 - ✓ Verificar la calidad de los datos (Documentar los hallazgos)
3. **Preparación de los datos:** el objetivo de esta fase es preparar los datos para el análisis. Los pasos de esta fase son:
 - ✓ Seleccionar los datos relevantes
 - ✓ Limpiar y preprocesar los datos

- ✓ Integrar los datos de diferentes fuentes
 - ✓ Generar nuevas variables y características
4. Modelado: el objetivo de esta fase es desarrollar un modelo predictivo o descriptivo. Los pasos de esta fase son:
- ✓ Seleccionar las técnicas de modelado apropiadas
 - ✓ Generar el diseño de prueba (Crear y entrenar los modelos)
 - ✓ Construir el modelo
 - ✓ Evaluar los modelos y seleccionar el mejor
5. Evaluación: el objetivo de esta fase es evaluar el modelo construido y determinar si cumple con los objetivos del proyecto. Los pasos de esta fase son:
- ✓ Evaluar los resultados (el modelo con nuevos datos)
 - ✓ Evaluar el desempeño del modelo.
 - ✓ Refinar el modelo si es necesario.
6. Despliegue: el objetivo de esta fase es implementar el modelo en producción. Los pasos de esta fase son:
- ✓ Planificar el despliegue
 - ✓ Seguimiento y mantenimiento del plan (Implementar el modelo)
 - ✓ Monitorear y mantener el modelo
 - ✓ Elaborar informe final

2.1.3 Realidad problemática

La universidad nacional San Luis Gonzaga, después de haberse negado el licenciamiento institucional Octubre 2019, tuvo que suspender sus procesos de admisión, lo que ocasionó una baja en la cantidad de estudiantes por lo que la comunidad universitaria integrada por autoridades, docentes y personal administrativo han realizado un gran esfuerzo y una serie de actividades con la finalidad de levantar las observaciones impuestas por la SUNEDU y poder cumplir con las condiciones básicas de calidad (CBC), con el objetivo principal de obtener el licenciamiento, siendo una de estas condiciones el CONTROL Y SEGUIMIENTO AL EGRESADO. El ansiado licenciamiento institucional se pudo lograr en enero del 2022. Habiendo cumplido con las exigencias de las condiciones básicas de calidad (CBC), se hace de suma importancia continuar con este estándar del licenciamiento y que requiere de una especial atención es el seguimiento a los egresados, establecidos en la Condición VII Mecanismos de mediación e inserción laboral, en la que se establece la Oficina y el Plan de seguimiento al

graduado, se debe prestar mucha atención que la licencia tiene un periodo de vigencia y después de ese periodo la universidad debe ser nuevamente evaluada. Esta problemática, hace necesario que los mismos esfuerzos realizados para conseguir el licenciamiento, sean potenciados para ir en busca de la acreditación. La mejora continua es una necesidad institucional, la mejora en los procesos académicos son de suma importancia, pero esto solo es posible si se conoce cuál es el estado actual del seguimiento de los egresados con la finalidad de poder tomar decisiones en el rumbo o marca de la carrera profesional. En la investigación se hará el estudio en la facultad de Ingeniería Pesquera.

La facultad de Ingeniería Pesquera ubicada en la provincia de Pisco de la Región Ica es una facultad que cuenta con dos menciones: Ingeniería Pesquera e Ingeniería de Alimento. La información de los egresados recogida en fichas físicas (Anexo 01), no hacen posible tener un conocimiento sobre la condición de los egresados, motivo por el cual se viene registrando información de los últimos años en un archivo de Excel, con mucha data faltante, según se reporta en el área de seguimiento al egresado (Anexo 02).

La facultad de Ingeniería Pesquera cuenta con dos (02) escuelas académicas: La escuela académica de pesquería y la escuela académica de Alimentos. En las Tablas I y II, se muestra el histórico de estudiantes desde los ciclos 2017-2021.

TABLA I
ESTUDIANTES POR CICLO ACADÉMICO PERIODO 2017-2021 – PESQUERÍA

CICLO	1° CICLO	2° CICLO	3° CICLO	4° CICLO	5° CICLO	6° CICLO	7° CICLO	8° CICLO	9° CICLO	10° CICLO	N° ESTUDIANTES
2017-I	25	08	26	12	20	07	30	13	32	19	192
2017-II	24	23	07	25	13	20	08	29	12	32	193
2018-I	43	24	18	7	25	13	19	9	29	12	199
2018-II	0	27	25	15	8	24	12	18	9	29	167
2019-I	38	20	16	26	15	7	25	12	17	10	186
2019-II	0	48	0	17	25	14	9	25	13	18	169
2020-I	19	0	0	39	16	22	13	9	25	12	155
2020-II	0	23	0	39	0	16	23	15	6	26	148
2021-I	0	0	17	0	30	0	11	20	16	4	98
2021-II	0	0	0	15	0	30	0	12	19	17	93

En la Tabla I, la cantidad de estudiantes de la escuela de pesquería se ha ido reduciendo, producto del no licenciamiento de la universidad nacional San Luis Gonzaga.

TABLA II
ESTUDIANTES POR CICLO ACADÉMICO PERIODO 2017-2021 – PESQUERÍA

CICLO	1° CICLO	2° CICLO	3° CICLO	4° CICLO	5° CICLO	6° CICLO	7° CICLO	8° CICLO	9° CICLO	10° CICLO	N° ESTUDIANTES
2017-I	42	19	39	06	34	13	34	14	58	18	277
2017-II	33	40	19	37	07	33	13	34	13	59	288
2018-I	72	34	39	15	37	10	31	15	32	13	298
2018-II	0	63	35	33	18	34	10	30	15	31	269
2019-I	34	25	63	32	35	16	34	11	31	13	294
2019-II	0	59	0	53	32	34	19	31	9	31	268
2020-I	18	0	57	0	51	30	34	18	31	9	248
2020-II	0	12	0	46	0	48	26	33	19	29	213
2021-I	0	0	16	0	52	0	46	28	38	13	193
2021-II	0	0	0	18	0	47	0	42	27	39	173

Al igual que la escuela de pesquería, la escuela de alimentos igualmente ha ido reduciendo su población estudiantil, debido al no licenciamiento que no permitía el ingreso de nuevos estudiantes, y que recién para el semestre 2022-I ingresaran los nuevos estudiantes, para un plan de estudios aprobados con la licencia institucional.

Para el mismo periodo del 2017 – 2021, han egresado estudiantes de las dos escuelas académica y que se hacen necesario tener información del seguimiento de estos egresados. La Tabla III, muestra por cada ciclo académico los egresados.

TABLA III
EGRESADOS POR CADA CICLO ACADÉMICO Y ESCUELA PROFESIONAL

CICLO	N° EGRESADOS	
	PESQUERÍA	ALIMENTOS
2017-I	19	18
2017-II	32	59
2018-I	12	13
2018-II	29	31
2019-I	10	13
2019-II	18	31
2020-I	12	9
2020-II	26	29
2021-I	4	13
2021-II	17	39

2.1.4 Formulación del Problema

La problemática descrita párrafos precedentes, nos llevaron a plantearnos la siguiente interrogante

2.1.4.1 Problema General

PG: ¿Cómo pueden las técnicas de minería de datos mejorar el seguimiento y análisis del desempeño laboral y trayectoria profesional de los egresados de la Facultad de Ingeniería Pesquera de la Universidad Nacional San Luis Gonzaga durante el período 2017-2021?

2.1.4.2 Problemas Específicos

PE 1: ¿Qué técnicas de minería de datos (como clustering, clasificación, análisis de asociación, etc.) son más efectivas para extraer información significativa sobre la trayectoria profesional y laboral de los egresados de la Facultad de Ingeniería Pesquera?

PE 2: ¿Cuál es el impacto de la información obtenida a través de la minería de datos en la toma de decisiones y en la mejora de las estrategias de formación y apoyo a los egresados en la Facultad de Ingeniería Pesquera?

2.1.5 Justificación e importancia de la investigación

Con la finalidad de poder valor la justificación e importancia del estudio se tomara en consideración los criterios de la justificación establecidos en el libro de Metodología de la Investigación [22] que establece lo siguiente:

Conveniencia: resulta conveniente la investigación ya que, con su desarrollo, los resultados servirán de parámetros iniciales de las características de la información con que se cuenta de los egresados de la facultad de pesquería.

Relevancia social: en este criterio es de resaltar lo relevante del tema, porque sus resultados aportaran a la toma de decisiones en la mejora la gestión académica y por ende en la población estudiantil. Además de ser un aporte al indicador de calidad establecido en el nivel universitario como condición básica de calidad.

Valor teórico: que tiene la investigación es de suma importancia porque se aporta en el conocimiento de la aplicación de la minería de datos en el campo del seguimiento del egresado.

Utilidad metodológica: es importante tener en consideración que la metodología empleada para la gestión de los datos en minería de datos sentará precedentes para futuros estudios en esta línea de investigación.

En base a estos criterios establecidos la investigación en el presente proyecto es de suma importancia por lo cual se justifica su estudio.

2.2 Objetivos

2.2.1 Objetivo general

OG: Desarrollar e implementar técnicas de minería de datos para mejorar el seguimiento y análisis del desempeño laboral y trayectoria profesional de los egresados de la Facultad de Ingeniería Pesquera de la Universidad Nacional San Luis Gonzaga durante el período 2017-2021, con el fin de proporcionar información valiosa para la mejora de las estrategias de apoyo a los egresados.

2.2.2 Objetivos específicos

OE 1: Seleccionar y aplicar técnicas de minería de datos adecuadas, como clustering, clasificación y análisis de asociación, para extraer patrones y tendencias significativas en la trayectoria profesional y laboral de los egresados.

OE 2: Evaluar el impacto de la información obtenida a través de la minería de datos en la toma de decisiones institucionales y en la formulación de estrategias para mejorar el apoyo y la formación de los egresados, proponiendo recomendaciones para la implementación de cambios basados en los hallazgos.

2.3 Hipótesis y variables de la investigación

2.3.1 Hipótesis General

HG: La aplicación de técnicas de minería de datos permite obtener información valiosa y precisa sobre la trayectoria profesional y laboral de los egresados de la Facultad de Ingeniería Pesquera de la Universidad Nacional San Luis Gonzaga durante el período 2017-2021, mejorando así el seguimiento y las estrategias de apoyo a los egresados.

2.3.2 Hipótesis Especificas

HE 1: Las técnicas de minería de datos, como clustering y clasificación, son capaces de identificar patrones y segmentos significativos en la trayectoria profesional de los egresados, permitiendo un análisis más detallado de su desempeño y evolución laboral.

HE 2: La información obtenida a través de la minería de datos tendrá un impacto positivo en la toma de decisiones y en la formulación de estrategias para mejorar el apoyo y la formación

de los egresados, contribuyendo al desarrollo profesional continuo y al ajuste de los programas educativos.

2.3.3 Variables de la investigación

2.3.4 Variable Independiente

Técnicas de Minería de Datos

2.3.5 Variable Dependiente

Información sobre la trayectoria profesional y mejora en el seguimiento y estrategias de Apoyo.

2.3.6 Operacionalización de las variables

Técnicas de Minería de Datos (Variable Independiente)

Indicadores:

- Tipos de técnicas aplicadas (clustering, clasificación, análisis de asociación, etc.).
- Herramientas y software utilizados
- Calidad y precisión de los modelos (errores de clasificación, precisión, recall).

Información sobre la Trayectoria Profesional (Variable Dependiente)

Indicadores:

- Empleabilidad (porcentaje de egresados empleados en su campo de estudio).
- Tipo de empleo (categorías de empleo: técnico, profesional, administrativo, etc.).
- Nivel de satisfacción laboral (medido mediante encuestas o cuestionarios).
- Continuación de estudios (porcentaje de egresados que realizan estudios de posgrado).

Mejora en el Seguimiento y Estrategias de Apoyo (Variable Dependiente)

Indicadores:

- Implementación de nuevas estrategias o políticas basadas en los datos (documentación de cambios en políticas o programas).
- Evaluación de la efectividad de las nuevas estrategias (resultados de encuestas o entrevistas con egresados y personal académico).

III. ESTRATEGIA METODOLÓGICA

3.1 Estrategia de la investigación

La estrategia de la investigación se desarrolló con una investigación aplicada tecnológica en la que aplicó los conocimientos sobre Minería de Datos, aplicando algoritmos no supervisados como K-means, clustering, utilizando la herramienta de minería de datos de uso libre Orange data mining. El nivel de la investigación se da en el nivel aplicado, con un diseño observacional, no experimental, de corte transversal para el periodo de egresados entre el 2017 al 2021, por la naturaleza de los datos almacenados en la oficina de estadística de la facultad de Ingeniería pesquera en sus dos escuelas académicas (pesquería y alimentos) siendo este diseño retrospectivo ya que los datos han sido registrados en base de datos en Excel.

La población estuvo conformada por todos los egresados entre el periodo de estudio 2017-2021. Siendo su unidad de análisis egresados de la facultad de Ingeniería Pesquera de la universidad nacional San Luis Gonzaga. La población estuvo conformada por 173 egresados de la escuela de pesquería y 249 egresados de la escuela de alimentos en total se tienen 422 egresados, la muestra que se va a procesar, es la misma que la población, por lo que no requiere de cálculo de la muestra.

3.2 Método de investigación

El proceso para la investigación se basará en la metodología CRISP-DM [23], ya que esta metodología desempeña un papel importante en la gestión de datos. Esta metodología incluye 6 fases, la misma que se esquematiza en la Fig. 1.

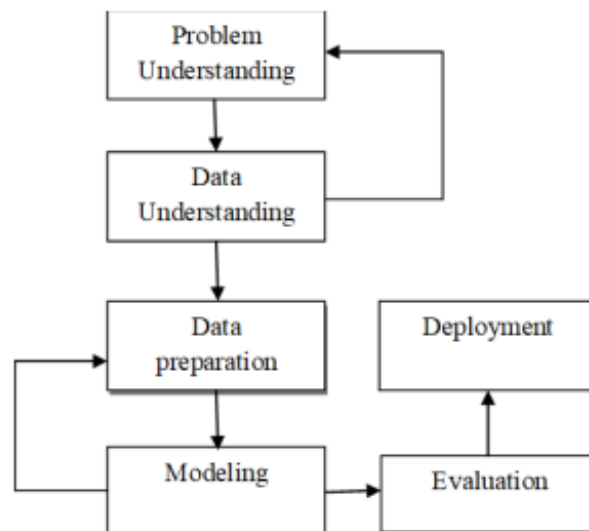


Fig. 1. CRISP-DM Methodology Process[23]

3.3 Aplicación del método de investigación

3.3.1 Comprensión del problema

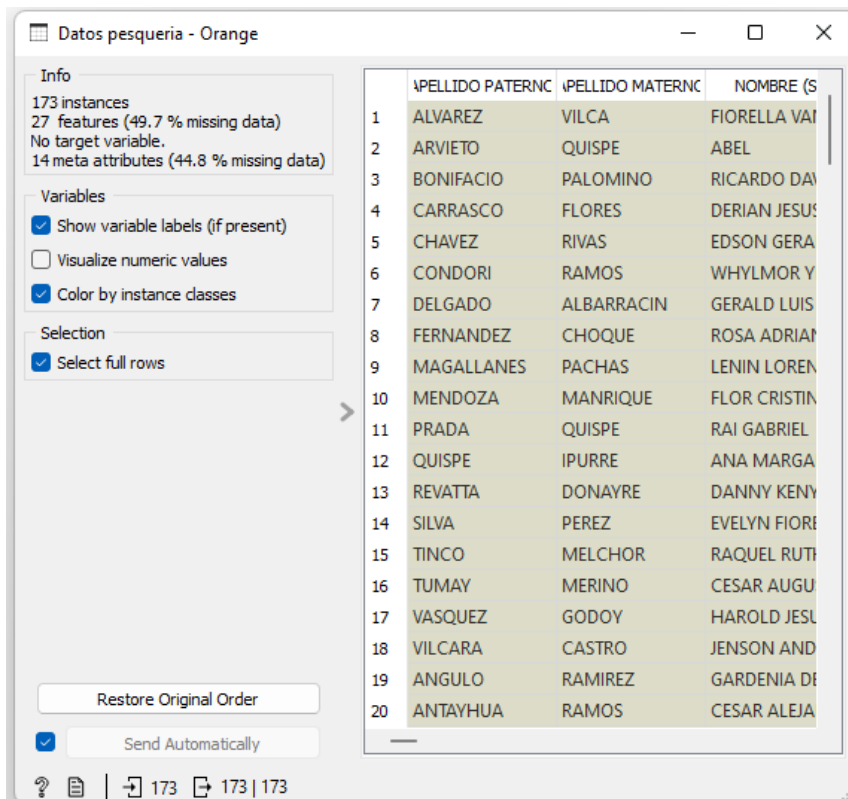
1. Determinar los objetivos empresariales y los criterios de éxito
2. Evaluar la situación (Identificar los requerimientos del negocio)

De los datos visualizados se desprende que existe datos registrados en un archivo de Excel, datos que se han obtenido de las fichas de registro de los egresados de la facultad de Ingeniería Pesquera y Alimentos. Sin embargo, no es posible tener un conocimiento sobre los egresados, tales como, por ejemplo:

1. Años transcurridos para obtener el bachiller desde que egreso
 2. Años transcurridos para obtener el título profesional
 3. Cantidad de estudiantes que han obtenido el título profesional
 4. Edad promedio de los egresados?
 5. Conocer que egresados están trabajando sobre todo en la especialidad
 6. Cantidad de egresados que se encuentran laborando, etc.
3. Determinar los objetivos de la minería de datos

Los objetivos que se persiguen con la minería de datos, es evaluar los datos registrados de los egresados del programa de ingeniería pesquera, diferenciadas en sus dos escuelas académicas: Ingeniería pesquera e Ingeniería de alimentos.

Igualmente, la evaluación debe buscar patrones de comportamiento en los resultados que es lo que se espera obtener.



	APELLIDO PATERNC	APELLIDO MATERNC	NOMBRE (S)
1	ALVAREZ	VILCA	IORELLA VAI
2	ARVIETO	QUISPE	ABEL
3	BONIFACIO	PALOMINO	RICARDO DA
4	CARRASCO	FLORES	DERIAN JESU
5	CHAVEZ	RIVAS	EDSON GERA
6	CONDORI	RAMOS	WHYLMOR Y
7	DELGADO	ALBARRACIN	GERALD LUIS
8	FERNANDEZ	CHOQUE	ROSA ADRIAN
9	MAGALLANES	PACHAS	LENIN LOREN
10	MENDOZA	MANRIQUE	FLOR CRISTIN
11	PRADA	QUISPE	RAI GABRIEL
12	QUISPE	IPURRE	ANA MARGA
13	REVATTA	DONAYRE	DANNY KENY
14	SILVA	PEREZ	EVERLYN FIORI
15	TINCO	MELCHOR	RAQUEL RUTI
16	TUMAY	MERINO	CESAR AUGU
17	VASQUEZ	GODOY	HAROLD JESU
18	VILCARA	CASTRO	JENSON AND
19	ANGULO	RAMIREZ	GARDENIA DE
20	ANTAYHUA	RAMOS	CESAR ALEJA

Fig. 2. Datos de la escuela de pesquería

Un primer acercamiento a los datos registrados en la escuela de Ingeniería Pesquera, muestran que existen 173 registros (instancias) registradas, con 27 características (atributos). De estos atributos o características se tiene que existen 49,7% de datos faltantes en algunas de estas características. Además, se cuenta con 14 atributos de meta data en las que se tiene 44,8% de datos faltantes o perdidos Fig.2.

	APELLIDO PATERNC	APELLIDO MATERNC	NOMBRE (S)
1	AGÜELA	ALCALA	LISSETH NOEM
2	AÑAMACO	ABARCA	RAISA
3	AVALOS	AVALOS	MARIA GUADA
4	CARRASCO	MENDIETA	AMELIA CAROI
5	CONISLLA	DAMIAN	HENRY JESUS
6	CRISOSTOMO	TINCO	ROCIO RAQUEL
7	DE LA CRUZ	HUALLANCA	ELIZABETH DEL
8	DE LA CRUZ	VILCA	LUZ KARINA
9	FELIPA	SANDOVAL	ALEXANDRA M
10	GALLEGOS	QUISPE	TESLY ALENY S
11	GARCIA	MUÑOZ	SHEMIRA YISBE
12	HUACCAMAITA	CONTRERAS	MARICRUZ
13	HUASASQUICHE	LOPEZ	JESUS MANUEL
14	IPURRE	TUEROS	ELIZABETH CRI
15	MATTA	HERENCIA	NOELIA ALEXA
16	PEÑA	MARTINEZ	ROXANA LISSE
17	PEVE	QUISPE	YADIMIR
18	SIGUAS	MOSQUERA	OLGA MELISSA
19	VILCA	TASAYCO	OMAR ANTONI
20	AJALCRIÑA	CABRERA	VICTOR JEANPI
21	AQUIJE	ASCONA	PATRIK KEVIN
22	ARAUJO	MANTARI	LUZ ADEL

Fig. 3. Datos de la escuela de alimentos

En la Fig. 3, similarmente, un primer acercamiento a los datos registrados en la escuela de Alimentos, muestran que existen 249 registros (instancias) registradas, con 31 características (atributos). De estos atributos o características se tiene que existen 54,6% de datos faltantes en algunas de estas características. Además, se cuenta con 11 atributos de meta data en las que se tiene 41,3% de datos faltantes o perdidos.

De las Fig. 2 y 3, se tiene que se cuenta con 422 registro de egresados y que serán analizados con las técnicas de minería de datos.

3.3.2 Comprensión de los datos

1. Recolectar los datos

Los datos ya han sido recolectados, por medio del registro de las fichas de egresados las mismas que se encuentran en un archivo de Excel.

Nº	CICLO Y AÑO DEL REGISTRO	COD. ALUMNO	APELLIDO PATERNO	APELLIDO MATERNO	NOMBRE (S)	ESCUELA PROFESIONAL	SEDE	FECHA NACIMIENTO	DIRECCION RESIDENCIA	TELEFONO / CELULAR
1										
2	2017-I	20122297	AGUIA	ALCALA	LISSETH NOEMI	INGENIERIA DE ALIMENTOS	PISCO	26/6/1994	CONTROL PORBLADO DE HERBA ALTO CAÑETE	954241681
3	2017-I	20114460	ASANACO	ABARCA	RAISA	INGENIERIA DE ALIMENTOS	PISCO	13/2/1992	ASOC MICAELA BASTIDAS MZ C LOT 5- VILLA T. PISCO	991180129
4	2017-I	20112613	AVALES	AVALES	MARIA GUADALUPE	INGENIERIA DE ALIMENTOS	PISCO	6/6/1994	CP EL PALOMAR MZ A LT 15	972205862
5	2017-I	20122302	CARRASCO	MENDETA	AMELIA CAROLINA	INGENIERIA DE ALIMENTOS	PISCO	2/7/1996	C.P CASA BLANCA MZ L LT 2 CERRO AZUL CAÑETE	902790380
6	2017-I	20122305	CONTELA	DAMIAN	HENRY JESUS	INGENIERIA DE ALIMENTOS	PISCO	27/12/1994	AV ISAC MONTALVAN S/N INDEPENDENCIA - PISCO	956573272
7	2017-I	20122962	CHRISTOTOMO	TINCO	ROCCO RAQUEL	INGENIERIA DE ALIMENTOS	PISCO	29/4/1995	J. P. JUNCHAYA S/N HUMAY PISCO	920536375
8	2017-I	20121172	DE LA CRUZ	HUALLANCA	ELIZABETH DEL ROSARIO	INGENIERIA DE ALIMENTOS	PISCO	17/1/1995	AV EL PORVENIR S/N SUNAMPE	985502772
9	2017-I	20122306	DE LA CRUZ	VILCA	LUZ KARINA	INGENIERIA DE ALIMENTOS	PISCO	14/2/1995	CP MIGUEL GRAU MZ B LT 8 CHINCHA BAJA	914879611
10	2017-I	20122308	PEÑA	SANNOVAL	ALEXANDRA MARITA	INGENIERIA DE ALIMENTOS	PISCO	6/7/1995	C. VILLA LEON MZ B LT 4 TAMBO DE MORA - CHINCHA	956572421
11	2017-I	20121174	GALEGOS	QUISEP	TESLY ALENY SYLBANIA	INGENIERIA DE ALIMENTOS	PISCO	27/11/1995	CALLE SAN AMRCCOS S/N CHAVIN - CHINCHA	954669084
12	2017-I	20121175	GAICIA	MAURIZO	SHEMIRA YSABEL	INGENIERIA DE ALIMENTOS	PISCO	24/5/1995	AV MIRAFLORES N88 SAN CLEMENTE	978572097
13	2017-I	20122309	HUACAMANTA	CONTEGAS	MARICRIZ	INGENIERIA DE ALIMENTOS	PISCO	29/7/1995	AAHN EL SALVADOR M2 L1 L1 PUEBLO NUEVO - CHINCHA	930604476
14	2017-I	20121177	IPURÉ	TUEROS	ELIZABETH CRISTINA	INGENIERIA DE ALIMENTOS	PISCO	29/4/1995	CALLE ABANCAY 505 SAN CLEMENTE - PISCO	931571379
15	2017-I	20121178	MATTA	HERENCIA	NOELIA ALEXANDRA	INGENIERIA DE ALIMENTOS	PISCO	10/7/1995	URB LA ALBORADA MZ X LT 1 - PISCO	968207834
16	2017-I	20114461	PEÑA	MARTINEZ	ROKANA LISSET	INGENIERIA DE ALIMENTOS	PISCO	3/5/1992	CARRETA DE TORO LATERAL 3 LT 29 - PISCO	959665869
17	2017-I	20122313	PEVE	QUISEP	YADIMIR	INGENIERIA DE ALIMENTOS	PISCO	28/8/1994	AAHN ALBERTO TATAJE MZ B LT 7	959137811
18	2017-II	20101838	SIGUAS	MOSQUERA	OLGA MELISSA	INGENIERIA DE ALIMENTOS	PISCO	22/9/1992	PASAJE LOS LENOS N 190	315031
19	2017-II	20121183	VILCA	TASATCO	OMAR ANTONIO	INGENIERIA DE ALIMENTOS	PISCO	18/4/1995	AV EMILIO ROMAN SARAVIA	945485077
21	2017-II	20120533	VALCARRERA	CABRERA	VICTOR JEANPIERE	INGENIERIA DE ALIMENTOS	PISCO	05/6/1995	P. JUVEN SAN MARTIN DE PORRAS B-1 ICA	953589351
22	2017-II	20123560	AQUIE	ASCONA	PATRIK KEVIN	INGENIERIA DE ALIMENTOS	PISCO	4/6/1995	CALLE MIRAFLORES N 401 SAN ANDRES	959804448

Fig. 4. Registro de datos de egresados1

N	EMAIL	SEXO	ESTADO CIVIL	UNI	FECHA EGRESO	FECHA DE INGRESO	GRADOS OBTENIDOS BACHILLER	ESPECIALIDAD/ MENCION	FECHA DE BACHILLER	GRADOS OBTENIDO TITULO	ESPECIALIDAD/ MENCION	FECHA DE TITULO	GRADOS OBTENIDO MAESTRIA/DOCTORADO	ESPECIALIDAD/ MENCION	FECHA DE MAESTRIA/DOCTORADO
1															
2		F	S	48289318	16/8/2017	26/8/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	05/10/2018						
3		F	S	48814478	16/8/2017	30/12/2011	BACHILLER	CIENCIA DE LOS ALIMENTOS	31/10/2018						
4	shariupe_88@hotmail.com	F	S	70171880	16/8/2017	17/7/2011	BACHILLER	CIENCIA DE LOS ALIMENTOS	28/05/2018						
5	carol_02car@hotmail.com	F	S	71521436	16/8/2017	26/8/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	05/10/2018	TITULO	INGENIERO DE ALI	22/12/2021			
6		M	S	75728084	16/8/2017	26/8/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	28/05/2018						
7		F	S	78179704	16/8/2017	26/8/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	28/05/2018						
8		F	S	75728280	16/8/2017	15/7/2012	EN TRAMITE	CIENCIA DE LOS ALIMENTOS							
9		F	S	48482411	16/8/2017	23/8/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	04/02/2019						
10		F	S	70400790	16/8/2017	26/8/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	28/05/2018	TITULO	INGENIERO DE ALI	12/07/2022			
11	luzgalego@gmail.com	F	S	78029924	12/12/2017	15/7/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	28/05/2018	TITULO	INGENIERO DE ALI	12/07/2022			
12	sthemmy24@hotmail.com	F	S	71131026	16/8/2017	15/7/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	28/05/2018	TITULO	INGENIERO DE ALI	15/02/2022			
13		M	S	11111108	16/8/2017	26/8/2012	EN TRAMITE	CIENCIA DE LOS ALIMENTOS	05/10/2018						
14		F	S	70203701	16/8/2017	15/7/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	17/05/2018						
15	shematta200@gmail.com	F	S	70140704	16/8/2017	15/7/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	21/05/2018	TITULO	INGENIERO DE ALI	12/07/2022			
16		F	S	70682347	16/8/2017	30/12/2011	BACHILLER	CIENCIA DE LOS ALIMENTOS	04/02/2019						
17		M	S	70281750	16/8/2017	26/8/2012	BACHILLER	CIENCIA DE LOS ALIMENTOS	28/05/2018	TITULO	INGENIERO DE ALI	12/07/2022			
18	yadmir_povequispe@gmail.com	M	S	72814754	16/8/2017	19/12/2015	BACHILLER	CIENCIA DE LOS ALIMENTOS	17/05/2018						
19		M	S	70722537	16/8/2017	15/7/2011	BACHILLER	CIENCIA DE LOS ALIMENTOS	28/05/2018				ESTUDIANTE	TECNOLOGIA DE ALIMENTOS EN LA AGRIC	
20															

Fig. 5. Registro de datos de egresados2

AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK
FECHA DE INGRESO 1er TRABAJ	FECHA DE CULMINACION 1er TRAB	LUGAR DE TRABAJO	DIRECCION DEL TRABAJO (DIRECCION, DISTRITO, PROVINCIA, REGION)			TIEMPO DURACION 1er TRABAJ	FECHA DE INGRESO ULTIMO TRABAJ	FECHA DE CULMINACION ULTIMO TRABAJ	EMPRESA Y/O INSTITUCION DE TRABAJO	DIRECCION DEL TRABAJO (DIRECCION, DISTRITO, PROVINCIA, REGION)
1										
2	2018		GYNARA	PROCESO DE ALCAHOFA	AKULLAR DE CALIDAD					
3										
4	2018		COMPLEJO AGRONINDUSTRIAL BETA	CONTROL DE CALIDAD	INSPECTOR DE ASESURAMIENTO DE LA CALIDAD		2018		BUREAU VERITAS	TAMBO DE MORA
5	2018		CERRO VERDE SAC		SUPERVISORA DE CALIDAD					
6										
7	2018		AGRICOLA HOJA REDONDA	EL CARMEN - CHINCHA	CONTROL DE CALIDAD				INSPECTOR DE CALIDAD	
8										
9										
10										
11			NO LABORAR							
12			NO LABORAR							
13										
14	2019		AGRICOLA CHAPI	CONTROL DE CALIDAD	INSPECTOR DE CALIDAD					
15										
16	2018		PRODUMIL	ANALISTA DE CALIDAD						
17	2018		FRUCHINCHA	ASEGURAMIENTO DE CALIDAD						
18	2018		AGRICOLA FONDO CALIFORNIA	CONTROL DE CALIDAD						
19	2018		PESQUERA DIAMANTE	PISCO	ANALISTA			2021	RVR AGRO	ICA
20										

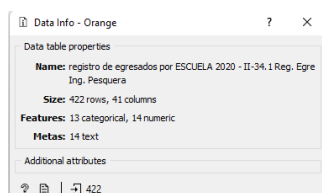
Fig. 6. Registro de datos de egresados3

	AL	AM	AN
	AREA DE TRABAJO	CARGO QUE DESEMPEÑA	TIEMPO DURACION ULTIMO TRABAJO
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18	CALIDAD	CALIDAD	
19			
20			

Fig. 7. Registro de datos de egresados4

Las Figuras de la 4 a la 7, muestran la información registrada en el archivo de Excel, en la que como se puede visualizar existe muy poca información sobre el tema laboral, además de mucha información que no aporta al conocimiento y que se desprenderán de las siguientes fases de la minería de datos.

2. Describir los datos (explorar y visualizar los datos)



	PELLIDO PATERNO	PELLIDO MATERNO	NOMBRE (S)	ECHA NACIMIENT	ECCEJÓN RESIDENC	E-MAIL
1	ALVAREZ	VILCA	FIORILLA VAN...	1993-09-08 00:00...	AAHH MIGUEL ...	?
2	ARIVITO	QUISPE	ABEL	1993-06-19 00:00...	SANTA CRUZ ...	?
3	BONIFACIO	PALOMINO	RICARDO DAVID	1994-12-01 00:00...	URB LEON DE V...	ricardo_bp94@...
4	CARRASCO	FLORES	DERIAN JESUS	1995-08-30 00:00...	RESIDENCIAL R...	?
5	CHAVEZ	RIVAS	EDSON GERAR...	1989-05-08 00:00...	AH PACHINGA	edsongch19@g...
6	CONDORI	RAMOS	WHYLMOR VEY...	1993-09-11 00:00...	CESAR VALLEJO...	?
7	DELGADO	ALBARRACIN	GERALD LUIS	1994-03-21 00:00...	PISCO - PISCO ...	geluis987@gm...
8	FERNANDEZ	CHOQUE	ROSA ADRIANA	1990-08-29 00:00...	SAN CLEMENT...	?
9	MAGALLANES	PACHAS	LENIN LORENZO	1995-07-14 00:00...	CRUZ BLANCA...	?
10	MENDOZA	MANRIQUE	FLOR CRISTINA	1994-12-19 00:00...	JIRON TAHUAN...	?
11	PRADA	QUISPE	RAI GABRIEL	1994-12-15 00:00...	RAMON ASPILL...	?
12	QUISPE	IPURRE	ANA MARGARITA	1993-10-19 00:00...	JORGE CHAVEZ...	annequispeipur...
13	REVATTA	DONAYRE	DANNY KENNY	2017-08-08 00:00...	LOTZACION D...	?
14	SILVA	PEREZ	EVELYN FIOREL...	1995-09-22 00:00...	AAHH SEÑOR ...	ebi_fiore_22@h...
15	TINCO	MELCHOR	RAQUEL RUTH	1992-06-08 00:00...	CALLE CAPULI ...	?
16	TUMAY	MERINO	CESAR AUGUSTO	1992-01-13 00:00...	PASAJE LOS CL...	?
17	VASQUEZ	GODOY	HAROLD JESUS	1994-08-28 00:00...	CALLE MANJUE...	?
18	VILCARA	CASTRO	JENSON ANDRES	1992-11-13 00:00...	CENTRO POBL...	?
19	ANGULO	RAMIREZ	GARDENIA DEL ...	1995-12-31 00:00...	AV SAN MARTI...	?
20	ANTAYHUIA	RAMOS	CESAR ALEJAN...	1995-05-17 00:00...	F. LEON N° 340...	?
21	ANYOSA	CHAMPE	CARLOS ANGEL	1995-12-05 00:00...	CALLE JOSE OL...	carlosanyosa_5...
22	APOLAYA	MENDEZ	WILMER EDWIN	1996-03-22 00:00...	CENTRO POBL...	wilmerapolaya...

Fig. 8. Información de los datos

Fig. 9. Información detallada de los datos.

De la información procesada y mostrada en la Fig. 8 se especifica que se cuenta con 422 registros de egresados, existiendo 41 columnas o características o atributos. De las 41 características 13 tienen datos categóricos (texto), 14 tienen datos numéricos y 14 presentan datos de tipo Metadato.

Ampliando en la información de la Fig. 9, existe en el registro de Excel de los 422 Instancias (registros de datos), en las que, de las 27 características o atributos, se presentan 43.3% de datos faltantes o perdidos. Igualmente, de los 14 meta atributos se presentan 42.2% de datos perdidos o faltantes.

3. Explorar los datos (Identificar problemas de calidad de datos)

	PELLIDO PATERNO	PELLIDO MATERNO	NOMBRE (S)	FECHA NACIMIENTO	DIRECCIÓN RESIDENC	E-MAIL	LUGAR DE TRABAJO	DIRECCION, DISTRIT	AREA DE TRABAJO (CARGO QUE DESEMPEÑA	INSTITUCION DE
1	ALVAREZ	VILCA	FIGRELLA VANE...	1993-09-06 00:0...	AAHH MIGUEL ...	?	NO LABORA	?	?	?	?
2	ARVIETO	QUISPE	ABEL	1993-06-19 00:0...	SANTA CRUZ - ...	?	NO LABORA	?	?	?	?
3	BONIFACIO	PALOMINO	RICARDO DAVID	1994-12-01 00:0...	URB LEON DE V...	ricardo_bp94@...	NO LABORA	?	?	?	?
4	CARRASCO	FLORES	DERIAN JESUS	1995-08-30 00:0...	RESIDENCIAL R...	?	NO LABORA	?	?	?	?
5	CHAVEZ	RIVAS	EDSON GERAR...	1989-05-08 00:0...	AH PACHINGA ...	edsongch19@g...	Empresa austral	?	analista de labo...	Supervisor de C...	?
6	CONDORI	RAMOS	WHYLMOR YEY...	1993-09-11 00:0...	CESAR VALLEJO...	?	EMPRESA PRIS...	PISCO	?	SUPERVISOR	AGROVICTORIA...
7	DELGADO	ALBARRACIN	GERALD LUIS	1994-03-21 00:0...	PISCO - PISCO ...	geluis987@gm...	MAR ANDINO ...	CERRO DE PAS...	ACUICULTURA	SUPERVISOR	?
8	FERNANDEZ	CHOQUE	ROSA ADRIANA	1990-08-29 00:0...	SAN CLEMENT...	?	NO LABORA	?	?	?	?
9	MAGALLANES	PACHAS	LENIN LORENZO	1995-07-14 00:0...	CRUZ BLANCA ...	?	NO LABORA	?	?	?	?
10	MENDOZA	MANRIQUE	FLOR CRISTINA	1994-12-19 00:0...	JIRON TAHUAN...	?	DIRECCION RE...	ICA	MEDIO AMBIEN...	?	?
11	PRADA	QUISPE	RAI GABRIEL	1994-12-15 00:0...	RAMON ASPILL...	?	EMPRESA ALIM...	ICA	ALGAR CULTIV...	LABORATORISTA	?
12	QUISPE	IPURRE	ANA MARGARITA	1993-10-19 00:0...	JORGE CHAVEZ...	annaquispeipur...	CORPORACIO...	AV FELIX CARRI...	ASEGURAMIENT...	SUPERVISOR ...	?
13	REVATTA	DONAYRE	DANNY KENYI	2017-08-08 00:0...	LOTIZACION D...	?	EMPRESA PRIS...	PISCO	?	SUPERVISOR D...	?
14	SILVA	PEREZ	EVELYN FIGRELLA...	1995-09-22 00:0...	AAHH SEÑOR ...	ebi_fiore_22@h...	NO LABORA	?	?	?	?
15	TINCO	MELCHOR	RAQUEL RUTH	1992-06-08 00:0...	CALLE CAPULI ...	?	NO LABORA	?	?	?	?
16	TUMAY	MERINO	CESAR AUGUSTO	1992-01-13 00:0...	PASAJE LOS CL...	?	DIAMANTE	?	?	OPERARIO DE T...	TRANSMARIANA
17	VASQUEZ	GODOY	HAROLD JESUS	1994-08-28 00:0...	CALLE MANUE...	?	NO LABORA	?	?	?	?
18	VILCARA	CASTRO	JENSON ANDRES	1992-11-13 00:0...	CENTRO POBL...	?	EMPRESA PRIS...	PISCO	PRODUCCION D...	?	?
19	ANGULO	RAMIREZ	GARDENIA DEL ...	1995-12-31 00:0...	AV SAN MARTI...	?	EMPRESA INFR...	?	?	?	?
20	ANTAYHUA	RAMOS	CESAR ALEJAN...	1995-05-17 00:0...	F. LEON N° 340...	?	FONCODES	?	?	SUPERVISOR	?
21	ANYOSA	CHAMPE	CARLOS ANGEL	1995-12-05 00:0...	CALLE JOSE OL...	carlosanyosa_5...	TRANSMARINA...	PARACAS	CONGELADOS	INSPECTOR ASE...	?
22	APOLAYA	MENDEZ	WILMER EDWIN	1996-03-22 00:0...	CENTRO POBL...	wilmerapolaya...	INVERSIONES P...	?	?	SUPERVISOR D...	?
23	ARONE	ALVITES	LISBETH YOSELIN	1995-10-22 00:0...	ASOC SAN PED...	?	?	?	?	?	?
24	ATUNCAR	PACHAS	JAVRO	1996-10-23 00:0...	AV MIGUEL GR...	atuncarpachasi...	EMPRESA: E.G ...	?	?	ADMINISTRATIVO	?
25	AURIS	SANCHEZ	FERNANDO	1994-04-01 00:0...	JR ARICA 181 C...	?	SAN FERNANDO	?	?	?	?
26	BULEJE	HUAMANI	DIANA KATHER...	1994-03-22 00:0...	AYACUCHO - S...	?	IQF S.A	PUEBLO NUEVO	BPM	CONTROL	?
27	CAMPOS	HERNANDEZ	CARMEN LILIBE...	1995-11-23 00:0...	CALLE LA ESPE...	?	?	?	?	?	?
28	CARRASCO	HUAMANI	ROSALINDA	1996-06-01 00:0...	28 DE JULIO MZ...	roze.carrasco@...	NO LABORA	?	?	?	?
29	CASTILLO	RIVERA	LUIS YUBAL	1992-06-09 00:0...	VISTA ALEGRE L...	?	MOVISTAR	?	VENTAS	?	?
30	CHAVEZ	FALCON	SANDRA MIREL...	1995-10-16 00:0...	PPIJ SAN MIGU...	?	NO LABORA	?	?	?	?

Fig. 10. Datos faltantes en los registros de los egresados

Como se puede comprobar en la Fig. 10, donde se muestran los datos, existen una cantidad de datos faltantes muy grande y que de una u otra manera deberán ser tratados o eliminados ya que no son posible su tratamiento. Sin embargo, la mayoría de esos datos faltantes no son relevantes para una analítica de estudio como por ejemplo los correos electrónicos, la dirección del egresado, otros si de importancia que están vacíos como área de trabajo (aunque no todos trabajan), cargo que desempeña, etc.

4. Verificar la calidad de los datos (Documentar los hallazgos)

	FECHA DE TITULO	ESTRA/2DA. ESPEC	IALIDAD/ MENCIO	IA/2DA ESPECIALI	DE INGRESO 1er TR	IE CULMINACION 1	DURACION 1er TR	INGRESO ULTIMO	CULMINACION ULT	URACION ULTIMO	Feature 1
1	2019-02-04 00:00...	?	?	?	?	?	?	?	?	?	?
2	?	?	?	?	?	?	?	?	?	?	?
3	2021-08-23 00:00...	?	?	?	?	?	?	?	?	?	?
4	?	?	?	?	?	?	?	?	?	?	?
5	2022-02-15 00:00...	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
6	?	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
7	2022-10-18 00:00...	?	?	?	?	2018-09-17 00:00...	?	?	?	?	?
8	2021-08-23 00:00...	?	?	?	?	?	?	?	?	?	?
9	?	?	?	?	?	?	?	?	?	?	?
10	2020-06-26 00:00...	?	?	?	?	2019-01-01 00:00...	?	?	?	?	?
11	?	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
12	?	?	?	?	?	2019-01-15 00:00...	?	?	?	?	?
13	?	?	?	?	?	2019-01-01 00:00...	?	?	?	?	?
14	2021-08-23 00:00...	?	?	?	?	?	?	?	?	?	?
15	?	?	?	?	?	?	?	2019-01-01 00:00...	?	?	?
16	?	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
17	2020-09-30 00:00...	?	?	?	?	?	?	?	?	?	?
18	?	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
19	2022-04-28 00:00...	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
20	2021-09-22 00:00...	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
21	2021-09-22 00:00...	?	?	?	?	2019-10-15 00:00...	?	?	?	?	?
22	2022-07-12 00:00...	?	?	?	?	2019-01-01 00:00...	?	?	?	?	?
23	?	?	?	?	?	?	?	?	?	?	?
24	2022-04-28 00:00...	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
25	?	?	?	?	?	2018-01-01 00:00...	?	?	?	?	?
26	?	?	?	?	?	2019-04-21 00:00...	?	?	?	?	?
27	?	?	?	?	?	?	?	?	?	?	?
28	?	?	?	?	?	?	?	?	?	?	?
29	?	?	?	?	?	2019-01-01 00:00...	?	?	?	?	?
30	?	?	?	?	?	?	?	?	?	?	?

Fig. 11. Calidad de los datos

Se ha encontrado por medio de la herramienta de minería de dato, que existen mucha información no registrada lo cual limita las posibilidades de realizar un tratamiento con estas características, y en su mayoría corresponden a la información laboral que pudo ser registrada siendo egresado, bachiller o titulado.

3.3.3 Preparación de los datos

1. Seleccionar los datos relevantes

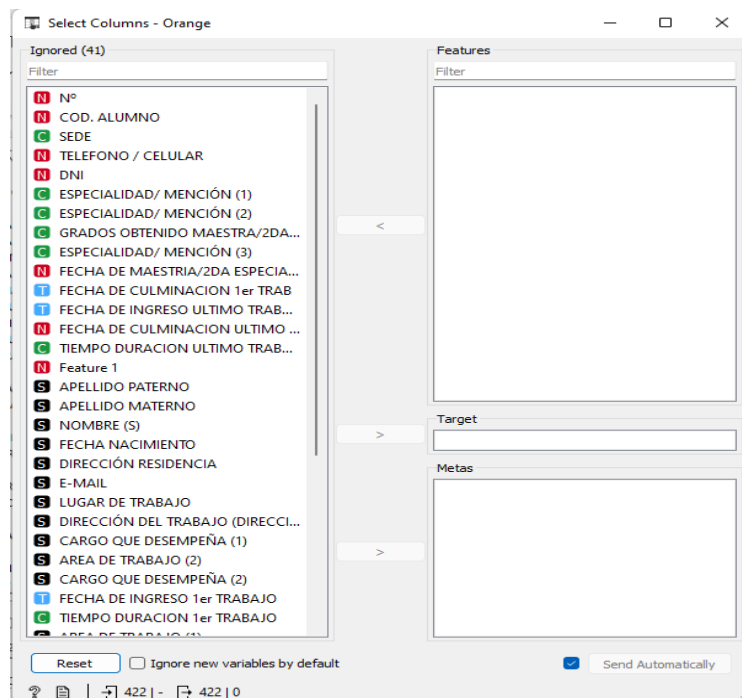


Fig. 12. Datos del Dataset

En el widget que permite seleccionar los datos, podemos apreciar en el lado izquierdo se muestran los datos que dispone el Dataset, en la parte superior derecha los datos que queremos seleccionar (Features), debajo el Target para minería de datos para conocer la variable para algoritmos supervisados, y en la parte superior los meta atributos.

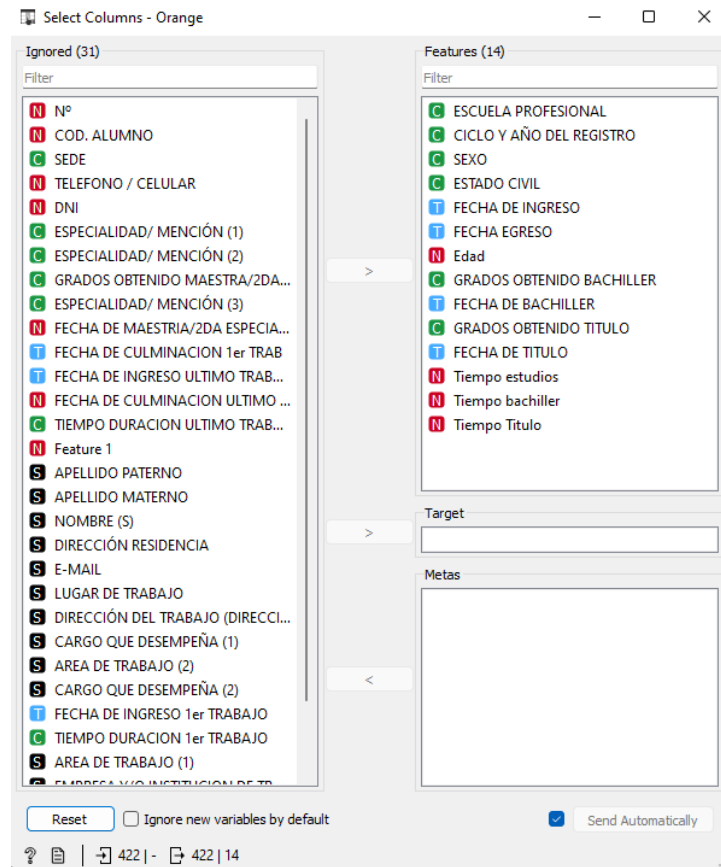


Fig. 13. Datos seleccionados

En la Fig. 13, se ha seleccionado los datos que tienen relevancia para el estudio, como se puede apreciar existen muchos datos, que lamentablemente no son de interés debido a que en la mayoría de los casos no tienen información, están incompletos existe muchos datos perdidos.

De la Fig.13 se desprende que se han seleccionado solo 14 características o atributos de los 44 existentes.

2. Limpiar y preprocesar los datos

Para realizar la limpieza de los datos debemos conocer los datos y sus características como lo que se presenta en la Fig. 14.

	FECHA NACIMIENT	CUOLA PROFESION	O Y AÑO DEL REGIS	SEXO	ESTADO CIVIL	FECHA DE INGRES	FECHA EGRESO	S OBTENIDO BACH	FECHA DE BACHILLE	S OBTENIDO TIT	FECHA DE TITULO
1	1993-09-06 00:00:00...	INGENIERIA PE...	2017 - I	F	S	2012-07-15 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-01-31 00:00:00...	TITULO	2019-02-04 00:00:00...
2	1993-06-19 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2012-08-26 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-05-21 00:00:00...	SIN TITULO	?
3	1994-12-01 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2012-07-15 00:00:00...	2017-08-23 00:00:00...	BACHILLER	2018-05-21 00:00:00...	TITULO	2021-08-23 00:00:00...
4	1995-08-30 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2012-07-15 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-05-21 00:00:00...	SIN TITULO	?
5	1989-05-08 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2006-07-23 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-08-31 00:00:00...	TITULO	2022-02-15 00:00:00...
6	1993-09-11 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2012-07-15 00:00:00...	2017-08-21 00:00:00...	BACHILLER	2018-05-21 00:00:00...	SIN TITULO	?
7	1994-03-21 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2012-07-15 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-05-21 00:00:00...	TITULO	2022-10-18 00:00:00...
8	1990-08-29 00:00:00...	INGENIERIA PE...	2017 - I	F	S	2008-08-10 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-05-21 00:00:00...	TITULO	2021-08-23 00:00:00...
9	1995-07-14 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2011-12-18 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2019-02-04 00:00:00...	SIN TITULO	?
10	1994-12-19 00:00:00...	INGENIERIA PE...	2017 - I	F	S	2012-07-15 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-01-31 00:00:00...	TITULO	2020-06-26 00:00:00...
11	1994-12-15 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2012-07-15 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-05-21 00:00:00...	SIN TITULO	?
12	1993-10-19 00:00:00...	INGENIERIA PE...	2017 - I	F	S	2012-08-26 00:00:00...	2017-08-21 00:00:00...	BACHILLER	2019-10-05 00:00:00...	SIN TITULO	?
13	2017-08-08 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2012-08-26 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-05-21 00:00:00...	SIN TITULO	?
14	1995-09-22 00:00:00...	INGENIERIA PE...	2017 - I	F	S	2012-08-26 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-01-31 00:00:00...	TITULO	2021-08-23 00:00:00...
15	1992-06-08 00:00:00...	INGENIERIA PE...	2017 - I	F	S	2010-08-08 00:00:00...	2017-08-21 00:00:00...	SIN TRAMITE	?	SIN TITULO	?
16	1992-01-13 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2011-08-14 00:00:00...	2017-08-08 00:00:00...	BACHILLER	2018-10-05 00:00:00...	SIN TITULO	?
17	1994-08-28 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2011-08-14 00:00:00...	2017-08-21 00:00:00...	BACHILLER	2018-01-31 00:00:00...	TITULO	2020-09-30 00:00:00...
18	1992-11-13 00:00:00...	INGENIERIA PE...	2017 - I	M	S	2009-12-27 00:00:00...	2017-08-23 00:00:00...	BACHILLER	2018-05-21 00:00:00...	SIN TITULO	?
19	1995-12-31 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2012-12-30 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2018-10-05 00:00:00...	TITULO	2022-04-28 00:00:00...
20	1995-05-17 00:00:00...	INGENIERIA PE...	2017 - II	M	S	2012-12-30 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2018-05-21 00:00:00...	TITULO	2021-09-22 00:00:00...
21	1995-12-05 00:00:00...	INGENIERIA PE...	2017 - II	M	S	2012-08-26 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2018-05-21 00:00:00...	TITULO	2021-09-22 00:00:00...
22	1996-03-22 00:00:00...	INGENIERIA PE...	2017 - II	M	S	2012-12-23 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2018-05-21 00:00:00...	TITULO	2022-07-12 00:00:00...
23	1995-10-22 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2013-04-03 00:00:00...	2017-12-12 00:00:00...	BACHILLER	2018-10-05 00:00:00...	SIN TITULO	?
24	1996-10-23 00:00:00...	INGENIERIA PE...	2017 - II	M	S	2012-08-26 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2018-05-21 00:00:00...	TITULO	2022-04-28 00:00:00...
25	1994-04-01 00:00:00...	INGENIERIA PE...	2017 - II	M	S	2012-08-26 00:00:00...	2017-12-16 00:00:00...	SIN TRAMITE	?	SIN TITULO	?
26	1994-03-22 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2012-12-30 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2019-02-04 00:00:00...	SIN TITULO	?
27	1995-11-23 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2012-12-23 00:00:00...	2018-02-01 00:00:00...	BACHILLER	2018-08-16 00:00:00...	SIN TITULO	?
28	1996-06-01 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2012-08-26 00:00:00...	2018-02-01 00:00:00...	BACHILLER	2019-01-21 00:00:00...	SIN TITULO	?
29	1992-06-09 00:00:00...	INGENIERIA PE...	2017 - II	M	S	2012-12-30 00:00:00...	2017-12-16 00:00:00...	EN TRAMITE	?	SIN TITULO	?
30	1995-10-16 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2012-12-30 00:00:00...	2018-02-01 00:00:00...	BACHILLER	2018-01-21 00:00:00...	SIN TITULO	?
31	1996-10-01 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2012-12-23 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2018-05-21 00:00:00...	TITULO	2022-04-28 00:00:00...
32	1995-04-21 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2012-12-30 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2018-10-05 00:00:00...	SIN TITULO	?
33	1996-08-22 00:00:00...	INGENIERIA PE...	2017 - II	F	S	2012-08-26 00:00:00...	2017-12-16 00:00:00...	BACHILLER	2018-08-16 00:00:00...	SIN TITULO	?
34	1986-10-07 00:00:00...	INGENIERIA PE...	2017 - II	M	S	2006-07-23 00:00:00...	2018-04-30 00:00:00...	BACHILLER	?	SIN TITULO	?

Fig. 14. Datos de las características o atributos seleccionados.

En la Fig. 14 se puede comprobar que algunos datos se encuentran vacíos, no porque falte información, sino porque la información aun no está disponible. Como ejemplo el registro de la fila 15, no tienen fecha de obtención de Bachiller, debido a que aún se encuentra en trámite por lo cual no existe fecha. Similarmente en la fecha de título porque aún no tiene título. Esto no lleva a la necesidad de tener la información por separada los que ya tienen el Bachiller y los que ya tienen el título.

Búsqueda de valores atípicos

Datos a procesar - Orange

Info
422 instances
14 features (8.1 % missing data)
No target variable.
No meta attributes.

Variables
 Show variable labels (if present)
Restore Original Order
 Send Automatically

	FECHA DE INGRES	FECHA EGRESO	Edad	S OBTENIDO BACH	FECHA DE BACHIL
220	2012-12-30 00:00:00...	1994-11-04 00:00:00...	0	BACHILLER	2019-01-21 00:00:00...
13	2012-08-26 00:00:00...	2017-08-08 00:00:00...	0	BACHILLER	2018-05-21 00:00:00...
249	2012-08-26 00:00:00...	2018-01-08 00:00:00...	20.6795	BACHILLER	2018-08-16 00:00:00...
41	2012-12-23 00:00:00...	2017-12-16 00:00:00...	21.1096	EN TRAMITE	
85	2013-12-28 00:00:00...	2018-12-19 00:00:00...	21.1562	BACHILLER	2019-11-28 00:00:00...
237	2012-12-30 00:00:00...	2018-01-08 00:00:00...	21.1616	BACHILLER	2018-10-05 00:00:00...
24	2012-08-26 00:00:00...	2017-12-16 00:00:00...	21.1616	BACHILLER	2018-05-21 00:00:00...

Fig. 15. Datos atípicos en la Edad

Como se puede comprobar en la Fig. 15, existen valores atípicos en la edad, la edad de un egresado no puede ser cero o como se muestra en la Fig. 16 valores perdidos.

Datos a procesar - Orange

Info
422 instances
14 features (8.1 % missing data)
No target variable.
No meta attributes.

Variables
 Show variable labels (if present)
Restore Original Order
 Send Automatically

	ECHA DE INGRESI	FECHA EGRESO	Edad	S OBTENIDO BAC	CHA DE BACHIL
203	2012-08-26 00:0...	2018-01-08 00:0...		? BACHILLER	2018-10-05 00:0
358	2015-04-16 00:0...	2021-09-08 00:0...		? BACHILLER	2022-08-24 00:0
163	2015-08-09 00:0...	2022-09-02 00:0...	42.3753	EN TRAMITE	
387	2016-12-28 00:0...	2022-09-01 00:0...	40.5589	EN TRAMITE	
88	1997-07-15 00:0...	2018-12-19 00:0...	39.7397	EN TRAMITE	
240	2012-08-26 00:0...	2018-01-08 00:0...	37.3096	EN TRAMITE	
356	2016-07-05 00:0...	2021-09-08 00:0...	34.3178	BACHILLER	2022-08-24 00:0

Fig. 16. Valores perdidos en la edad

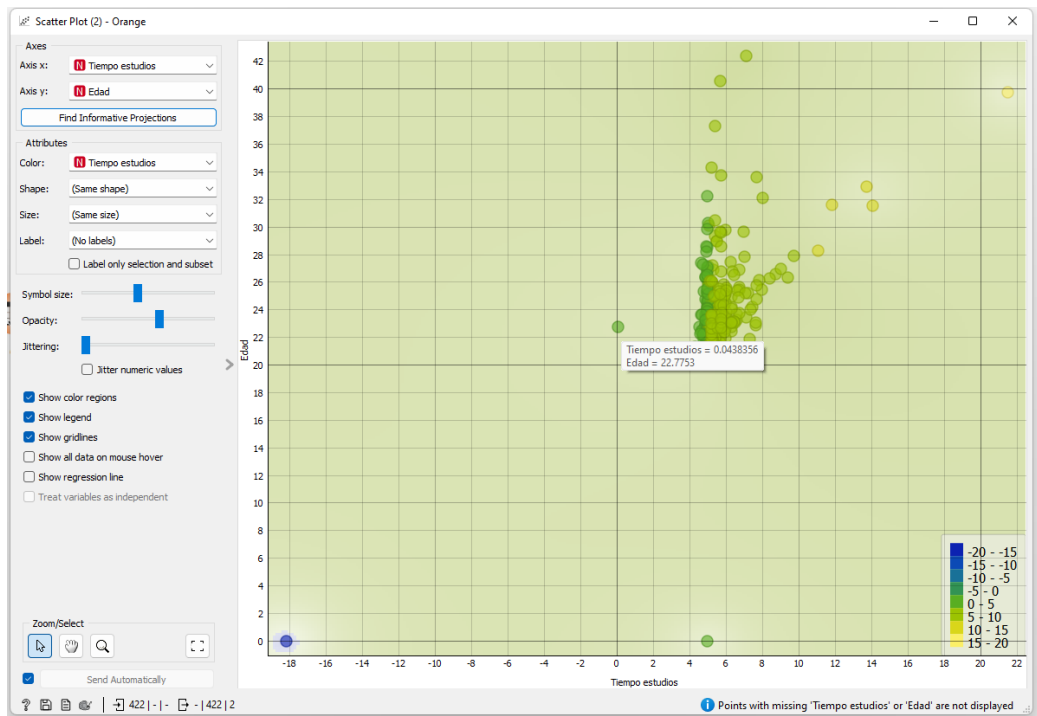


Fig. 17. Scatter plot de los datos

En la relación del tiempo de estudios y la edad, se puede apreciar de manera gráfica que existen valores que están fuera de los rangos donde se concentra la información, lo que nos deduce que estos valores no son valores válidos, si se pone el puntero del mouse en el punto se verá la información del registro.

Este mismo comportamiento se puede apreciar en la Fig. 18.

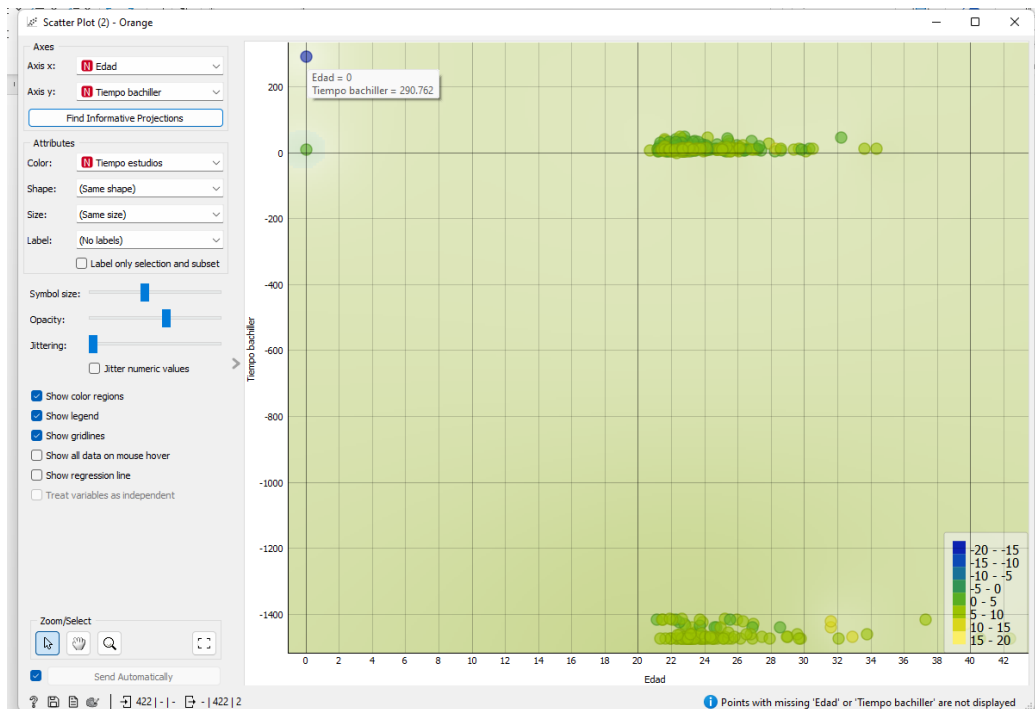


Fig. 18 Scatter plot de la relación Edad y tiempo del bachiller

Esta situación que se presentan de valores atípicos de debe resolver antes de poder realizar el procesamiento de los datos. para no cometer errores por la información mostrada, es importante que se realice previo a la eliminacion de estos valores, la separación de los bachilleres y los titulados de manera separada y despues eliminar los datos realmente atípicos, actividades que se realizan en las fases siguientes.

3. Integrar los datos de diferentes fuentes

Los datos existentes provienen de la única fuente de datos que existen que es el archivo de Excel donde se registra a los egresados, información que sale de la Ficha de egresados motivo por lo cual este paso no se hace necesario al existir una sola fuente de datos.

4. Generar nuevas variables y características

Los datos presentados en la Fig. 12 nos muestra la información disponible, pero que de esta información disponible se debe obtener más información que se desprende de ellas. Conocer en cuanto tiempo acabo la carrera profesional, o en que tiempo se obtuvo el bachiller, cuantos de ellos sacaron el bachiller, cuantos tienen título y en qué tiempo obtuvieron esos bachilleres es de suma importancia.

En la Fig. 11, se puede apreciar que se ha seleccionado 4 atributos que han sido generados para poder tener datos complementarios estos datos son:

Edad: $(\text{Fecha de egreso} - \text{Fecha de nacimiento}) / 365$

Tiempo estudios: $(\text{Fecha de egreso} - \text{Fecha de ingreso}) / 365$

Tiempo bachiller: (Fecha de bachiller – Fecha de egreso) /365

Tiempo título: (Fecha bachiller – Fecha título) / 365

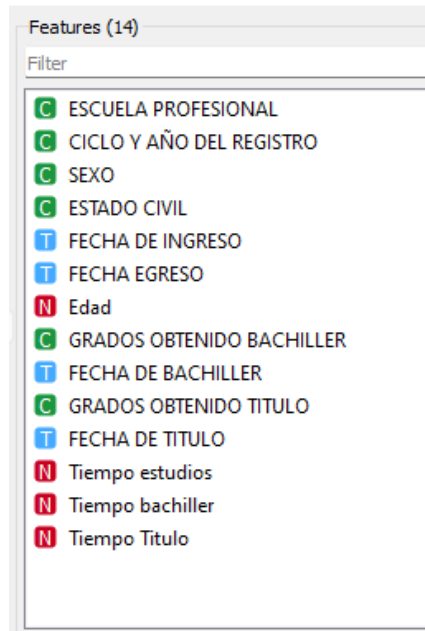


Fig. 19 Nuevas características generadas

3.3.4 Modelización

1. Seleccionar las técnicas de modelado apropiadas

A. Caracterización de los datos de los egresados

En el presente apartado se realiza una serie de caracterización de los datos de los egresados para lo cual se emplea el widget **Pivote Table**. Este widget nos permite analizar y obtener datos de la combinación de diversas características como las escuelas académicas por ciclo de egreso y otras combinaciones que se presentan a continuación:



Fig. 20. Egresados de las escuelas académicas por ciclo de egreso

En la Fig. 20 se muestra que existen 249 egresados de la escuela profesional de alimentos y 173 egresados de la escuela de ingeniería pesquera. Por lo que desde el ciclo 2017-II hasta el semestre 2021-II han egresado 422 estudiantes de ambas escuelas.

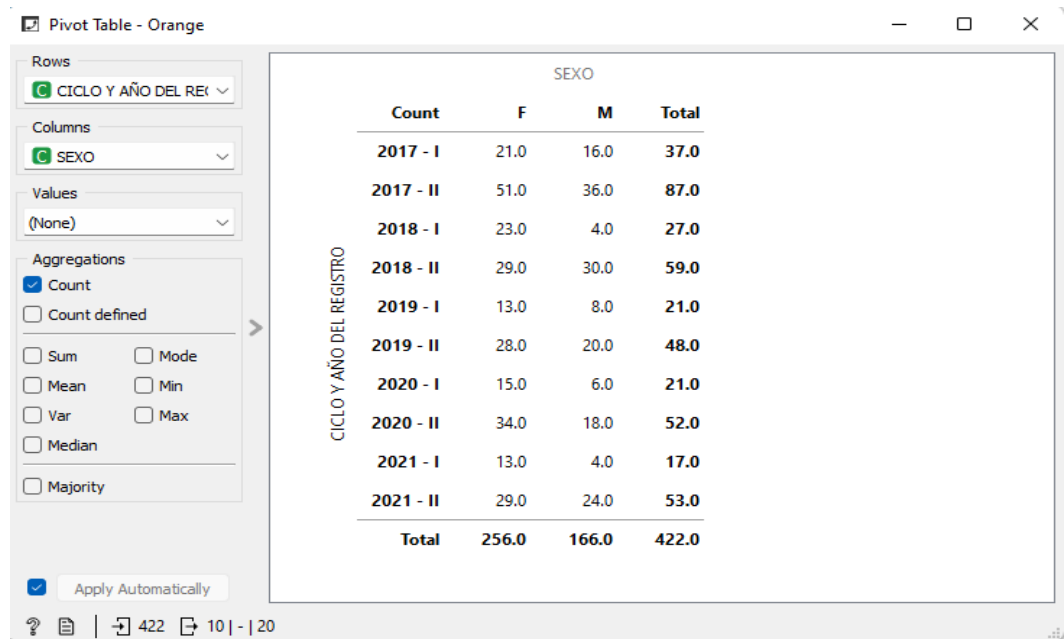


Fig. 21. Egresados por sexo de cada ciclo académico

Con relación al sexo, se ha encontrado que existen 256 egresadas del sexo femenino y 166 egresados del sexo masculino presentados en la Fig. 21, siendo mayor la cantidad de egresadas del sexo femenino, esto se debe a que en la escuela de ingeniería de alimentos

existen una mayor población de estudiantes y por lo tanto mayor cantidad de egresados, con predominancia en el sexo femenino, como se muestra en la Fig. 20.

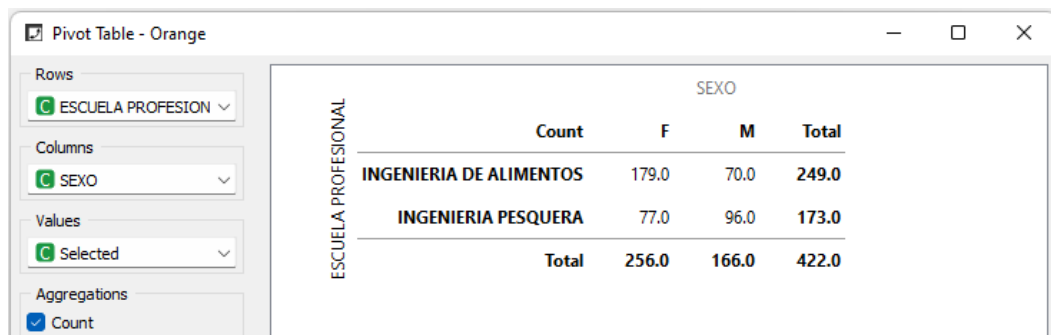


Fig. 22. Distribución de egresado por escuela académica y sexo

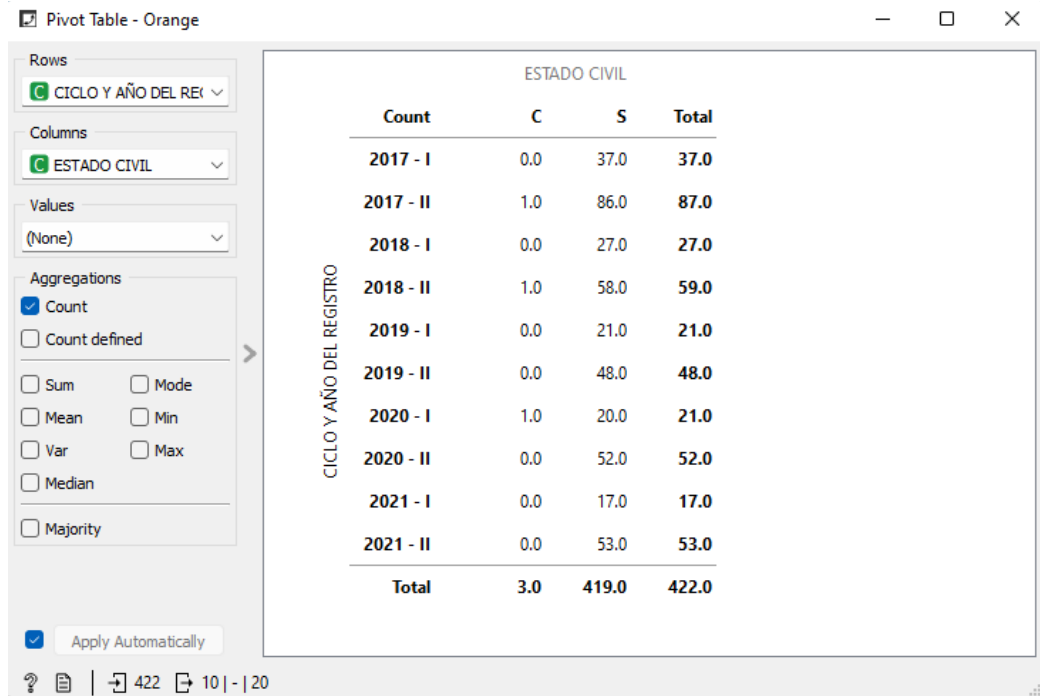


Fig. 23. Egresado según su estado civil

Como se muestra en la Fig. 23, en los resultados del widget Pivot Table, existen 03 egresados que tienen la condición de casados, dos que se presentan en la escuela de ingeniería de alimentos y 1 en la escuela de ingeniería pesquera como se muestra en la Fig. 24. Además, en la Fig. 25 se muestra que los tres egresados de estado civil casado son del sexo femenino.

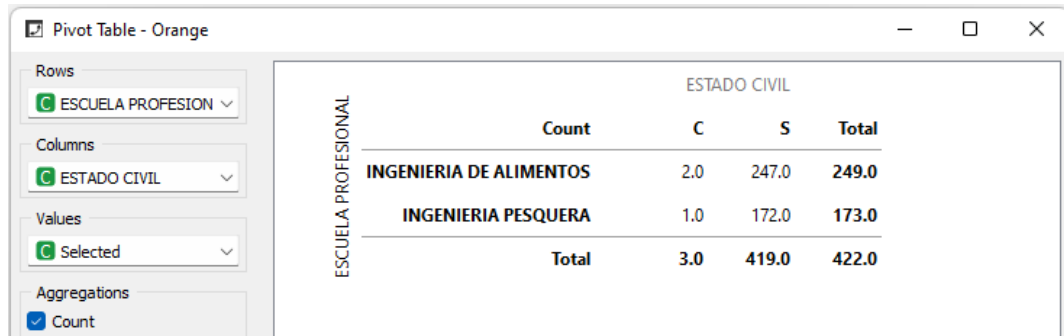


Fig. 24. Distribución de egresados según escuela académica y estado civil

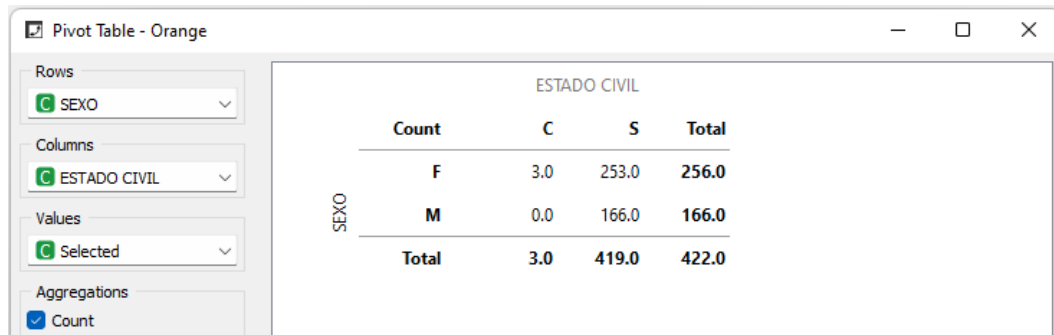


Fig. 25. Distribución de egresados según el sexo y estado civil

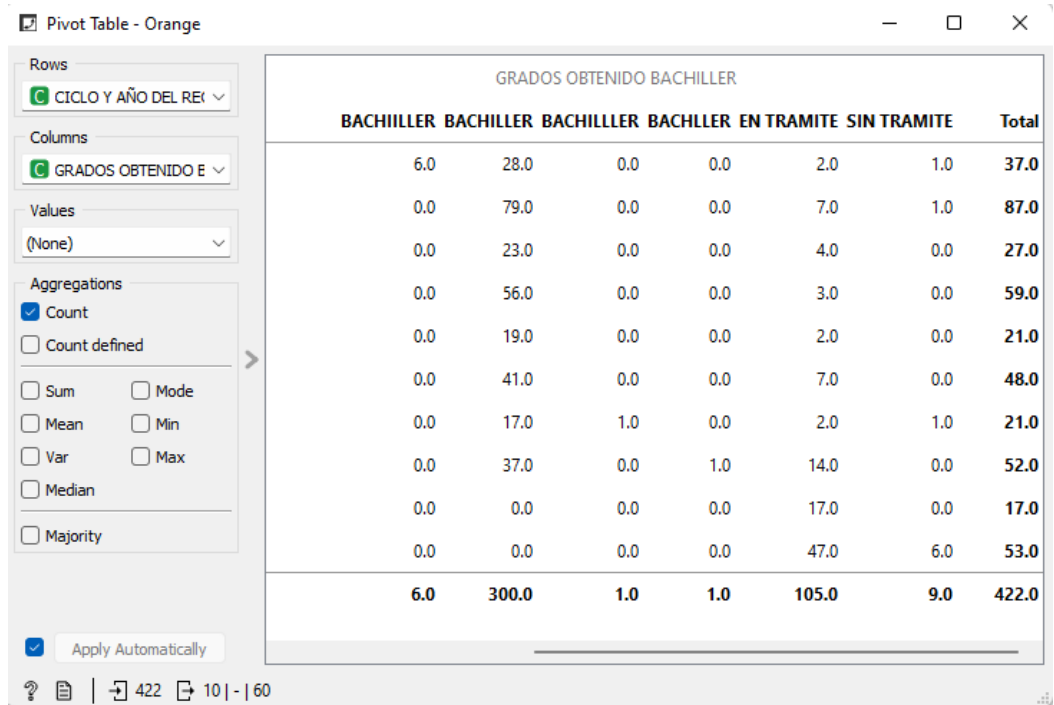


Fig. 26. Egresados con estado del bachillerato

En la Fig. 26 se puede apreciar que al realizar el Pivoteo, estos datos muestran que existe una deficiencia en el registro de los datos, el dato Bachiller figura en diversas notaciones por lo cual se necesita solucionar el registro. La retroalimentación de corrección se realiza en el archivo base de Excel para contar una base de datos más sólida. Después de la corrección se muestra el resultado en la Fig. 27.

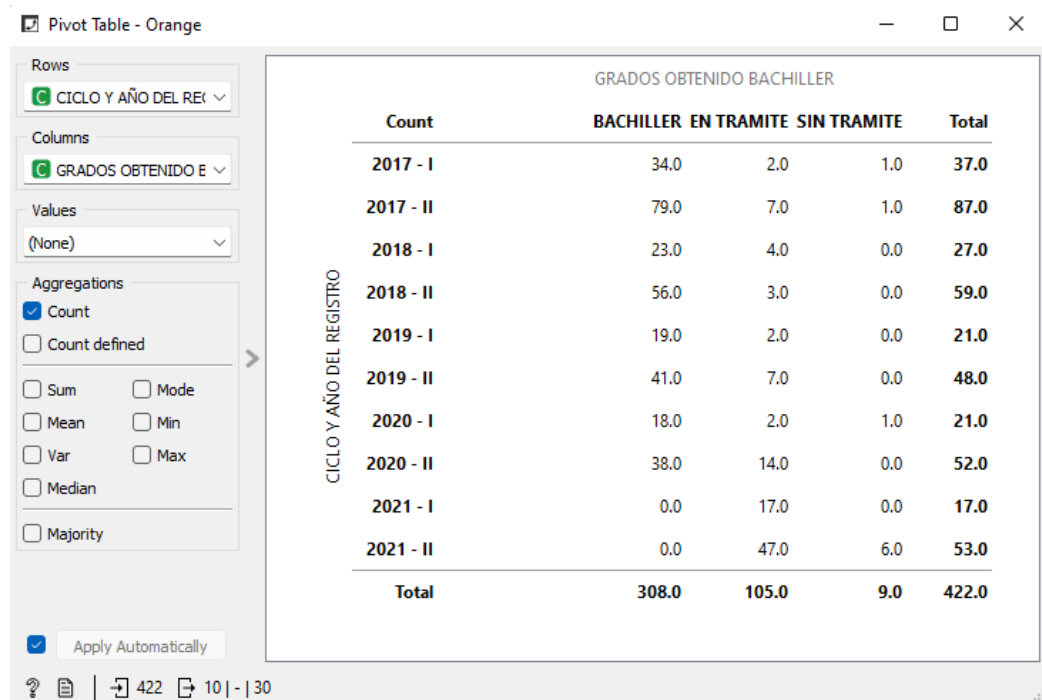


Fig. 27. Estado de los bachilleres a la fecha del procesamiento

En cuanto a los bachilleres, se tiene que, de los 422 registros, 308 ya han obtenido su bachiller, 105 se encuentran en trámite y solo 9 no han realizado el trámite del bachiller. De los bachilleres en trámite el mayor volumen se encuentra entre el semestre 2020-II al 2021-II. con 88 bachilleres en trámite.

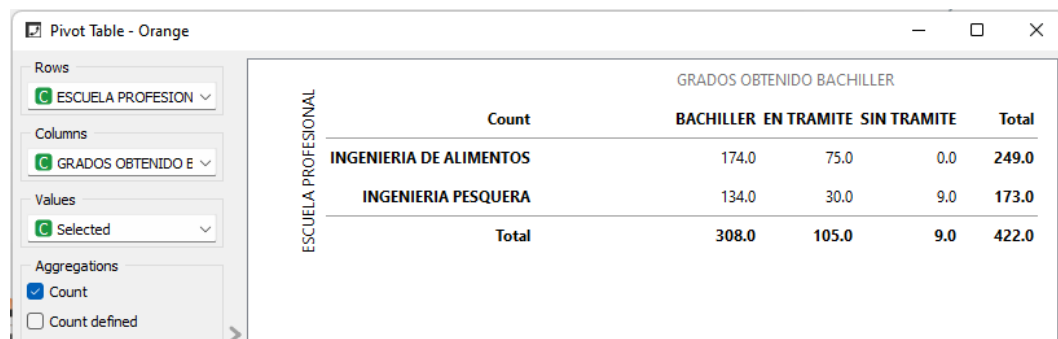


Fig. 28. Estado de los egresados con relación al bachiller

La Fig. 28 muestra que existen 174 bachilleres de la escuela de Ingeniería de alimentos de los 249 egresados del periodo estudiado (69.88%), y 134 bachilleres de los 173 egresados de la escuela de Ingeniería pesquera (77.46%).

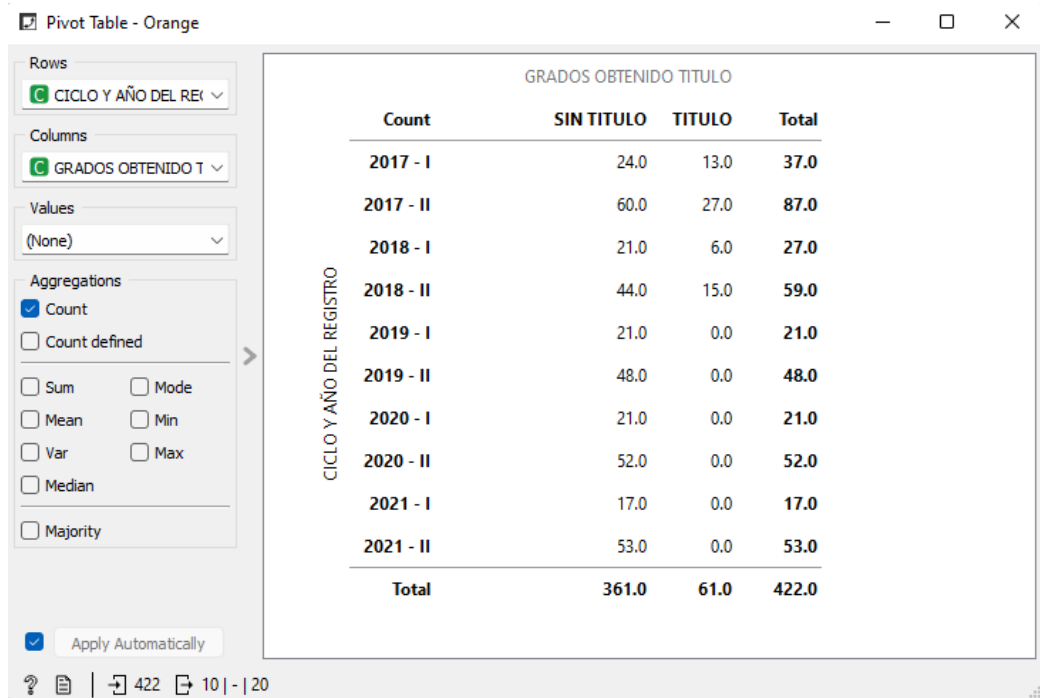


Fig. 29. Estado de los titulados a la fecha del procesamiento

Como se presenta en la información del Pivot Table de la Fig. 29, solo se han titulado a la fecha del procesamiento 61 egresados de los 422, y de las promociones más alejadas (2017-II – 2018-II). esta cantidad solo refleja un 14.45% de titulados siendo un porcentaje bastante bajo. Lo que indica que la facultad no está realizando las acciones necesarias para cumplir con la función principal, tener profesionales, no solo bachilleres.



Fig. 30. Estado de los egresados en relación con el título profesional

En la Fig. 30, los bachilleres titulados durante el periodo en estudio, se cuenta con 39 titulados de la escuela de Ingeniería de alimentos de 249 egresados, vale decir solo

15.66%. De otro lado en la escuela de Ingeniería pesquera se cuenta con 22 titulados de los 173 egresados solo 12.72%.

B. Identificación de información valiosa

Para la identificación se va a emplear widgets como el K-means que es un algoritmo no supervisado que establece agrupamiento de los datos según sus características, y el resultado de este agrupamiento se visualiza con el Scatter plot que es la creación de los puntos de distribución de los datos.

Debemos considerar que de la información base sobre la cantidad de bachilleres (308), y realizar la eliminación de los valores atípicos o perdidos, nos quedamos con un Dataset de 301 registros de egresados. En el caso de los titulados de los 61 registros no hubo valores que eliminar.

	Selected	ESCUELA PROFESIONAL Y AÑO DEL REGISTRO	SEXO	ESTADO CIVIL	Edad	Tiempo estudios	Ti
1	No	INGENIERIA PE...	2017 - I	F	S	23.937	5.06849
2	No	INGENIERIA PE...	2017 - I	M	S	24.1534	4.95342
3	No	INGENIERIA PE...	2017 - I	M	S	22.7425	5.10959
4	No	INGENIERIA PE...	2017 - I	M	S	21.9562	5.06849
5	No	INGENIERIA PE...	2017 - I	M	S	28.2712	11.0521
6	No	INGENIERIA PE...	2017 - I	M	S	23.9589	5.10411
7	No	INGENIERIA PE...	2017 - I	M	S	23.4	5.06849
8	No	INGENIERIA PE...	2017 - I	F	S	26.9616	9
9	No	INGENIERIA PE...	2017 - I	M	S	22.0849	5.64384
10	No	INGENIERIA PE...	2017 - I	F	S	22.6521	5.06849
11	No	INGENIERIA PE...	2017 - I	M	S	22.663	5.06849
12	No	INGENIERIA PE...	2017 - I	F	S	23.8548	4.98904
13	No	INGENIERIA PE...	2017 - I	F	S	21.8932	4.95342
14	No	INGENIERIA PE...	2017 - I	M	S	25.5863	5.98904
15	No	INGENIERIA PE...	2017 - I	M	S	22.9973	6.02466
16	No	INGENIERIA PE...	2017 - I	M	S	24.7918	7.66027
17	No	INGENIERIA PE...	2017 - II	F	S	21.9753	4.96438
18	No	INGENIERIA PE...	2017 - II	M	S	22.6	4.96438
19	No	INGENIERIA PE...	2017 - II	M	S	22.0466	5.30959
20	No	INGENIERIA PE...	2017 - II	M	S	21.7507	4.98356
21	No	INGENIERIA PE...	2017 - II	F	S	22.1562	4.69589
22	No	INGENIERIA PE...	2017 - II	M	S	21.1616	5.30959
23	No	INGENIERIA PE...	2017 - II	F	S	23.7534	4.96438
24	No	INGENIERIA PE...	2017 - II	F	S	22.2082	5.11233

Fig. 31. DataSet de bachilleres

En la parte superior derecha de la información que se muestra en la Fig. 31, se cuenta con las 301 instancias o registros, donde no existe datos perdidos (no missing data) y se seleccionó 7 características o atributos.

Titulados - Orange

Info
61 instances (no missing data)
7 features
Target with 2 values
No meta attributes.

Variables
 Show variable labels (if present)
 Visualize numeric values
 Color by instance classes

Selection
 Select full rows

Restore Original Order

Send Automatically

	UELA PROFESION	Y Y AÑO DEL REGI	SEXO	ESTADO CIVIL	Edad	Tiempo estudios	Tiempo Titulo
17	INGENIERIA PE...	2017 - II	M	S	22.7041	5.42466	4.6137
4	INGENIERIA PE...	2017 - I	M	S	23.4	5.06849	4.4137
36	INGENIERIA DE ...	2017 - II	M	S	30.1041	5.0274	4.39452
40	INGENIERIA DE ...	2017 - II	M	S	23.737	5.04658	4.23836
12	INGENIERIA PE...	2017 - II	M	S	21.7507	4.98356	4.14521
26	INGENIERIA DE ...	2017 - I	F	S	22.1178	5.09041	4.14521
27	INGENIERIA DE ...	2017 - I	M	S	22.9836	4.97534	4.12603
41	INGENIERIA DE ...	2017 - II	F	S	22.7507	5.0274	4.12603
34	INGENIERIA DE ...	2017 - II	F	S	22.6548	5.0274	4.12603
24	INGENIERIA DE ...	2017 - I	F	S	22.0575	5.4137	4.12603
14	INGENIERIA PE...	2017 - II	F	S	21.2219	4.98356	3.93973
13	INGENIERIA PE...	2017 - II	M	S	21.1616	5.30959	3.93973
43	INGENIERIA DE ...	2017 - II	M	S	22.2219	5.3726	3.92055
39	INGENIERIA DE ...	2017 - II	F	S	21.8356	5.0274	3.90411
49	INGENIERIA DE ...	2018 - I	F	S	22.5918	5.04932	3.90411
48	INGENIERIA DE ...	2018 - I	F	S	23.3918	5.07397	3.90411
32	INGENIERIA DE ...	2017 - II	F	S	23.1479	5.0274	3.76986
35	INGENIERIA DE ...	2017 - II	F	S	21.3178	5.3726	3.76986
31	INGENIERIA DE ...	2017 - II	F	S	21.2356	5.3726	3.76986
25	INGENIERIA DE ...	2017 - I	F	S	22.2466	5.09041	3.72329
9	INGENIERIA PE...	2017 - II	F	S	21.9753	4.96438	3.56438
37	INGENIERIA DE ...	2017 - II	M	S	24.6932	5.48767	3.56438

Fig. 32. DataSet de los titulados

Similarmente en esta Fig. 32, se muestra que existen 61 instancias o registros sin datos perdidos, y con 7 características o atributos seleccionados para los titulados.

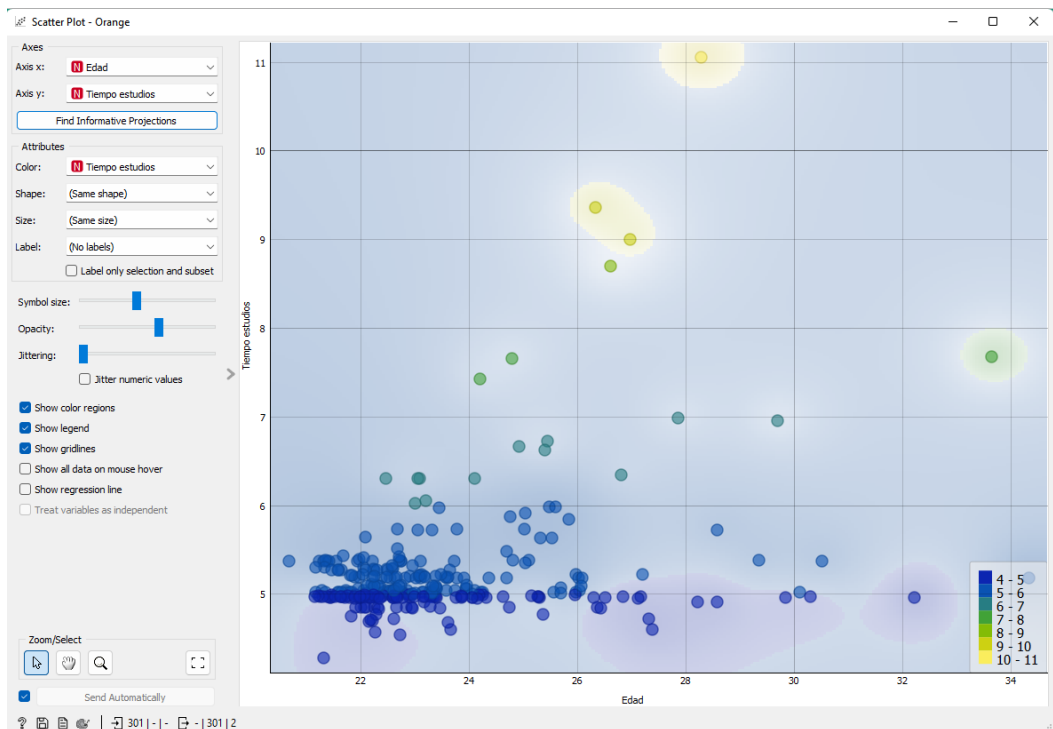


Fig. 33. Distribución de los egresados según edad y tiempo de estudio

La distribución de los egresados según la edad se concentra entre el rango de 21 a 24 siendo tiempo de estudios entre 4.7 a 5.5 años de estudios como se aprecia en la Fig. 33.

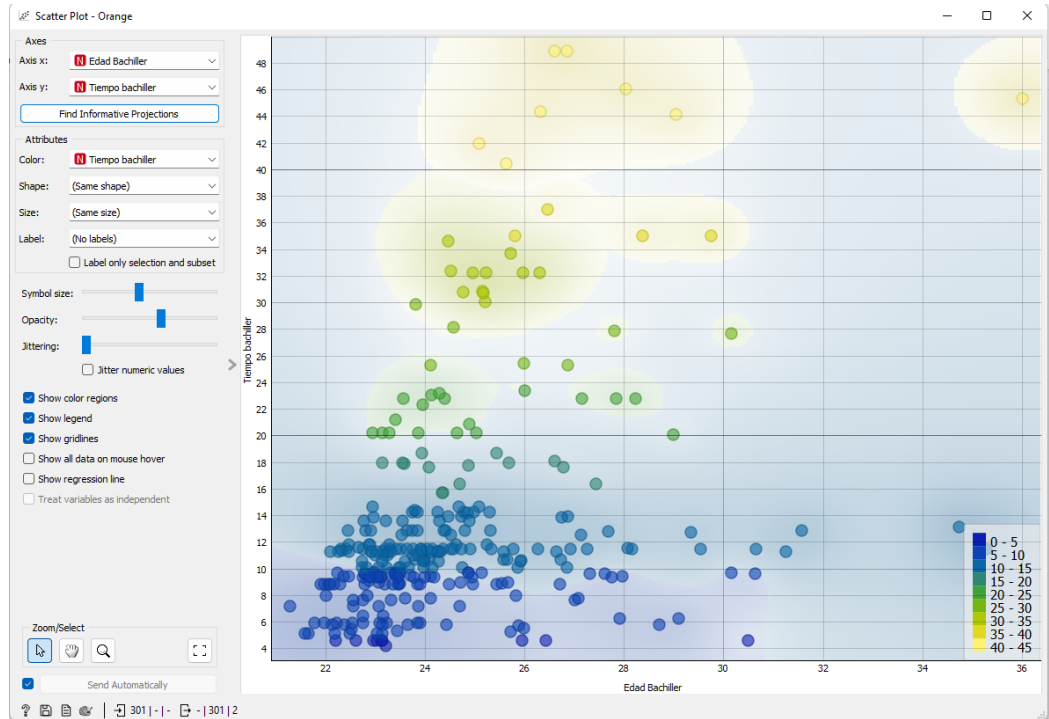


Fig. 34. Distribución de los bachilleres según edad y tiempo de obtención

En la Fig. 34 se puede apreciar que hay una mayor concentración de bachilleres entre los rangos de edad de 22 a 25 años y el tiempo en obtener el bachiller entre 4 a 14 meses.

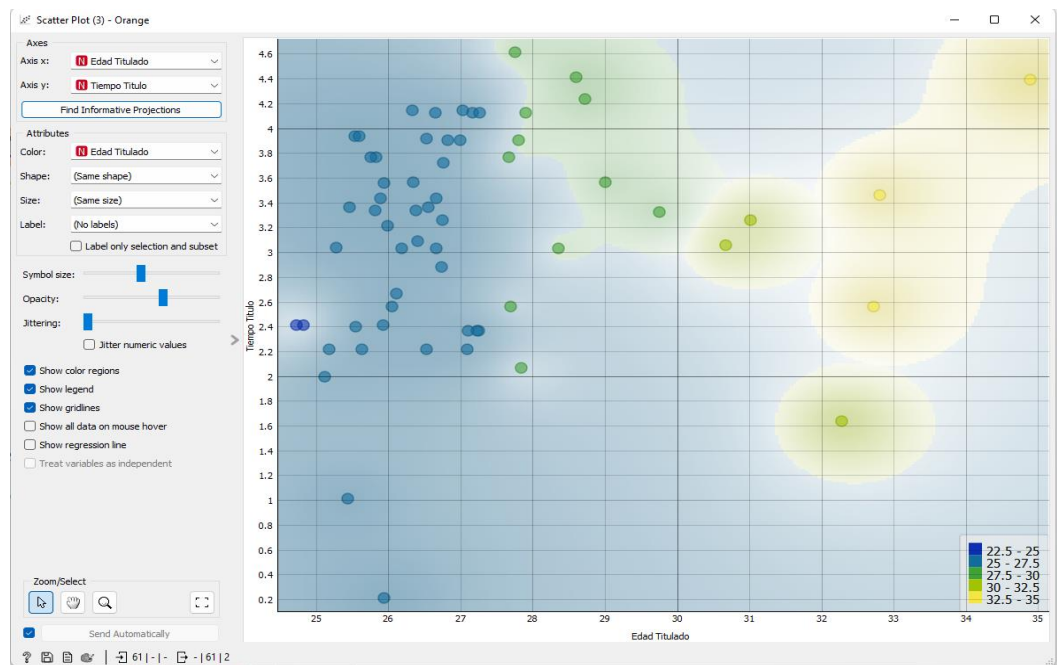


Fig. 35. Distribución de los titulados según la edad y tiempo en obtener el título

En la Fig. 35, de los pocos bachilleres que se han titulado, se puede apreciar que estos están dispersos en el rango de edades de 25 a 27 años y un tiempo de obtenerlo entre 2.2 a 4.2 años.

Adicionalmente a esta identificación con k-mean y Scatter plot, vamos a emplear un widget especializado Interactive K-means, que procesa dinámicamente los puntos de los datos estableciendo nuevos centroides.



Fig. 36 Distribución de los bachilleres según su edad con un centroide

En la Fig. 36 el widget establece una distribución de manera automática y que por medio de este widget Interactive K-means, se puede establecer otros centroides para recomputar los datos. en el caso se han formado tres centroides en las que se agrupan los valores cercanos al centroide. El centroide celeste en un rango de edad entre 21 a 28 años aproximadamente con un tiempo de 4 a 14 meses. El centroide rojo ubica entre los 23-25 años en un tiempo de 15-22 meses, y el centroide color verde un rango de edades entre 24-30 años en un tiempo de 25-50 meses.

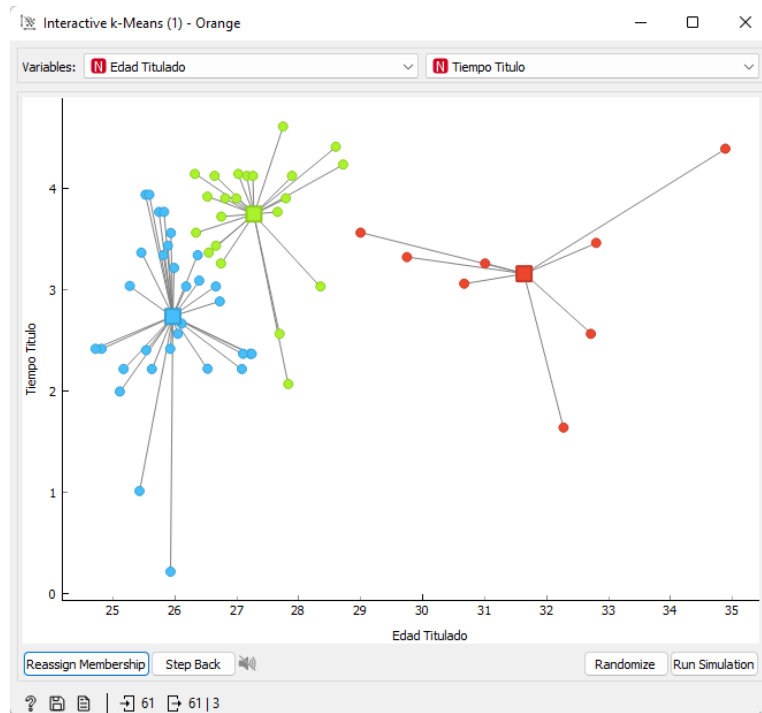


Fig. 37. Distribución de los Titulados con tres centroides

Los centroides de mayor predominancia de los titulados presentados en la Fig. 37, muestran al centroide de color celeste y el de color verde.

C. Estimar los datos de los egresados

Edad de los egresados

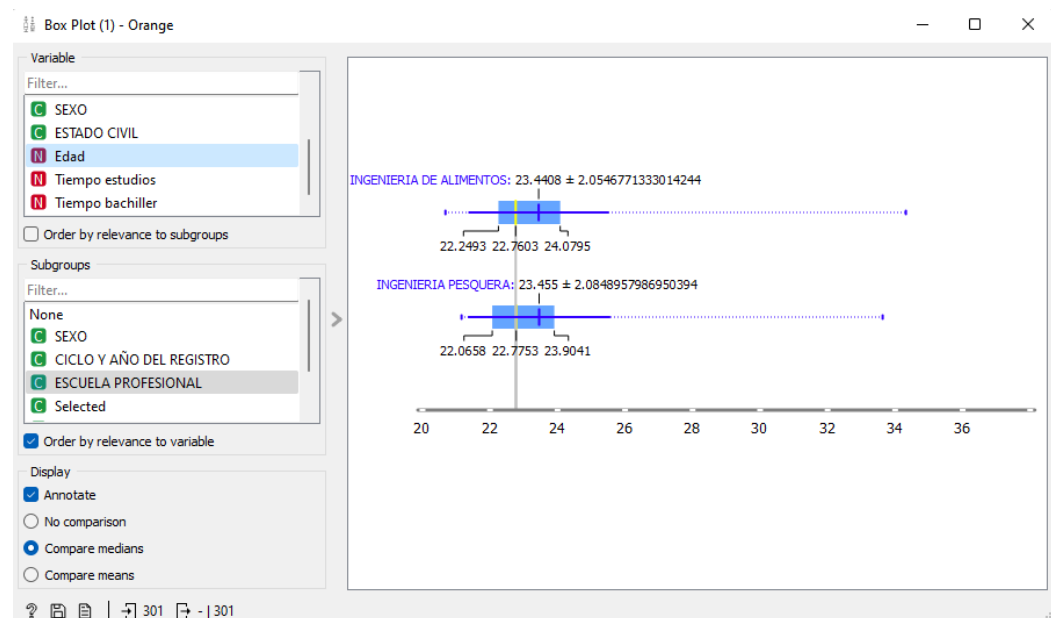


Fig. 38. Edad de los egresados por escuela académica

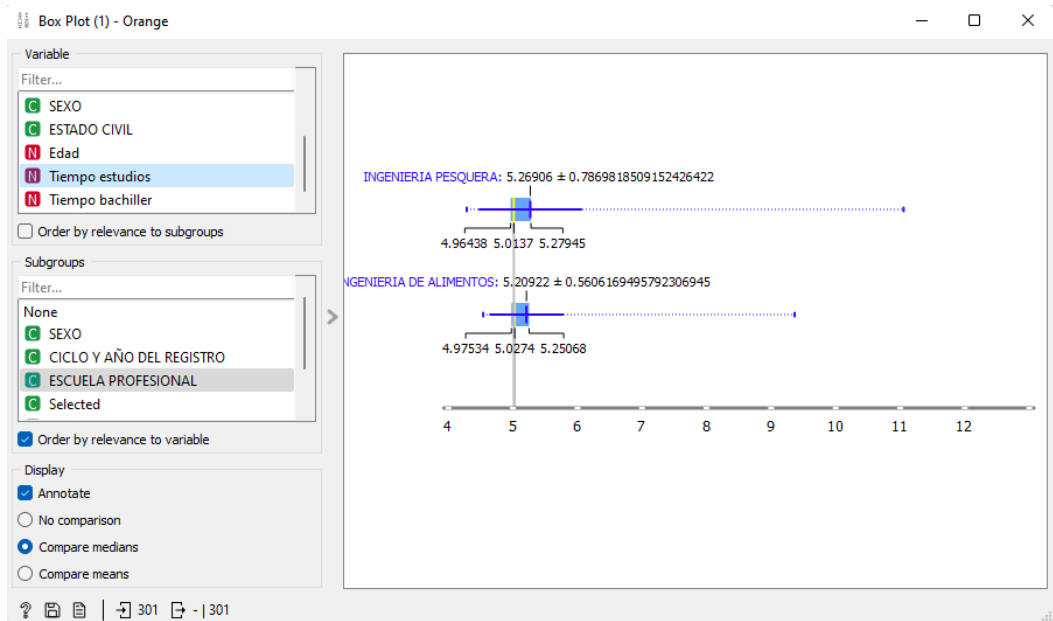


Fig. 39. Tiempo de estudios de los egresados por escuela académica

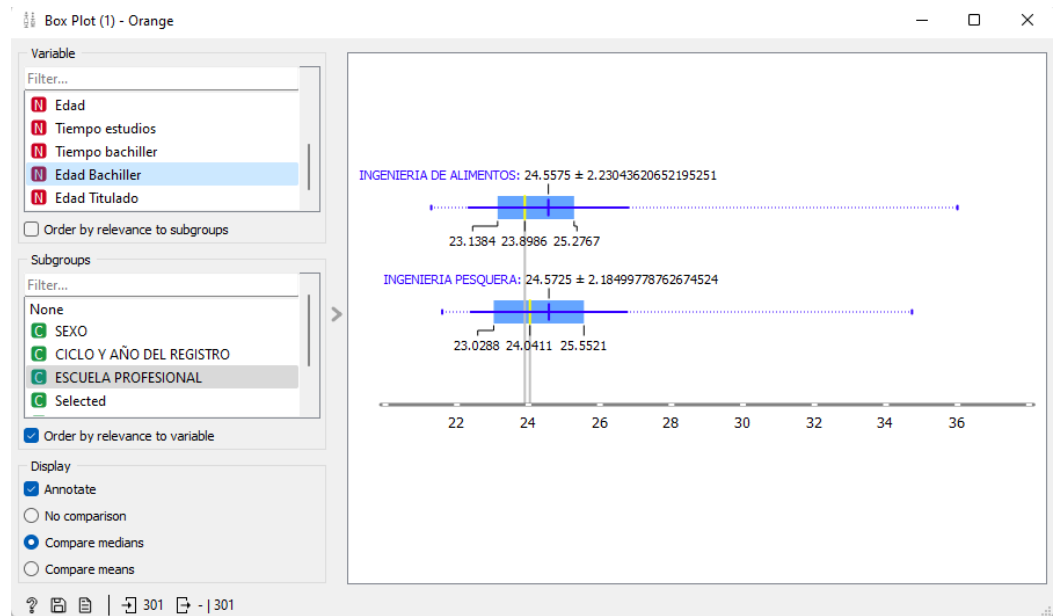


Fig. 40. Edad del Bachiller al obtener el bachiller según la escuela académica

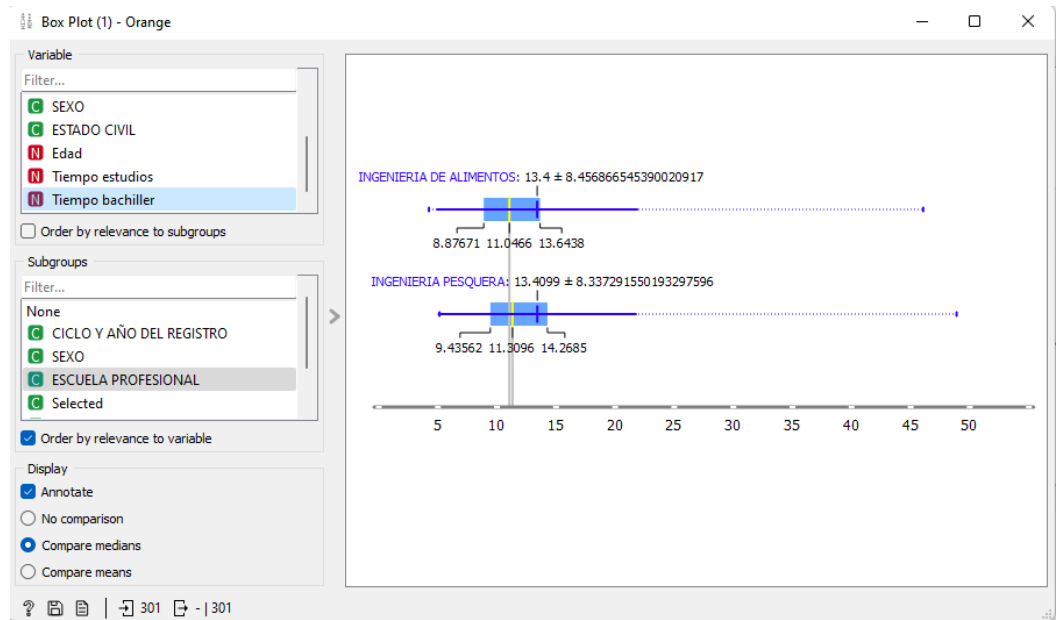


Fig. 41. Tiempo en obtener el bachiller según la escuela académica
Para los Titulados

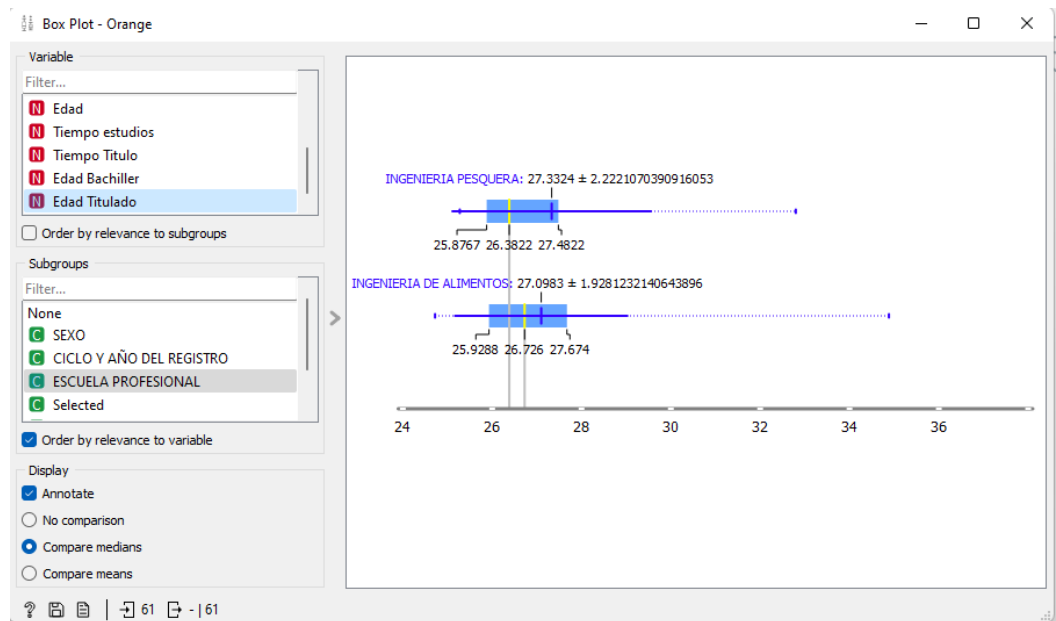


Fig. 42. Edad del titulado al obtener el título profesional por escuela académica

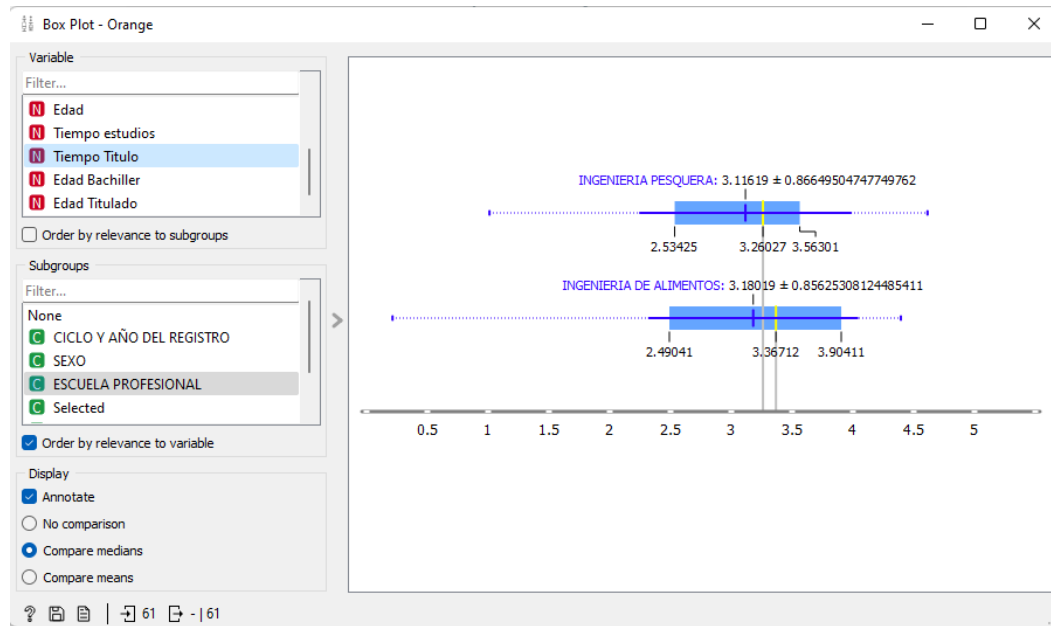


Fig. 43. Tiempo en obtener el título por escuela académica

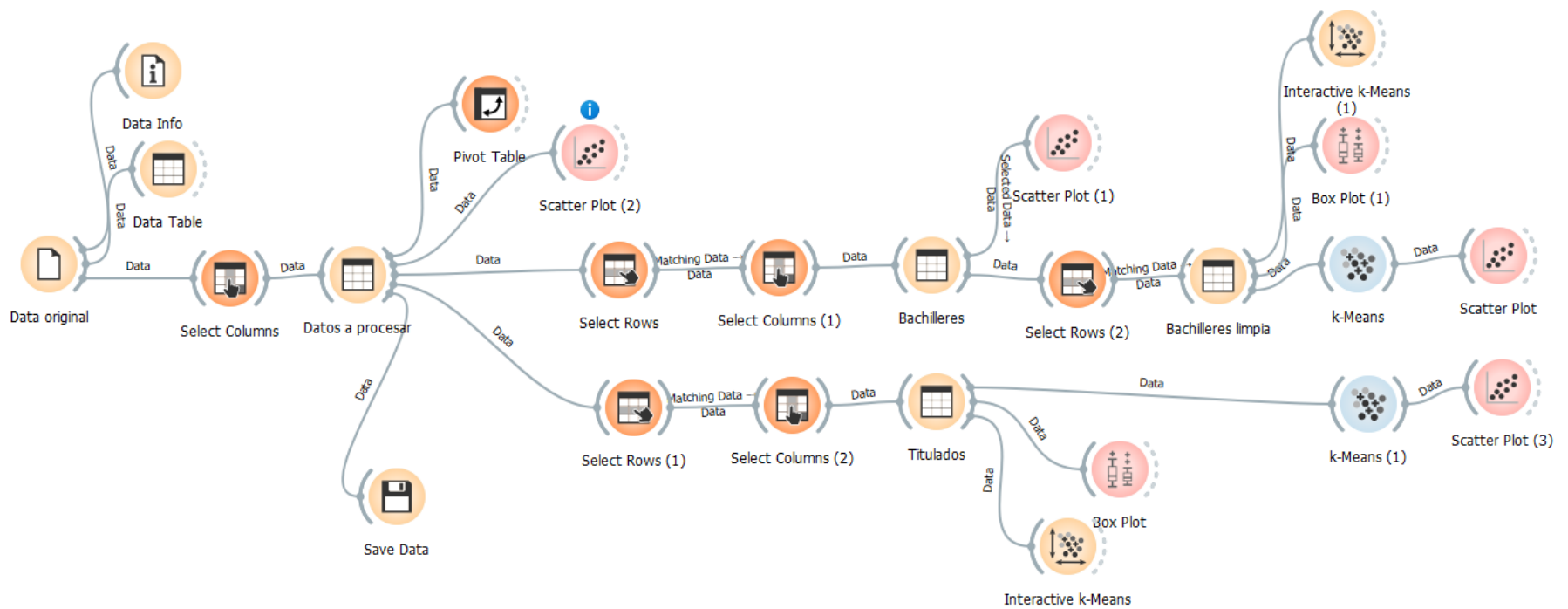


Fig. 44. Modelo de minería de datos empleado

El modelo de minería de dato desarrollado parte de cargar los datos por medio del widget File (Data original), a continuación, se seleccionan las características que son relevantes con el widget Select columns, que luego se pasan al widget Data table (Datos a procesar). Los Datos por procesar se subdividen en dos dataSet con el widget Select Rows que culmina en la selección de un dataSet para los Bachilleres y dataSet para los titulados. A estos dataSet se realizan las evaluaciones respectivas con el modelo no supervisado de Machine Learning con K-Means, y su representación con el widget Scatter Plot.

El widget de los Datos a procesar, se realiza la descripción de los datos por medio del widget Pivot Table.

2. Generar, Construir y evaluar el diseño de prueba

Con la finalidad de realizar una evaluación complementaria de los datos procesados con la minería de datos de la Fig. 44. Se realizó un modelo de evaluación predictiva con diversos algoritmos de Machine Learning y Deep Learning como se muestra en la Fig. 45.

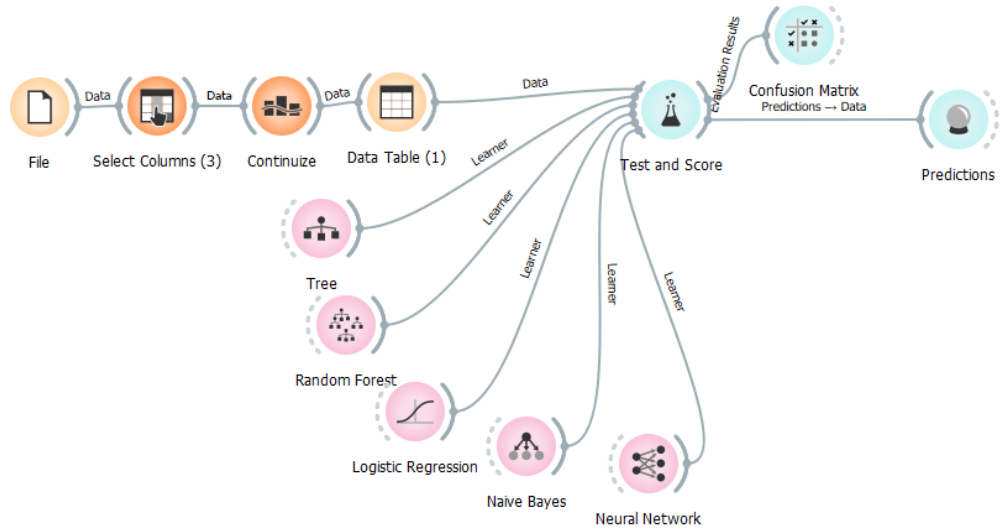


Fig. 45. Diseño de modelo de prueba

Se realizó la evaluación de los algoritmos más importantes de la minería de datos para el trabajo de predicción: los algoritmos Treem Random Forest, Logistic Regression, Naive Bayes y Neural Network se someten por medio del widget test and Score para determinar cual de ellos es el mejor predictor. Como se aprecia en el modelo se tuvo que convertir los datos categóricos en numéricos con el widget continuize, sus resultados se muestran en la Fig. 45.

Data Table (1) - Orange

Info
422 instances
8 features (0.1 % missing data)
Target with 3 values
No meta attributes.

Variables
 Show variable labels (if present)
 Visualize numeric values
 Color by instance classes

Selection
 Select full rows

	¿S OBTENIDO BACH	OBTENIDO TITULO	SIONAL=INGENIERIA	SEXO=M	ESTADO CIVIL=S
1	BACHILLER	1	1	0	1
2	BACHILLER	0	1	1	1
3	BACHILLER	1	1	1	1
4	BACHILLER	0	1	1	1
5	BACHILLER	1	1	1	1
6	BACHILLER	0	1	1	1
7	BACHILLER	1	1	1	1
8	BACHILLER	1	1	0	1
9	BACHILLER	0	1	1	1
10	BACHILLER	1	1	0	1
11	BACHILLER	0	1	1	1
12	BACHILLER	0	1	0	1
13	BACHILLER	0	1	1	1

Fig. 46. Conversión de datos categóricos en datos numéricos.

los mismos que los valores originales (color beige oscuro) como se comprueba con los valores de los registros enmarcados en color rojo.

IV. RESULTADOS

Caracterización de los egresados de la facultad de Ingeniería pesquera y alimentos



Fig. 49. Egresados por semestre y escuela académica

La gráfica de la Fig. 49, muestra los egresados por cada escuela académica en los semestres del periodo en estudio y donde se puede apreciar semestres con pocos egresados como en el 2018-I, 2019-I, 2020-I y 2021-I. este comportamiento es el arrastre desde el semestre 2018-I y se va agudizando en los semestres posteriores.

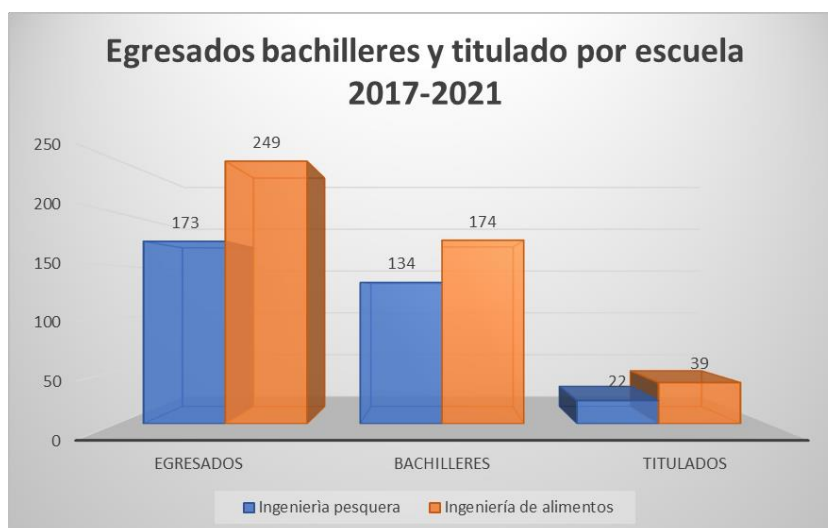


Fig. 50. Egresados, bachilleres con título profesional

La gráfica de la Fig. 50, muestra que, de los 173 egresados de la escuela de ingeniería pesquera, han obtenido su bachiller 134, y de esos 134 bachilleres solo 22 han logrado el título profesional. Similarmente en la escuela de ingeniería de alimentos, de los 249 egresados han obtenido el bachiller solo 174, de los cuales solo 39 han logrado obtener el título profesional.



Fig. 51. Egresados por escuela y sexo

En la gráfica de la Fig. 51, se muestra una predominancia del sexo femenino en la escuela de ingeniería de alimentos con relación a la escuela de ingeniería pesquera.



Fig. 52. Egresados por escuela y estado civil

Con se aprecia en la gráfica de la Fig. 52, el estado civil de los egresados de ambas escuelas es mayoritariamente de estado civil soltero. Del rango de la evaluación solo 3 egresados son casados.

HE 1: Las técnicas de minería de datos, como clustering y clasificación, son capaces de identificar patrones y segmentos significativos en la trayectoria profesional de los egresados, permitiendo un análisis más detallado de su desempeño y evolución laboral.

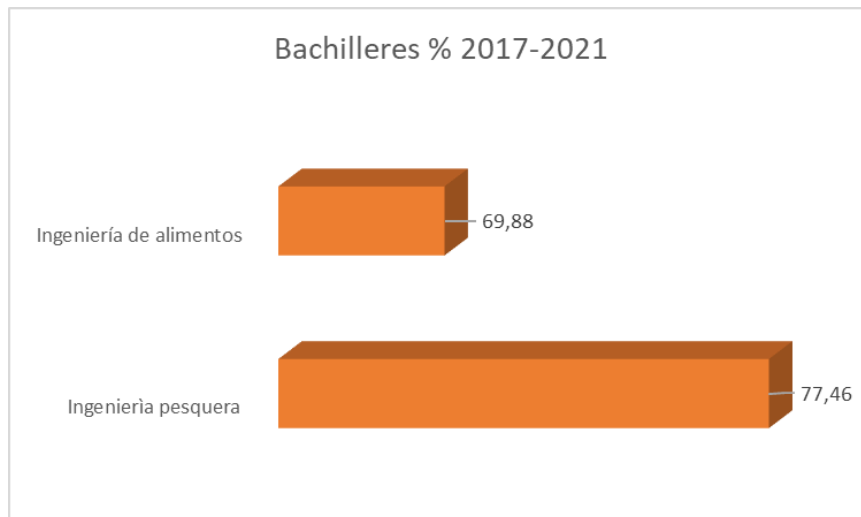


Fig. 53. Bachilleres por escuela académica

En la gráfica de la Fig. 53, se puede apreciar que existe un mayor porcentaje de egresados que han obtenido su bachiller en la escuela de ingeniería pesquera con relación a la escuela de ingeniería de alimentos. Sin embargo, es importante resaltar la preocupación que se debe tener debido siendo el bachiller automático, este porcentaje debería ser mucho mayor.

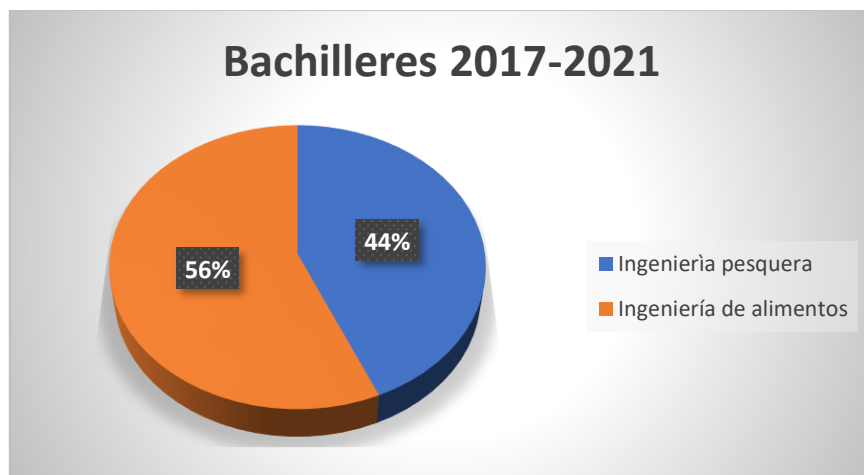


Fig. 54. Relación entre bachilleres por escuela

La Fig. 54, muestra la relación entre los bachilleres de las escuelas, en la que se tiene un mayor porcentaje 56% de bachilleres en la escuela de Ingeniería de alimentos. A pesar de ello debe ser una preocupación para la facultad esta situación.

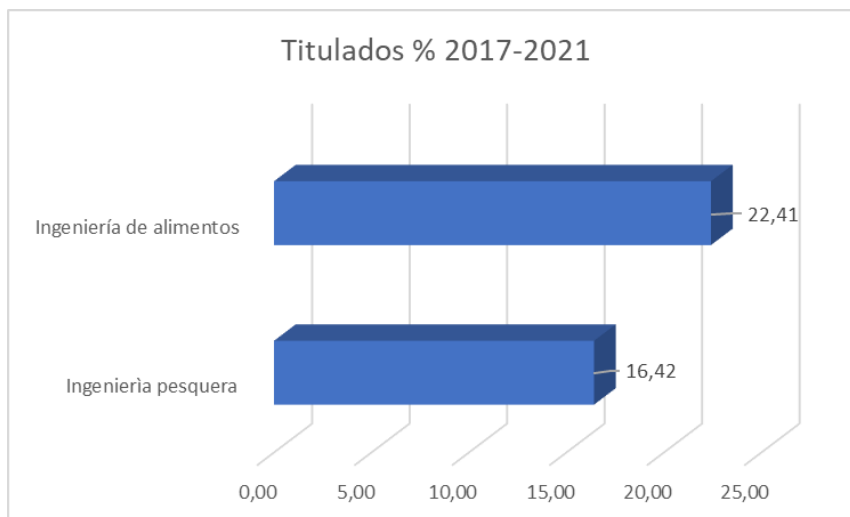


Fig. 55. Titulados por escuela académica

La Fig. 55, demuestra un reducido porcentaje (menos del 23%) de bachilleres que han obtenido su título profesional en ambas escuelas académicas,



Fig. 56. Relación de titulados por escuela académica

En la Fig. 56, la gráfica de la reducida cantidad de titulados el mayor porcentaje se da en la escuela de ingeniería de alimentos con relación a la de ingeniería pesquera.

HE 2: La información obtenida a través de la minería de datos tendrá un impacto positivo en la toma de decisiones y en la formulación de estrategias para mejorar el apoyo y la formación de los egresados, contribuyendo al desarrollo profesional continuo y al ajuste de los programas educativos.

TABLA IV
ESTIMACION ESTADISTICA DE LOS DATOS DE INGENIERIA PESQUERA

Ingeniería pesquera	Media	Mediana	Q1 (25%)	Q2(75%)
Edad del egresado	23,46	22,78	22,07	23,9
Tiempo de estudios	5,27	5,01	4,96	5,28
Edad del bachiller	24,57	24,04	23,03	25,55
Tiempo en obtener bachiller	13,41	11,31	9,44	14,27
Edad del titulado	27,33	23,38	25,88	27,48
Tiempo en obtener el titulo	3,12	3,26	2,53	3,56

En la tabla IV, se tiene la información estadística de los datos los egresados, edad en la que egresaron, el tiempo de estudios, al igual que la edad cuando obtuvieron su bachiller y el tiempo que duro obtener su bachiller. Igualmente, la edad que tenían cuando obtuvieron su título y el tiempo que duró obtener su título profesional. Información de la escuela de pesquería.

TABLA V
ESTIMACION ESTADISTICA DE LOS DATOS DE INGENIERIA DE ALIMENTOS

Ingeniería de alimentos	Media	Mediana	Q1 (25%)	Q2(75%)
Edad del egresado	23,44	22,76	22,49	24,08
Tiempo de estudios	5,21	5,03	4,98	5,25
Edad del bachiller	24,56	23,90	23,14	25,28
Tiempo en obtener bachiller	13,40	11,05	8,88	13,64
Edad del titulado	27,10	26,73	25,93	27,67
Tiempo en obtener el titulo	3,18	3,36	2,49	3,90

En la Tabla V, se tiene la información estadística de los datos los egresados, edad en la que egresaron, el tiempo de estudios, al igual que la edad cuando obtuvieron su bachiller y el tiempo que duro obtener su bachiller. Igualmente, la edad que tenían cuando obtuvieron su título y el tiempo que duró obtener su título profesional. Información de la escuela de alimentos.

V. DISCUSIÓN

La minería de datos es un excelente aliado para poder obtener información valiosa sobre este proceso del control del perfil del egresado con técnicas de aprendizaje no supervisado como K-means, con ella encontrar patrones en los datos de los egresados, similarmente el estudio de [14]. La información obtenida debe servir para mejorar los procesos académicos entre ellos los procesos de bachilleres y titulados ya que los titulados son los que determinan la calidad de la institución, evaluación también obtenida en el estudio de [5]. Al mismo tiempo el seguimiento de los egresados ayuda a comprender su desempeño profesional [3], sin embargo existe mucha información que la facultad de pesquería no cuenta sobre este desempeño profesional, lo que dificulta la retroalimentación de los planes académicos, como también lo establece la investigación de [24].

VI. CONCLUSIONES

Como conclusión general:

- La aplicación de técnicas de minería de datos en el seguimiento de egresados de la Facultad de Ingeniería Pesquera de la Universidad Nacional San Luis Gonzaga durante el período 2017-2021 ha demostrado ser una herramienta valiosa para obtener información precisa y significativa sobre la trayectoria profesional de los egresados. Las técnicas empleadas permitieron identificar patrones y tendencias relevantes que facilitaron un análisis más profundo de la empleabilidad, el tipo de empleo, el nivel de satisfacción laboral y la continuación de estudios de los egresados.

Como conclusiones específicas:

- Las técnicas de minería de datos, como clustering y clasificación, demostraron ser efectivas para identificar patrones significativos en la trayectoria profesional de los egresados. Estos métodos permitieron descubrir segmentos relevantes dentro de la población de egresados, proporcionando una base sólida para ajustar las estrategias de formación y apoyo. Los resultados sugieren que la minería de datos puede ser una herramienta poderosa para el análisis de datos educativos y profesionales.
- La información obtenida a través de la minería de datos ha tenido un impacto positivo en la toma de decisiones institucionales y en la formulación de estrategias para el apoyo a los egresados. Los hallazgos derivados del análisis han facilitado la implementación de políticas más informadas y efectivas, contribuyendo a una mejora en los programas educativos y de apoyo a los egresados.

VII. RECOMENDACIONES

- Establecer Protocolos de Recolección de Datos: Implementar procedimientos estandarizados para la recolección y actualización de datos sobre los egresados. Asegurarse de que la información recolectada sea completa, precisa y actualizada.
- Desarrollar un Sistema Integral de Gestión de Datos: Crear o mejorar una base de datos centralizada que almacene información detallada sobre los egresados, incluyendo datos sobre empleabilidad, tipo de empleo, satisfacción laboral y continuación de estudios.
- Capacitar al Personal en Gestión de Datos: Proporcionar formación continua al personal encargado de la recolección y gestión de datos para asegurar la correcta entrada, mantenimiento y actualización de la información.
- Implementar Técnicas de Minería Avanzadas: Continuar explorando y aplicando técnicas avanzadas de minería de datos, como redes neuronales y análisis predictivo, para obtener insights más profundos y detallados sobre la trayectoria de los egresados.
- Validar y Refinar Modelos: Regularmente evaluar y ajustar los modelos de minería de datos para mejorar su precisión y utilidad. Utilizar métricas de rendimiento como la precisión, el recall y la F1-score para asegurar que los modelos estén proporcionando resultados válidos.
- Monitorear y Evaluar Resultados: Establecer indicadores clave de rendimiento (KPI) para evaluar el impacto de las nuevas estrategias y políticas implementadas basadas en los hallazgos de la minería de datos. Realizar revisiones periódicas para ajustar las estrategias según sea necesario.
- Recoger Retroalimentación de Egresados: Implementar encuestas y mecanismos de retroalimentación para obtener opiniones directas de los egresados sobre la efectividad de las nuevas políticas y programas. Utilizar esta retroalimentación para realizar ajustes y mejoras continuas.
- Establecer Alianzas con Empresas: Fomentar colaboraciones con empresas y organizaciones del sector para obtener información adicional sobre la empleabilidad y las necesidades del mercado laboral. Utilizar esta información para adaptar los programas educativos y de apoyo a las demandas actuales del mercado.
- Crear Redes de Contacto: Desarrollar redes de contacto entre egresados, empleadores y la facultad para facilitar oportunidades de empleo, mentoría y colaboración profesional.

VIII. REFERENCIAS BIBLIOGRÁFICAS

- [1] M. Bander, M. Medina, G. Tapia, y S. Tissome, «LA EDUCACIÓN PERMANENTE DEL EGRESADO PARA SATISFACER LAS DEMANDAS LABORALES ACTUALES EJE», en *PROYECTO POLITICO DE LA EDUCACION SUPERIOR*, Argentina: Universidad Nacional de la Plata, 2020, p. 8.
- [2] J. C. Cevallos, M. C. Escobar, J. E. Falcones, y W. J. Cevallos, «Modelado laboral de los egresados de la Facultad de Ciencias Informáticas de la Universidad Técnica de Manabí (Ecuador)», *Inf. Tecnológica*, vol. 32, n.º 6, pp. 111-122, 2021, doi: <http://dx.doi.org/10.4067/S0718-07642021000600111>.
- [3] H. Coello Caballero, M. A. Cruz Barrios, y J. A. Furones Mourelle, «Desempeño profesional de egresados de la maestría en Farmacoepidemiología», *Educ. Médica Super.*, vol. 35, n.º 2, p. 3, 2021.
- [4] M. C. J. Erick *et al.*, «PLATAFORMA WEB CON INTEGRACIÓN DE MINERÍA DE DATOS Y REDES SOCIALES PARA EL SEGUIMIENTO A GRADUADOS DEL PROGRAMA DE INGENIERÍA DE SISTEMAS DE LA UNIVERSIDAD DE CUNDINAMARCA, EXTENSIÓN FACATATIVÁ. JULIÁN», *Rev. CENIC Cienc. Biológicas*, vol. 152, n.º 3, p. 28, 2016.
- [5] J. Gómez García y Z. Juan, «EL SEGUIMIENTO AL EGRESADO DE UNISABANETA EN EL MERCADO LABORAL», Corporación Universitaria de Sabaneta, 2018.
- [6] E. Gonzales Morales, S. J. José, y D. L. R. Tonyse, «Seguimiento al egresado y graduado, Caso de Estudio Ingeniería en Gestión Empresarial en la Universidad Metropolitana, Ecuador», *Rev. Científica Multidisciplinaria*, vol. 3, n.º 1, pp. 82-87, 2018.
- [7] H. Lamos-Díaz, A. Ramírez-Sierra, y M. L. Rangel-Granados, «Seguimiento a graduados del programa Ingeniería Industrial de la Universidad Industrial de Santander mediante análisis multivariado y la red profesional LinkedIn », *Rev. Educ. En Ing.*, vol. 15, n.º 29, pp. 73-82, 2019.
- [8] F. Y. Monsalve Villa, «Modelo para la gestión de egresados integrando minería de datos y redes sociales», Tecnológico de Antioquia, 2017.
- [9] L. Román Fernández, L. Díaz González, y D. Covarrubias Melgar, «Seguimiento de egresados de la unidad académica de enfermería no. 1, cohortes 1999- 2010, universidad autónoma de guerrero. un aporte de evidencias», n.º 117, pp. 87-105, 2016.
- [10] P. L. Obando y H. H. V. Bolívar, «Modelo de seguimiento a egresados con énfasis en empleabilidad», *Rev. Electrónica Gest. Las Pers. Tecnol.*, vol. 12, n.º 35, pp. 17-26, 2019.
- [11] J. P. Durand Villalobos, «Maestría en Innovación Educativa , periodo 2011 - 2017 Study of graduates of the Master in Educational Innovation , period 2011-2017», *RECIE Rev. Electrónica Científica Investig. Educ.*, vol. 4, p. 1040, 2017.
- [12] G. Quijada, S. Aguayo, y L. Monárrez, «El seguimiento de egresados de la licenciatura en Historia . Percepción de su formación académica y oportunidades de cambio y mejora del Plan de estudios .», México, 2019.
- [13] T. Esmeralda. Rivera Alvarenga, «Seguimiento de egresados del programa de Salud Ocupacional del CIES UNAN Managua periodo 2010-2016», *Bibl. Responsável*, p. 72, 2018.
- [14] E. M. Asto Rodriguez, «FRAMEWORK BASADO EN MINERÍA DE DATOS PARA LA OBTENCIÓN DEL PERFIL DE EGRESO DE LOS ESTUDIANTES DEL PROGRAMA DE INGENIERIA MECATRÓNICA DE LA UNIVERSIDAD NACIONAL DE TRUJILLO AÑO 2019», Universidad Privada Antenor Orrego, 2020.
- [15] J. R. Colquehuanca Achulli, «Sistema de información de egresados del Instituto de Educación Superior Pedagógico Público de Chimbote y mejora de la empleabilidad. Estudio de Caso a partir del Proyecto “Diseño de un sistema de seguimiento a egresados” en la Provincia del Santa - Chimbo», Pontificia Universidad Católica del Perú- Lima, Perú, 2020.

- [16] D. Mendoza Jiménez, «Sistema de información transaccional (SITE) para mejorar el seguimiento al egresado de la facultad de ingeniería industrial y sistemas de la Universidad Nacional Hermilio Valdizan de Huánuco», Universidad Nacional Daniel Alcides Carrión - Cerro Pasco, Perú, 2019.
- [17] J. Mori *et al.*, *Diseño de un sistema de seguimiento de egresados y una estrategia para la implementación de dicho sistema*. 2018.
- [18] P. Neira, J. Ramírez, y V. Mabel, «FORMATIVO DEL PROFESIONAL DE ENFERMERÍA THE FOLLOW-UP OF GRADUATES AND ITS IMPACT ON THE NURSING INTRODUCCIÓN En el Perú , la educación es un derecho fundamental y servicio público esencial que toma en cuenta la realidad y continuar aprendiendo durante t», pp. 0-2.
- [19] L. Damien, *Data Mining*, Kindle. 2019.
- [20] C. Chearer, «(4) CRISP-DM The New Blueprint for Data Mining Colin Shearer (Fall 2000) | ankita bhagwat - Academia.edu». Accedido: 3 de abril de 2023. [En línea]. Disponible en: https://www.academia.edu/42079490/CRISP-DM_The_New_Blueprint_for_Data_Mining_Colin_Shearer_Fall_2000_?auto=download
- [21] P. Chapman *et al.*, «CRISP-DM 1.0: Step-by-step data mining guide», 2000. Accedido: 3 de abril de 2023. [En línea]. Disponible en: <https://www.semanticscholar.org/paper/CRISP-DM-1.0%3A-Step-by-step-data-mining-guide-Chapman-Clinton/54bad20bbc7938991bf34f86dde0babfd2d5a72>
- [22] R. Hernandez-Sampieri, C. Fernández-Collado, y P. Baptista-Lucio, *Metodología de la investigación*, 6ta Ed. Mexico D.F., 2014.
- [23] C. Weng, R. B. Ghazali, S. A. Mustafa, A. N. Kareem, y B. A. Khalaf, «Weather Forecasting for Batu Pahat Using Neural Network», *Fusion Pract. Appl.*, vol. 6, n.º 2, pp. 64-70, 2021, doi: 10.54216/fpa.060204.
- [24] J. R. Colquehuanca Achulli, «Sistema de información de egresados del Instituto de Educación Superior Pedagógico Público de Chimbote y mejora de la empleabilidad. Estudio de Caso a partir del Proyecto “Diseño de un sistema de seguimiento a egresados” en la Provincia del Santa - Chimbo», Pontificia Universidad Católica del Perú- Lima, Perú, 2020.

IX. ANEXOS

Anexo 01: FICHA DE REGISTRO DEL EGRESADO

APELLIDO PATERNO	APELLIDO MATERNO	NOMBRES

DNI	ESTADO CIVIL	SEXO	LUGAR DE NACIMIENTO	FECHA DE NACIMIENTO	NACIONALIDAD

DIRECCIÓN(Av./Jr./Calle/Numero)	NÚMERO DE REGISTRO

DISTRITO	PROVINCIA	DEPARTAMENTO	PAÍS

ESCUELA ACADÉMICO PROFESIONAL	TELÉFONO O CELULAR	EMAIL

DATOS ACADÉMICOS

FECHA DE INGRESO A LA UNSLG	FECHA DE EGRESO A LA UNSLG	CÓDIGO DE MATRICULA			

SEGUIMIENTO DEL EGRESADO

Marque X si Usted se encuentra: 1. Trabajando 2. Prácticas Profesionales

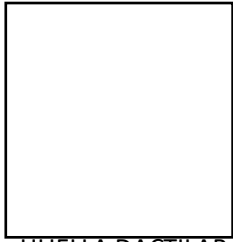
NOMBRE DE LA EMPRESA O INSTITUCIÓN	DIRECCIÓN DE LA EMPRESA

ÁREA EN QUE REALIZA SU LABOR	FUNCIÓN QUE DESEMPEÑA	FECHA DE INGRESO A LA EMPRESA	

Ica, ___ de _____ de 20__

REGISTRADO POR

FIRMA DEL EGRESADO



HUELLA DACTILAR

