

**ANÁLISIS DEL COMPORTAMIENTO LABORAL EN EL SECTOR TIC SEGÚN  
EL TIPO DE FORMACIÓN ACADÉMICA CULMINADA EN COLOMBIA**

**RONALD YESID MEDINA QUIROGA**

**Tutora**

**Isabel Andrea Mahecha Nieto**

**INSTITUCIÓN UNIVERSITARIA POLITÉCNICO GRANCOLOMBIANO**

**FACULTAD DE INGENIERÍA, DISEÑO E INNOVACIÓN**

**INGENIERÍA DE SISTEMAS**

**ENERO, 2021**

## LISTA DE CONTENIDO

1.	TÍTULO .....	5
2.	JUSTIFICACIÓN.....	6
3.	PLANTEAMIENTO DEL PROBLEMA.....	7
4.	INTRODUCCIÓN .....	8
5.	OBJETIVOS .....	9
5.1.	Objetivo General .....	9
5.2.	Objetivos Específicos.....	9
6.	LEVANTAMIENTO DE LA INFORMACIÓN.....	10
7.	MARCO TEÓRICO .....	11
8.	METODOLOGÍA .....	21
9.	DESARROLLO DE LA SOLUCIÓN.....	28
10.	ANÁLISIS DE LOS RESULTADOS .....	50
10.1.	Proyecciones nacionales.....	51
10.2.	Proyecciones departamentales.....	58
11.	VISUALIZADOR DE RESULTADOS.....	67
12.	CONCLUSIONES .....	69
13.	REFERENCIAS .....	72

## LISTA DE ILUSTRACIONES

Ilustración 1. Regresión lineal simple .....	13
Ilustración 2. Árbol de decisión. ....	15
Ilustración 3. Etapas de la metodología CRISP-DM.....	22
Ilustración 4. Formación de Colombianos en carreras TIC en diferentes niveles.....	29
Ilustración 5. Ficha de Asistencia Educativa.....	33
Ilustración 6. Población económicamente activa. ....	34
Ilustración 7. Número de inscripciones en Educación Superior. ....	34
Ilustración 8. Número de graduados en educación superior.....	35
Ilustración 9. Tasa de ocupación por nivel de estudios. ....	35
Ilustración 10. Población ocupada en relación a su último título académico.....	36
Ilustración 11. Oferentes registrados en portales de empleo.....	36
Ilustración 12. Estructura intermedia base de datos demanda laboral.....	37
Ilustración 13. Estructura final demanda laboral 2015-2020. ....	38
Ilustración 14. Estructura intermedia base de datos asistente cursos TIC.....	38
Ilustración 15. Estructura final base de datos asistente cursos TIC. ....	38
Ilustración 16. Estructura profesionales y técnicos trabajando en áreas TIC.....	39
Ilustración 17. Estructura empleados sin título formal trabajando en áreas TIC. ....	39
Ilustración 18. Base de datos inicial Graduados TIC. ....	40
Ilustración 19. Estructura graduados TIC.....	41
Ilustración 20. Formato establecido para RapidMiner .....	42
Ilustración 21. Modelo creado para el preprocesamiento de los datos.....	43
Ilustración 22. Modelo creado para el procesamiento de los datos. ....	45
Ilustración 23. Configuración Módulo Holt-Winters para los cursos de capacitación TIC. ....	46
Ilustración 24. Configuración Modulo Apply Forecast para los cursos de capacitación TIC.....	47
Ilustración 25. Estructura de la proyección Asistentes cursos TIC. ....	48
Ilustración 26. Estructura de la proyección de profesionales, tecnólogos y técnicos trabajando en el sector TIC.....	48
Ilustración 27. Estructura de la proyección de personal sin educación superior trabajando en el sector TIC.....	49

Ilustración 28. Comportamiento personas mayores de 18 años que asisten a cursos de educación TIC. ....	51
Ilustración 29. Comportamiento nivel de estudios personal TIC. ....	53
Ilustración 30. Distribución empleados TIC 2019. ....	55
Ilustración 31. Comportamiento nivel de vacantes y graduados TIC a nivel nacional. ....	56
Ilustración 32. Departamentos sin instituciones de educación superior. ....	59
Ilustración 33. Comportamiento nivel de vacantes y graduados en Antioquia. ....	60
Ilustración 34. Comportamiento nivel de vacantes y graduados en Bogotá D.C. ....	61
Ilustración 35. Comportamiento nivel de vacantes en Bolívar. ....	62
Ilustración 36. Comportamiento nivel de vacantes y graduados en Magdalena. ....	63
Ilustración 38. Comportamiento nivel de vacantes y graduados en Valle. ....	65
Ilustración 43. Interfaz inicial herramienta de visualización. ....	68
Ilustración 44. Interfaz comportamiento graduados y vacantes a nivel departamental. ....	69

### **LISTA DE TABLAS**

Tabla 1. Similitud entre CRISP-DM y KDD .....	23
---	----

## **1. TÍTULO**

Análisis del comportamiento laboral en el sector TIC según el tipo de formación académica culminada en Colombia.

## **2. JUSTIFICACIÓN**

En un mundo afectado por la aparición del COVID-19 y las medidas preventivas decretadas por el Gobierno Nacional se fortalece la posición estratégica de la selección de las TIC como el eje fundamental de la economía nacional, es por esta razón que se evidencia la necesidad del país de generar talento humano en relación a esta área de conocimiento específica, no solo por su gran aporte a la disminución de los indicadores de contagio al utilizarse para la realización efectiva de las tareas diarias, sino porque desde tiempo anterior, se observa que será un subsector que jalonará la economía mundial, y aprovechara su rápida innovación para mejorar la productividad de otras actividades comerciales.

Es por esta razón que este trabajo de grado tiene como objetivo realizar un análisis pormenorizado de la situación actual de los jóvenes colombianos en relación a la formación en Tecnologías de la Información y las Comunicaciones, y su incorporación al mercado laboral, entendiendo el eje de la investigación en las necesidades futuras de las empresas de la industria y la imperiosa necesidad de fortalecer los mecanismos de educación sobre el tema.

Con base en lo anterior, este proyecto se realiza en búsqueda de establecer las necesidades primordiales de la industria para dar respuesta a las nuevas políticas de fortalecimiento del área TIC que promueve el Gobierno Nacional ofreciendo herramientas que permitan conocer cuál es el tipo de conocimiento que se requiere para dar respuesta a las necesidades del mercado, al igual que establecer el impacto que se produce al presentar el fenómeno de formar jóvenes en campos de acción que no serán utilizados en su máxima

capacidad por las empresas nacionales. (Ministerio de Tecnologías de la Información y las Comunicaciones, 2020)

### **3. PLANTEAMIENTO DEL PROBLEMA**

En Colombia, la brecha de talento humano con conocimiento en Tecnologías de la Información y Comunicaciones crece con el paso del tiempo, las empresas se encuentran en búsqueda de profesionales que ofrezcan el conocimiento básico en programación y sus áreas afines, pero las instituciones de educación superior han visto disminuido el número de profesionales en áreas relacionadas con las TIC desde 2016 por primera vez desde su aparición. (Educacion Revista Semana, 2019)

Pero cual puede ser la razón para esta disminución de egresados, tal vez la evolución de las tecnologías que se utilizan para el desarrollo de aplicaciones hace que las carreras extensas cuya duración se promedian entre cuatro y cinco años se hacen no pertinentes para el mercado laboral de esta área, puesto que el tiempo que dura el proceso académico logra que las empresas demanden otras competencias, pero en este sentido, el mismo futuro esperan los estudiantes que seleccionan la profundización de conocimientos sobre una herramienta en particular, puesto que en su lugar deben iniciar nuevamente con su proceso académico enfocado a herramientas diferentes que requieren un tipo de conocimiento general para su rápido y eficaz entendimiento.

Es aquí, donde se presenta la disyuntiva sobre qué tipo de formación es la más adecuada para estos nuevos tiempos donde la tecnología es casi una obligación, y la innovación es parte del paradigma general, y en este contexto son las necesidades de la

industria y las exigencias del país quienes deben enfatizar en los perfiles necesarios para el desarrollo sostenido del subsector TIC.

#### **4. INTRODUCCIÓN**

En la actualidad el proceso académico de los jóvenes nacionales que muestran su preferencia hacia las áreas TIC se ve enfocado hacia dos corrientes diferentes en términos de costos y tiempo, por un lado, se presentan las carreras profesionales o tecnológicas de las áreas específicas de las tecnologías de la información y los cursos cortos enfocados hacia el conocimiento de herramientas específicas.

Para los dos casos las preferencias de los bachilleres nacionales para el momento de su elección están delimitados al conocimiento básico de programación y sus posibles campos de acción, por lo cual, las variables que más ejercen influencia en la elección son el costo de la formación, el tiempo de estudio y la inmediatez de ingresar al campo laboral. (SITEAL/TIC, 2019)

En este orden de ideas, sería acertado afirmar que el número de estudiantes que optan por los cursos cortos sería superior a las demás áreas de conocimiento, pero es posible afirmarlo con el solo análisis de las variables que sustentan su elección; también se debe entender que los bachilleres eligen estos cursos cortos como su preparación inicial para enfrentar la vida laboral, o seleccionan programas universitarios completos con el objetivo de obtener una preparación más fuerte antes de ingresar al mercado profesional.

Por medio del análisis de la información acertada sería posible encontrar cual es la mejor opción para el país en la selección del personal idónea de acuerdo a los puestos de

trabajo requeridos, o también establecer si las empresas realizan sus procesos de contratación enfocadas hacia el personal requerido, es interesante observar cómo estas respuestas pueden dar las señales efectivas sobre el comportamiento actual y el futuro a corto plazo del desarrollo de estas áreas de conocimiento en el país.

## **5. OBJETIVOS**

### **5.1. Objetivo General**

- Analizar el comportamiento laboral de los graduados en áreas correspondientes a ciencias TIC frente a las vacantes de la industria para identificar las estrategias que permitan suplir las necesidades de talento humano para el futuro.

### **5.2. Objetivos Específicos**

- Desarrollar un análisis exploratorio de los datos y obtener un informe estadístico sobre la situación de oferta y demanda de recurso humano en el sector de las TIC en Colombia.
- Generar la base de datos que permita el almacenamiento y la extracción de la información para su procesamiento.
- Implementar un modelo de minería de datos para procesar los conjuntos de datos con el fin de generar modelos predictivos, análisis de comportamiento y tendencias.
- Implementar una herramienta de visualización de información interactiva que permita la generación de diferentes reportes.

## 6. LEVANTAMIENTO DE LA INFORMACIÓN

Con el objetivo de conseguir la información correspondiente para el inicio de la investigación es importante definir las organizaciones de tipos privada y pública que se encargan de la recopilación de los datos correspondientes al tema de la investigación:

- Observatorio laboral del Ministerio de Educación. Información disponible en:  
<https://www.mineducacion.gov.co/sistemasinfo/Informacion-a-la-mano/212400:Estadisticas>
- Departamento Administrativo Nacional de Estadística. (DANE). Bases de datos disponibles en: <https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion>
- Ministerio de Tecnologías de la Información y las Telecomunicaciones (MINTIC). Estadísticas disponibles en: <https://colombiatic.mintic.gov.co/679/w3-propertyvalue-36342.html>
- Datos Abiertos. Portal Publico de información estadística, con información disponible en:  
<https://www.datos.gov.co/browse?q=Educacion%20Profesional&sortBy=relevance>
- Informe estadístico de oferta laboral del servicio público de empleo, a disposición en:  
<https://www.serviciodeempleo.gov.co/estudios-e-investigacion/oferta-y-demanda-laboral/anexo-estadistico-de-oferta-laboral-buscadores>
- Ministerio de Trabajo Nacional por medio de su herramienta de Fuente de Información Laboral en Colombia (FILCO, obtenida desde:  
<http://filco.mintrabajo.gov.co/FILCO/faces/estadisticas.jsf>
- Estudio de identificación de brechas de capital humano para el sector TIC. Alianza TIC, Ministerio del trabajo, Ministerio de las TIC, Ministerio de Educación Nacional,

SENA y mesa sectorial, a disposición en:  
<https://herramientas.datos.gov.co/sites/default/files/ESTUDIO%20DE%20BRECHA%20SECTOR%20TIC.pdf>

## 7. MARCO TEÓRICO

### **Descubrimiento del conocimiento en bases de datos (KDD)**

Es entendido como el proceso de encontrar patrones válidos, novedosos y comprensibles que permitan entender el comportamiento de un grupo de datos, estos datos son necesariamente un conjunto de hechos verificables, que para el caso de estudio son bases de datos presentadas por entidades del Gobierno nacional o entidades comerciales reconocidas, y el patrón que se busca se presenta como un término definido para una porción de los datos que pueden ser caracterizados por medio de un modelo que tiene por lo menos una característica en común. (Gorbea Portal & Madera Jaramillo, 2017)

Cuando se define este concepto como proceso se subdivide en diferentes pasos, en referencia al término analizado estos pasos son: recolección de datos, preparación, búsqueda de patrones, verificación del conocimiento y preparación, todos continuos por medio de múltiples iteraciones. Este tipo de proceso KDD se realiza de una búsqueda o inferencia, pues su conocimiento no se realiza por medio de cálculos u operaciones previamente definidas de las cuales se obtiene un resultado deseado, estos resultados se obtiene por medio de las observaciones y determinaciones del realizador, por supuesto estas conclusiones deben verificarse como válidas en otros datos similares con un alto grado de confiabilidad. (Flores Lagla, Cadena Moreano, & Villa Quisphe, 2019)

## **Minería de datos**

Esta ciencia tiene sus inicios en la década de los 60 tiene como función extraer información valiosa de grandes volúmenes de datos, tiene sus orígenes en la estadística y sus procesos de análisis de datos sin previa hipótesis.

Desde su aparición ha presentado cambios en sus nombres desde arqueología de datos, extracción de conocimiento entre otros, su función principal es el descubrimiento de nuevas y significativas relaciones, tendencias y patrones de una gran cantidad de datos.

Tiene sus orígenes en la actividad rutinaria de las organizaciones de acumular información, toda compañía se encarga de generar a diario mucha información de sus diferentes modelos de negocio, pero cómo es posible elegir qué información es vital y cual no, la minería de datos persigue el descubrimiento automático de la información importante y requerida para la creación de contenido, para lograrlo se utilizan diferentes técnicas y herramientas con capacidad de procesar grandes cantidades de datos, para ser analizados por medio de diferentes algoritmos que se encargan de realizar las tareas adecuadas sobre el conjunto de datos.

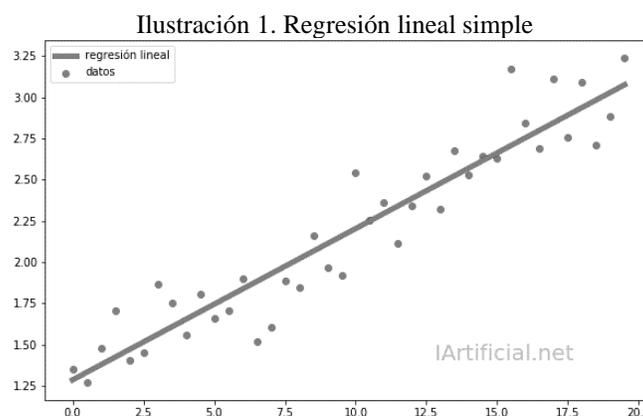
Para desarrollar un proyecto de minería de datos es necesario su división en fases, estas subrutinas son de forma general, las siguientes, selección de datos, preprocesamiento de los datos, transformación de datos, selección de patrones y modelos, evaluación de la nueva información. (Romero Romero, 2018)

## **Técnicas en la minería de datos**

Para obtener excelentes resultados en el proyecto es imperativo, seleccionar la técnica más adecuada en referencia al tipo de proyecto y los resultados necesarios, con el

objetivo de seleccionar la más adecuada es imperativo entender que existen técnicas predictivas, descriptivas y auxiliares, en razón al objetivo de la investigación y por la búsqueda del pronóstico de diferentes variables independientes se hace necesario presentar las técnicas más importantes de tipo predictivo, así como los algoritmos que son más utilizados para esta misión, a continuación se exponen las técnicas más utilizadas en la actualidad:

Regresión lineal, es la técnica más utilizada cuando existe la necesidad de establecer relaciones entre los diferentes datos, por esta característica se pueden establecer relaciones entre datos numéricos y presentar modelos de regresión, esta técnica permite predecir a partir de un muestreo aleatorio de datos, puesto que su característica principal es la determinación de relaciones entre diferentes variables, la precisión de esta técnica depende en gran valor a la cantidad de datos que se manejen, es decir, que su precisión es directamente proporcional al número de datos que se manejen (Uvidia Fassler, Cisneros Barahona, Méndez Naranjo, & Villa Yáñez , 2017), a continuación se observa en la ilustración No. 1 una gráfica correspondiente a la regresión lineal simple.



Nota. Adaptado de IArtificial.net, regresión lineal – Error cuadrático (MSE), por Martínez Heras, José.

Reglas de asociación, como su nombre lo indica esta técnica busca la aparición de asociaciones entre los resultados y un grupo de datos, es decir se evalúan primero el soporte o fuente (datos) y se contrasta con los resultados obtenidos en búsqueda de verificar la confianza de los resultados, esta técnica se sustenta en la ocurrencia de ciertos eventos dentro de un conjunto de datos que producen un resultado común. (Ruiz & Myriam, 2019)

Agrupamiento (Clustering), esta técnica permite entender como cierta información agrupada o caso, pertenecen a diferentes clases y grupos lo que significa que comparten algunas características, desde allí el análisis se realiza desde otra nueva agrupación con el objetivo de encontrar relaciones no verificadas con los datos iniciales. (Flores Lagla, Cadena Moreano, & Villa Quisphe, 2019)

Redes neuronales, este tipo de redes que tienen su disposición en base con la distribución del cerebro humano y su capacidad de aprendizaje se fundamenta en el pasado y las experiencias anteriores, que se utilizan para la resolución de nuevos problemas.

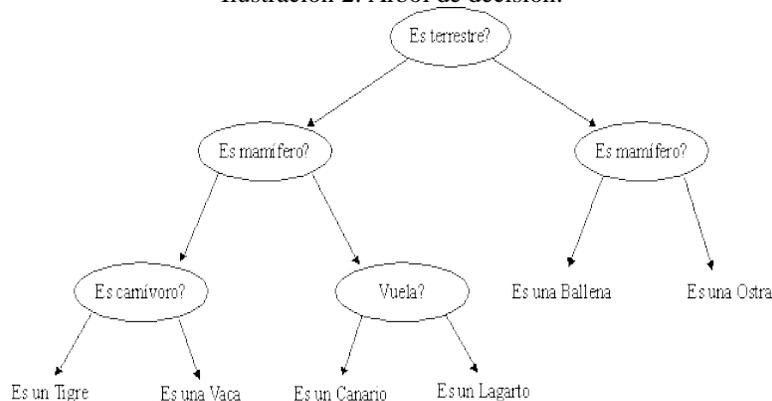
Este tipo de técnica es capaz de aprender y detectar patrones y características comunes en los datos y con la particularidad de hacer previsiones y calificaciones una vez se encuentran completamente implementadas. Esta técnica en razón a su capacidad de aprendizaje, la sencillez en su utilización, la capacidad de adaptación ante una realidad muy cambiante es utilizada para el reconocimiento de imágenes y la búsqueda específica en amplias cantidades de ubicaciones y datos. (Pincay Ponce, Angulo Murillo, Herrera Tapia, & Delgado Muentes, 2020)

Arboles de decisión, esta herramienta analítica tiene como objetivo el descubrimiento de reglas y relaciones, por medio de la representación de los datos en

diagramas de flujo, que para esta técnica se presenta con forma de árbol pues los nodos internos hacen representación a las diferentes validaciones y sus ramas son las salidas desde los nodos anteriores por cada una de las posibles rutas que tienen como salida la respuesta a la afirmación anterior, esta representación se presenta a continuación en la ilustración No. 2 árbol de decisión.

Estos árboles son utilizados para la predicción por medio de condiciones sucesivas en base con la limitación de las opciones de respuesta y las condiciones necesarias para ofrecer las medidas correctas de acuerdo con el escenario. (Camana, 2016)

Ilustración 2. Árbol de decisión.



Nota. Adaptado de Informática UV, Practica 6, ejemplo de árbol de decisión. Guía 1, practica 6.

Forecast: este método tiene como finalidad la estimación y análisis de un grupo de datos agrupados como una serie de tiempo con el objetivo de pronosticar el comportamiento de la variable de interés, el método utiliza diferentes algoritmos estadísticos para realizar el análisis y el pronóstico de los datos, este tipo de pronóstico depende del modelo proporcionado, este método tiene la particularidad de permitir la selección del número de periodos que se desean pronosticar. (Uvidia Fassler, Cisneros Barahona, Méndez Naranjo, & Villa Yáñez , 2017)

Algoritmo K-means: este algoritmo se utiliza para el análisis de diferentes grupos, su funcionamiento es el siguiente, divide los datos recogidos en diferentes grupos, en los cuales se agrupan en razón a las características comunes, con lo cual se busca encontrar información que, aunque se encuentra en los datos no se tiene en cuenta pues no datos no se encuentran agrupados en relación a sus diferentes clústeres.

Algoritmo EM: este algoritmo define automáticamente un parámetro de acuerdo al análisis de los datos que componen la data, y en razón a este predice la probabilidad de salida de un evento aleatorio dentro de los parámetros de los datos.

Algoritmo AdaBoost: normalmente este algoritmo funciona dentro de otro algoritmo que anticipa un comportamiento de acuerdo con los datos observados, con la diferencia que tiene en cuenta los extremos estadísticos, se utiliza para crear un mayor porcentaje de efectividad de las predicciones. (Gutierrez O. & Molina , 2018)

Algoritmo Holt-Winters: este algoritmo es la versión mejorada del enfoque del suavizamiento exponencial, tiene la ventaja de ser fácil de adaptarse a medida que nueva información real está disponible, se utiliza para realizar modelos de pronóstico con triple suavizado exponencial para los componentes de nivel, tendencia y estacionalidad, esto se logra con la definición de tres parámetros Alpha, beta y gamma, Alpha especifica el coeficiente que se utilizara para el suavizamiento de nivel, beta es el coeficiente para el suavizado de tendencias y gamma se utiliza para representar el suavizado estacional. Este tipo de procedimientos reducen los cálculos y requerimientos en términos de almacenamiento en las bases de datos, un factor de alta importancia cuando se están prediciendo muchas series de tiempo. (Gonzalez Farroñan, 2018)

## **Tecnologías de la información y comunicación (TIC)**

La concepción de este término data de la aparición de definiciones como nuevas tecnologías, tecnologías de aprendizaje y conocimiento entre otros, la definición de estos términos se debe entender como el conjunto de las tecnologías que permiten la adquisición, producción, comunicación, tratamiento y registros de información con estilo de voz, imágenes y contenidos en los diferentes tipos de señales.

Esta definición cambia constantemente al mismo tiempo que aparecen nuevas tecnologías que tienen como fin el tratamiento de la información, por eso este entendido ha cambiado múltiples veces desde su aparición en 1985 hasta la fecha, pero el objetivo es el mismo, el proceso de almacenamiento de los datos para ser utilizados por las organizaciones y las personas para su beneficio propio o general.

Las TIC comparten varias características, inmaterialidad, que se refiere a que el insumo de las tecnologías no es material, pues siempre esta materia prima es la información, un conjunto de datos que permite su almacenamiento, consulta, modificación etc.

Interactividad, pues las tecnologías de la información y la comunicación permiten a los usuarios una interactividad de los usuarios con la información de manera total, la instantaneidad de la información, un factor que rompe con la distancia entre empresas y naciones. (Grande, Cañon , & Cantón , 2016)

### **Alianza TIC colombiana**

Desde el año 2018 el país vio la aparición de una alianza publica de diferentes organismos que buscan la realización de diferentes estudios que buscan identificar la

demanda y oferta de capital humano en el sector TIC, así como la presentación de estrategias que buscan aumentar la fuerza laboral en esta industria.

Esta alianza está compuesta por el SENA, el Ministerio de las telecomunicaciones y comunicaciones, El Ministerio de Educación Nacional, la Mesa Sectorial de gestión de Tecnología y Talento Digital, de su implementación se lograron sincronizar acciones para elaborar de manera conjunta diferentes estudios donde se identificaron las principales brechas de capital humano y la prospectiva laboral del sector, así como las tendencias que marcarán el futuro de la industria en lo correspondiente con las nuevas necesidades de competencias de tipo general y específicos para el empleo en los próximos años.

De los estudios realizados se pueden presentar de forma alarmante que de un total de 75 cargos TIC identificados existen brechas en todos los campos de conocimiento, siendo la totalidad los siguientes, un 52% en los programas educativos profesionales, 31% para el posgrado y un 17% en programas técnicos y tecnológicos. (Ministerio de Tecnologías de la Información y las Comunicaciones, 2020)

También es fundamental resaltar que en la actualidad las competencias técnicas que son más solicitadas por las empresas del subsector TIC son análisis de sistemas de información, lenguajes de programación y SQL, PHP y Angular, lo que concluye que en el desarrollo de software la mayor cantidad de competencias son requeridas por las empresas nacionales y las Instituciones de Educación Superior no están proveyendo el conocimiento específico más necesario para los estudiantes de las carreras afines. (Ministerio de las Tecnologías de información y la telecomunicaciones, 2019)

## **Nuevos paradigmas de contratación**

En esta nueva economía del conocimiento es importante para todos los Gobiernos adaptar el perfil de egreso de los colegios secundarios y las universidades a las demandas del mundo actual, en este sentido la mayor preocupación en la actualidad es la modificación en los procesos productivos generados por la revolución informática y el incesante cambio tecnológico, para atender esta preocupación los gobiernos y las empresas deben entender que el esfuerzo debe estar mancomunado entendido en recuperar la inversión de estos actores en el sistema educativo con la graduación de capital humano a nivel, técnico, tecnológico y profesional renovado y general, con los conocimientos mínimos, habilidades y actitudes para enfrentar el desenfrenado avance tecnológico y la globalización que se presenta en los periodos actuales. (Silva Montes , 2018)

Las economías y sociedades actuales basan su funcionamiento y éxito en el tratamiento del conocimiento, en este orden de ideas, deben buscar capital humano que permita aprender de su operación, almacenamiento y clasificación, con el objetivo de no retrasar los indicadores macroeconómicos de las diferentes industrias, el tipo de trabajo para el futuro estará enfocado en el saber interpretar antes del saber hacer, porque seguramente el hacer será modificado para que sea realizado por diferentes dispositivos que crean un menor costo para las organizaciones.

## **Implementación de la política de Gobierno Digital**

En lo referente a Gobierno Digital, Colombia establece una política de Estado que ha venido evolucionando en lo referente a su alcance e implementación, todo lo anterior en base con la conclusión y el reconociendo para el uso de las Tecnologías de la Información y

las Comunicaciones, como el pilar fundamental para el mejoramiento de la gestión pública y la relación del Estado con los ciudadanos. No por nada se ha creado, la estrategia de Gobierno en línea, y de su éxito total, el ejecutivo presenta una nueva propuesta denominada Política de Gobierno Digital, con el objetivo de utilizar la tecnología para el monitoreo de todas y cada una de las entidades del estado y la presentación de propuestas innovadoras que permitan utilizar la tecnología para reducir los tiempos de los procesos y la aumenten la productividad en cada organización. (MinTIC, 2020)

Pero la anterior revolución no para allí, el Gobierno observando el panorama le ha apostado al incentivo de la educación y el emprendimiento en el área de las tecnológicas de la información y comunicaciones, teniendo en cuenta las próximas necesidades de las organizaciones y la modernización actual, anticipándose a los cambios que las economías emergentes deben aplicar para no quedar rezagado en el mercado internacional, posicionando el país como una gran posibilidad para recibir la inversión extranjera y aumentar la probabilidad de exportar servicios de tipo informático, preparando el camino ha impulsado varias alternativas como Talento Digital, enfocada en la formación de técnicos, tecnólogos y profesionales para las áreas TIC con créditos condonables a través de aplicaciones específicas que fortalezcan las instituciones públicas, pero además estrategias como Semilleros SEED certificados enfocados a ciudadanos de región que quieran aprender a validar su idea de negocio a través de metodologías ágiles, Spinit, programa de acompañamiento para que emprendedores digitales vendan más y consoliden su mercado, entre otros que buscan la consolidación de los emprendedores del sector.

Por lo anterior, es importante establecer si el esfuerzo del Gobierno tiene el impacto deseado en la academia, y si los estudiantes de educación superior pueden abastecer las necesidades de este subsector de la economía, con lo cual se articularían los esfuerzos entre academia, gobierno y comercio para fortalecer la industria con mayor potencial a futuro. (Ministerio de Tecnologías de la Información y las Comunicaciones , 2018)

## **8. METODOLOGÍA**

El proyecto es una investigación prescriptiva de tipo aplicada que se inicia con una revisión documental de las diferentes bases de datos de las entidades públicas y privadas reconocidas que manejan las variables de interés para la investigación, número de egresados de las instituciones de educación superior del país enfocados hacia el área TIC, así como las personas que optan por la realización de cursos cortos de diferentes plataformas de educación, al igual que las tendencias de las empresas del subsector en relación a la necesidad académica de sus diferentes puestos de trabajo, también se presenta información correspondiente a estudios sectoriales que buscan la identificación de las diferentes brechas de conocimiento en el área TIC, estas variables en general relacionadas con el paso del tiempo, con el objetivo de identificar las modificaciones que se han presentado de un periodo a otro y entender las nuevas necesidades del mercado.

Por último, para esta fase se establecen las políticas públicas para el área TIC a nivel nacional y la normatividad relacionada con la promoción e incentivo para los estudiantes y emprendedores de la industria.

Con la información anterior se realiza el proceso de minería de datos por medio de la técnica CRISP-DM (Cross-Industry Standard Process for Data Mining), para obtener la comprensión del problema, la exploración de los datos, preparación de la información, transformación de los datos, modelado de la información y la evaluación del conocimiento.

Métodos que permitirán determinar si las necesidades de recurso humano actuales y a futuro para las empresas del sector TIC deben focalizarse en el aprendizaje de tecnologías informáticas independientes o la búsqueda de personal en el área informática de tipo tecnológico y profesional.

En este aparte se presentan las consideraciones para cada una de las etapas de la metodología CRISP-DM, el detalle de cada uno y la selección de los modelos y algoritmos de minería de datos utilizados para la obtención de resultados.

La metodología seleccionada propone el desarrollo de todo proyecto de minería de datos a través de diferentes etapas, en las que cada una de estas implementa tareas específicas con el fin de cumplir con los objetivos previamente establecidos.

Estas etapas se desarrollan de forma cíclica e iterativa, con la característica de estar en capacidad de devolverse a la etapa anterior cuando sea necesario.

Ilustración 3. Etapas de la metodología CRISP-DM



Nota. Adaptado de publicaciones Research Gate, Herramienta HERMINWEB, por Ordoñez, Yoanni y Horacio, Darian.

De acuerdo con la ilustración No. 3 etapas de la metodología CRISP-DM, los pasos que se siguen en la metodología se pueden resumir en los siguientes: comprensión del negocio, comprensión y preparación de los datos, modelamiento y evaluación, para finalizar con el despliegue del proyecto.

Con estas etapas se evidencia claramente que las etapas de esta metodología son concordantes con el proceso KDD, agrupan en dos de las etapas de CRISP varias capas del modelo KDD, pero manteniendo la misma estructura con se presenta en la tabla No. 1 similitud entre CRISP-DM y KDD.

Tabla 1. Similitud entre CRISP-DM y KDD

<b>KDD</b>	<b>CRISP-DM</b>
Identificación del problema	Comprensión del negocio
Selección e integración de los datos	Comprensión de los datos
Preprocesamiento de los datos	
Transformación de los datos	Preparación de los datos
Selección e implementación de la minería de datos	
Interpretación y evaluación	Modelamiento y evaluación
Post KDD	Despliegue del proyecto

Fuente: Elaboración propia.

Aunque si bien las metodologías son muy parecidas es importante establecer que KDD es el primer modelo aceptado por la comunidad científica para la extracción y explotación de información, y se basa en la minería de datos como la etapa donde se realiza la extracción de patrones a partir de un grupo de datos inicial, pero no al proceso completo de descubrimiento de conocimiento, con lo anterior se puede establecer que el método CRISP profundiza en mayor detalle las tareas y actividades que se deben realizar en cada etapa de proceso de minería de datos y no solo una idea general del modelo, tal vez el mayor beneficio de la metodología CRISP es la comprensión del negocio, pues no solo se entiende el problema que se presenta sino que se investiga el contexto en que se presenta,

se busca el entorno amplio de las circunstancias que los causan y la delimitación de sus consecuencias, con lo cual se entiende no solo el problema como en KDD sino que se aborda información más amplia de las características del problema mismo y su ambiente.

Teniendo en cuenta lo anterior, y además la flexibilidad de la metodología CRISP, es la seleccionada para la realización de la investigación pues, es posible entenderla como un conjunto de tareas y actividades específicas que si bien tienen un orden preestablecido y una jerarquía, la metodología en sí misma no explica cómo desarrollarlos y obliga a la realización en un orden determinado, generando alta flexibilidad en los procesos y aumentando la velocidad de todos los procesos metodológicos, pues si bien la metodología se divide en fases esta sucesión no es completamente rígida, pero además las fases se descomponen en tareas generales, que están compuestas por tareas específicas, pero esta especialización en tareas no propone un modelo para realizarlas ordenadamente.

A continuación, se explican las etapas que componen la metodología seleccionada.

### **Comprensión del negocio**

Esta etapa se enfoca en el conocimiento del problema, esto se entiende como la definición del problema del negocio, con el objetivo de entender los requisitos que tendrá el proyecto en general. De esta clara definición se desprenden decisiones con la misma importancia como la selección de los algoritmos utilizados de las diferentes presentadas en la actualidad.

Por lo anterior, y teniendo en cuenta el auge de los nuevos modelos de enseñanza enfocados en la enseñanza de nuevas tecnologías individuales, recae la importancia de

establecer las necesidades profesionales del país en términos de talento humano, realizando el análisis del comportamiento académico y laboral del país para establecer el tipo de educación necesaria para enfrentar los nuevos retos laborales del país.

Con las características expuestas se puede determinar que el problema cuenta con las características de un proceso de negocio, pues tiene un conjunto de acciones y decisiones que son necesarios para presentar resultados que permitan seleccionar el método académico más adecuado para el futuro de los estudiantes nacionales.

Los desafíos que presenta esta etapa son los siguientes, la capacidad económica de los estudiantes, las necesidades del mercado laboral actual, los diferentes modelos académicos y las competencias específicas necesarias para el futuro de la industria.

### **Comprensión de los datos**

Con los resultados de la etapa anterior, es decir, con la definición del problema y los desafíos del negocio, se inicia con la búsqueda de la información correcta para realizar el proceso de la minería de datos, esta etapa inicia con la búsqueda de la información relacionada con las variables necesarias para la generación de los indicadores del proceso.

Cuando se seleccionan estos indicadores la etapa comprende la selección de las fuentes confiables de estadísticas y la recolección inicial de los datos, para posteriormente establecer las relaciones entre estos.

Esta etapa presenta la mayor demanda de tiempo en la realización del proyecto, las principales tareas de son:

Recolección de datos, estableciendo las fuentes desde donde se obtienen.

Descripción de los datos, en esta actividad se definen los volúmenes de la información y los formatos en que se encuentran.

Exploración de los datos, presentado una estructura para la información y comprobando la frecuencia y la distribución de todos los datos.

Verificación de los datos, en esta tarea se determina la calidad de los datos, comprobando su consistencia, la existencia de datos nulos e irregularidades en la exactitud.

### **Preparación de los datos**

Cuando se finaliza la etapa anterior se tiene la información necesaria y se ha analizado y validado completamente, es decir se tiene la certeza de contar con datos de calidad para el análisis, en este instante se deben preparar para adaptarlos a las técnicas de minería de datos seleccionada.

Esta etapa se relaciona con el modelamiento y la aplicación de las técnicas de minería de datos, con relación a que según la técnica seleccionada es necesario procesar la información de una determinada manera.

De acuerdo con la técnica los pasos a seguir son los siguientes:

Selección de datos, ubicando un subconjunto de los datos recolectados de la etapa anterior.

Limpieza de datos, se realiza la preparación para el modelado, por lo general por medio de técnicas de normalización, valores nulos etc.

Estructuración de datos, generando nuevos atributos o transformando con los que se cuenta.

Integración de los datos, agrupando la información que tiene una característica en común por medio de una estructura que las pueda contener.

Formateo de los datos, transformando los datos con el fin de ajustarlos para las técnicas seleccionadas.

### **Modelamiento**

Como el nombre lo indica en esta etapa se seleccionan las técnicas de modelado más apropiadas para el proyecto de acuerdo con el problema que se debe resolver, esto se realiza de acuerdo a la información que ya está procesada y validada de acuerdo a los procesos de la etapa anterior.

Las variables para tener en cuenta en la elección de este modelado se deben tener en cuenta los requisitos del problema que se encuentra en estudio y el conocimiento previo del funcionamiento de la técnica, pues esto genera un punto a favor debido a que se conocen los parámetros que se deben ajustar.

Los pasos que se realizan en esta etapa son:

Selección de la técnica: esta selección se realiza de acuerdo con el tipo de problema y el objetivo principal del proyecto.

Construcción del modelo: de acuerdo con la técnica seleccionada se aplica a los datos recopilados para la generación de uno o más modelos.

Evaluación del modelo: se realiza esta evaluación con la interpretación de estos modelos en base al conocimiento previo y los criterios de éxito.

Evaluación de los resultados: evaluándolos con los objetivos del proyecto.

### **Despliegue del proyecto**

De todas las etapas de la metodología se obtiene conocimiento, con esta información y el modelo ya validado, se establecen recomendaciones para el proceso de negocio.

Se debe destacar que el proyecto de minería de datos debe tener bien documentados sus resultados, pues el objetivo de este tipo de proyectos es la creación de conocimiento a partir de nueva información obtenida.

## **9. DESARROLLO DE LA SOLUCIÓN**

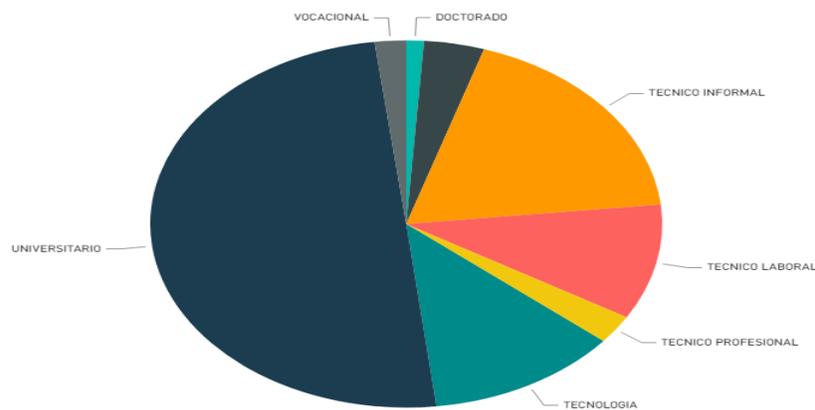
### **Comprensión del negocio**

En esta etapa del proyecto se define el problema fundamental de la investigación, observando el panorama laboral en el país y las políticas nacionales es evidente la necesidad de suplir varias vacantes en el corto y mediano plazo referenciadas hacia las tecnologías TIC, esto en constante crecimiento debido a la aparición del fenómeno de salud pública actual permiten inferir un contexto que vislumbra la necesidad imperiosa de formar empleados con las competencias necesarias para abordar los problemas tecnológicos que las empresas de las diferentes industrias necesitan.

En este orden de ideas, existen dos vertientes académicas actuales que permiten fortalecer el capital humano nacional en aras de suplir las vacantes necesarias para la transformación digital de las empresas, la primera, el ámbito general compuesto por diferentes programas técnicos, tecnólogos y profesionales de las áreas de la información, y segundo, la aplicación de programas cortos enfocados a diferentes paradigmas y lenguajes de programación que permiten la rápida inclusión del estudiante en el ámbito laboral.

Según el Ministerio de las Tecnologías de la Información y las Comunicaciones nacional en el sector TIC es evidente que la preparación superior es muy importante para el ingreso a la industria, en este orden de ideas según las estadísticas de la cartera ya mencionada es muy poco el personal que se presenta en las empresas TIC son conocimientos de técnicos informales a continuación se presenta un gráfico correspondiente a la distribución para el periodo 2018 del nivel de formación en la industria.

Ilustración 4. Formación de Colombianos en carreras TIC en diferentes niveles.



Nota. Adaptado de DANE, formación de los nacionales en carreras TIC para diferentes niveles, por Estadísticas del Ministerio de las Tecnologías y la información.

De acuerdo con las estadísticas de estos números las personas que cursan diferentes estudios de tipo superior en áreas TIC estaban para 2018 distribuidos de la siguiente forma, universitario 49.91%, tecnológico 12.21%, técnico laboral 10.17%, técnico profesional 2.53% y técnico informal el 18.39%. (Ministerio de las Tecnologías de la Información y las Comunicaciones, 2018)

Para alcanzar el objetivo de la investigación es necesario utilizar la información pública de los organismos interesados en los ámbitos académicos e industriales con el objetivo de presentar un sistema de información que presenta a la población objetivo la información específica con referencia al ámbito laboral TIC del país segmentado al tipo de educación

recibida y las competencias desarrolladas durante el proceso académico. En esta etapa se encuentran diferentes bases de datos e información con un alto grado de desorden y características no adecuadas para el desarrollo del proyecto, sin embargo, presentan información muy valiosa que debe ser organizada para hacer parte de la investigación.

### **Comprensión de los datos**

En esta fase se establece la base fundamental de todo el sistema, los datos sujetos de análisis y exploración, para este estudio es importante identificar datos de fuentes confiables que permitan obtener resultados fiables, estos deben tener su fundamento en organizaciones reconocidas y estadísticas nacionales.

Durante esta fase se encuentran múltiples falencias, iniciando con la contradicción de datos de la misma índole seguramente en razón al ámbito revisado, así como cifras parcializadas de acuerdo con el organismo emisor, después de la organización y verificación de los datos se encuentra información muy valiosa que aporta a la investigación en frentes como las necesidades laborales futuras, el número de graduados actuales y los perfiles de búsqueda más solicitados por parte de las empresas.

La información recolectada para la utilización de la investigación es la siguiente:  
Sistema Nacional de Información de la Educación Superior

El nuevo sistema de consulta interactiva del Ministerio de Educación es de dominio público y cuenta con toda la información necesaria para establecer el panorama académico de la educación superior. Para el análisis se utilizan las estadísticas de las personas que fueron admitidas a las diferentes IES pues son muchos los inscritos pero los verdaderos datos de personas en formación hacen parte de los admitidos por las

instituciones. De esta base de datos se resaltan las variables de nivel nacional y regional del número de graduados en relación con los periodos 2010 a 2019, así como la fecha de presentación de los datos.

Este organismo público de estadística ofrece la información correspondiente a los principales indicadores del mercado laboral de la población colombiana de acuerdo con aspectos relacionados con su educación, como la asistencia a una institución educativa y el título logrado. De esta fuente de información las variables con mayor importancia son el número de graduados en los diferentes niveles de educación, técnico, tecnológico y profesional TIC, y las fechas de graduación correspondientes.

También se utiliza la encuesta realizada en términos de la dinámica de formación para el trabajo, sus principales características y la participación en procesos de certificación en competencias de la población de 15 años y más, en esta base de datos es importante destacar la variable correspondiente al número anual de graduados en todos los cursos de educación TIC y la fecha de graduación correspondiente.

Plataforma publica datos.gov

En esta plataforma se presenta información correspondiente a las matrículas en estadísticas de Educación Superior por nivel de formación y de acuerdo con los diferentes municipios, además del número de instituciones educativas con ofertas activas, de esta fuente de información es importante ubicar la variable correspondiente al número de matriculados en áreas TIC en la IES de los diferentes municipios, y el periodo correspondiente a las matrículas.

### Portal público del servicio de empleo

Este portal público tiene información estadística de los buscadores de empleo de todas las regiones del país, así como información relacionada con la edad, nivel educativo, área ocupacional, experiencia laboral, aspiración salarial, entre otras. De esta base de datos se determinará como variable principal el número de oferentes y vacantes para el sector TIC en los diferentes periodos desde 2014 a 2020.

### Fuente de Información Laboral de Colombia

Herramienta suministrada por el Ministerio del Trabajo para presentar las estadísticas correspondientes a la cartera, en este portal se utiliza la estadística enfocada a la población que termina un programa académico de una institución de educación superior, lo que sugiere un indicador que muestra la oferta de trabajadores calificados en el territorio. De esta base de datos es importante determinar el número de graduados en áreas TIC con respecto al fin de los diferentes periodos desde 2010 a 2019.

### Informe estudio sectorial Alianza TIC

De este informe se extrae la información correspondiente a los datos sobre la identificación de las brechas de capital humano para el sector TIC con enfoque en la explotación de datos y prospectiva, así como la identificación de las competencias generales y específicas para el sector a partir de un estudio realizado en diferentes municipios del país a través de encuestas individuales con empresas de diferentes tamaños del sector TIC a nivel nacional.

### **Preparación de los datos**

La preparación de los datos se inicia con la selección de la información más amplia de las bases de datos seleccionadas, en este sentido de siete (07) bases de datos previamente seleccionadas se han eliminado dos (02) teniendo en cuenta que la información es similar, y seleccionando los datos de mayor completitud y fiabilidad, por estas razones se han eliminado del análisis las bases de datos del DANE en lo correspondiente a la formación para el trabajo, puesto que los indicadores son muy generales y observando la base de datos la ficha más específica divide la población nacional de acuerdo a su nivel de estudio, datos que se encuentran más específicos en otra fuente de información, en la ilustración No. 4 ficha de asistencia educativa se presenta la distribución del archivo correspondiente al módulo de formación para el trabajo a nivel nacional.

Ilustración 5. Ficha de Asistencia Educativa

Total nacional		2013	2014	2015	2016	2017	2018	2019
Concepto		Abr - jun						
Población que asiste o asistió		3.305	3.218	3.324	3.615	3.503	3.495	3.734
Ninguno		122	135	121	126	113	116	111
Educación básica primaria		418	382	432	405	400	367	416
Educación básica secundaria		224	202	221	190	203	188	169
Educación media		1.270	1.250	1.288	1.560	1.488	1.498	1.632
Educación superior <sup>^</sup>		1.254	1.238	1.251	1.334	1.298	1.327	1.407
No informa		18	11	12	1	0	0	0

Fuente: DANE, Gran Encuesta Integrada de Hogares (GEIH) - Módulo de Formación para el Trabajo.  
 Nota: <sup>^</sup> Las categorías de educación son elaboradas a partir de la Clasificación Internacional Normalizada de la Educación adaptada para Colombia (CINE – 2011 A.C.). Esta fue establecida y adoptada mediante la Resolución 776 de 2015 del

Nota. Adaptado de Estadísticas Departamento Nacional de Estadística, Modulo de formación para el trabajo, por DANE.

Así como la base de datos correspondiente a la población económicamente activa que asiste a los diferentes niveles educativos del país, teniendo en cuenta que la información es irrelevante para los intereses y la determinación de resultados de la investigación, también se encuentran fuentes de información más fiables y con un mayor número de indicadores que proveen valor a la investigación.

En la ilustración No. 5 población económicamente activa, se evidencia la información que presenta el módulo correspondiente a esta población y se observa que presenta información innecesaria para la investigación.

Ilustración 6. Población económicamente activa.



Total	Concepto	Año									
		2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
% población en edad de trabajar		82,9	83,0	83,2	83,5	83,7	83,9	84,1	84,3	84,4	
TGP		62,7	63,7	64,5	64,2	64,2	64,7	64,5	64,4	64,0	63,3
TO		55,4	56,8	57,8	58,0	58,4	59,0	58,5	58,4	57,8	56,6
TD		11,8	10,8	10,4	9,6	9,1	8,9	9,2	9,4	9,7	10,5
Población total mayores de 3 años		41.890	42.443	42.983	43.506	44.006	44.528	45.018	45.588	46.081	46.621
Población en edad de trabajar		34.706	35.248	35.781	36.307	36.827	37.342	37.851	38.355	38.861	39.355
Población económicamente activa		21.777	22.446	23.091	23.292	23.654	24.173	24.405	24.697	24.863	24.902
Ocupados		19.213	20.020	20.696	21.048	21.503	22.017	22.156	22.383	22.457	22.287
Desocupados		2.564	2.426	2.394	2.243	2.151	2.156	2.249	2.314	2.406	2.615
Inactivos		12.929	12.802	12.690	13.015	13.172	13.169	13.446	13.658	13.998	14.453

Nota. Adaptado de Departamento Nacional de Estadística, por población económicamente activa, DANE Estadísticas.

En este sentido, las bases de datos que serán fuentes de información para el posterior análisis son las siguientes:

Número de inscripciones en programas de educación superior del Ministerio de Educación Nacional.

Ilustración 7. Número de inscripciones en Educación Superior.



Código de la Institución	IES PADRE	Institución de Educación Superior (IES)	Principal o Seccional	ID Sector IES	Sector IES	ID Carácter	Carácter IES	Código del Departamento o IES	Departamento de domicilio de la IES	Código del Municipio (IES)	Municipio de domicilio de la IES	Código SNIES del programa	Programa Académico	ID Nivel Académico	Nivel Académico
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	0	SIN PROGRAMA ESPECIFICO	0	NO APLICA
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	0	SIN PROGRAMA ESPECIFICO	0	NO APLICA
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	0	SIN PROGRAMA ESPECIFICO	0	NO APLICA
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	1	INGENIERIA AGRONÓMICA	1	PREGRADO
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	1	INGENIERIA AGRONÓMICA	1	PREGRADO
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	1	INGENIERIA AGRONÓMICA	1	PREGRADO
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	10	NUTRICIÓN Y DIETÉTICA	1	PREGRADO
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	10	NUTRICIÓN Y DIETÉTICA	1	PREGRADO
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	10	NUTRICIÓN Y DIETÉTICA	1	PREGRADO
1101	1101	UNIVERSIDAD NACIONAL DE COLOMBIA	Principal	1	OFICIAL	4	Universidad	11	BOGOTÁ D.C.	11001	BOGOTÁ D.C.	10	NUTRICIÓN Y DIETÉTICA	1	PREGRADO

Nota. Adaptado de Ministerio de Educación Nacional, numero de inscripciones a programas de educación superior, por MinEducación.

Número de graduados en educación superior de la Fuente de Información Laboral de Colombia.

Ilustración 8. Número de graduados en educación superior.

Periodicidad	Año	Geográfica	Temática	Valor
Anual	2017	Colombia	Superior o universitari	462367
Anual	2016	Colombia	Superior o universitari	423182
Anual	2015	Colombia	Superior o universitari	374738
Anual	2014	Colombia	Superior o universitari	359607
Anual	2013	Colombia	Superior o universitari	351790
Anual	2012	Colombia	Superior o universitari	344904
Anual	2011	Colombia	Superior o universitari	299792
Anual	2010	Colombia	Superior o universitari	227378
Anual	2009	Colombia	Superior o universitari	212445

Nota. Adaptado de Ministerio de Educación Nacional, numero de graduados a nivel nacional, por Estadísticas MinEducacion.

Tasa de ocupación de acuerdo con el nivel de estudios de la Fuente de Información Laboral de Colombia.

Ilustración 9. Tasa de ocupación por nivel de estudios.

Periodicidad	Año	Geográfica	Temática	Valor
Anual	2019	Colombia	Hasta básica primaria	53,2
Anual	2019	Colombia	Basica secundaria	38,3
Anual	2019	Colombia	Media	64,3
Anual	2019	Colombia	Superior o universitari	68,2
Anual	2018	Colombia	Hasta básica primaria	55,1
Anual	2018	Colombia	Basica secundaria	39,7
Anual	2018	Colombia	Media	65,4
Anual	2018	Colombia	Superior o universitari	68,9
Anual	2017	Colombia	Hasta básica primaria	56
Anual	2017	Colombia	Basica secundaria	40,3
Anual	2017	Colombia	Media	65,7

Nota. Adaptado de Ministerio de Educación Nacional, tasa de ocupación nacional, por MinEducacion Estadísticas.

Población ocupada de acuerdo con su último título obtenido del Departamento de Estadística Nacional.

Ilustración 10. Población ocupada en relación a su último título académico

DANE		El futuro es de todos		Gobierno de Colombia							
Gran Encuesta Integrada de Hogares											
Población ocupada según último título obtenido y oficio principal											
Total Nacional											
Año 2010 - 2019											
Concepto	Año										
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	
<b>Total Ocupados</b>	19.213	20.020	20.696	21.048	21.503	22.017	22.156	22.383	22.457	22.287	
Profesional y técnico	1.794	1.873	1.978	2.222	2.267	2.302	2.359	2.436	2.448	2.595	
Directores y funcionarios públicos	465	480	510	509	545	564	568	594	496	479	
Personal Administrativo	1.612	1.665	1.755	1.759	1.859	1.920	1.923	1.903	1.966	1.764	
Comerciantes y Vendedores	3.451	3.559	3.630	3.765	3.746	3.801	3.824	3.770	3.771	3.750	
Trabajadores de servicios	3.418	3.566	3.837	3.976	4.102	4.270	4.294	4.301	4.280	4.330	
Trabajadores agropecuarios y forestales	3.469	3.560	3.554	3.485	3.440	3.472	3.497	3.669	3.660	3.461	
Trabajadores y operarios no agrícolas	4.858	5.164	5.268	5.155	5.370	5.503	5.496	5.526	5.666	5.719	
No informa	146	152	164	178	175	186	195	184	169	190	
<b>Ninguno</b>	10.527	10.744	10.890	10.590	10.532	10.581	10.273	10.157	9.872	9.436	
Profesional y técnico	79	69	90	91	74	73	49	50	52	44	
Directores y funcionarios públicos	73	73	89	67	58	52	56	62	55	44	
Personal Administrativo	268	264	250	234	226	232	193	196	184	149	

Nota. Adaptado de Departamento Nacional de Estadística, población ocupada según título obtenido por Estadísticas DANE.

Oferentes registrados en los portales de empleo de acuerdo con el nivel educativo, del sistema de empleo público nacional.

Ilustración 11. Oferentes registrados en portales de empleo

Vacantes registradas por año y mes según departamento y Clasificación Sectorial CIU 4AC																
2015 - 2020																
Desagregación		Clasificación Sectorial CIU 4AC														
Código DIVIPOL	Departamento	Sección	Nombre Sección	2016												
				No	Dic	Total	En	Fel	Ma	Ab	Ma	Jur	Jul	Ag	Set	Oc
00	Total nacional	SECCIÓN A	Agricultura, ganadería, caza, silvicultura y pesca	1545	1330	7962	1242	1057	841	1384	1039	1183	1007	1213	1133	1294
00	Total nacional	SECCIÓN B	Explotación de minas y canteras	1087	844	8510	952	946	795	1173	1173	1589	1389	1834	1774	2174
00	Total nacional	SECCIÓN C	Industrias manufactureras	12.161	9.052	84.649	13.849	14.523	11.933	14.254	12.697	13.321	11.420	14.366	13.003	11.853
00	Total nacional	SECCIÓN D	Suministro de electricidad, gas, vapor y aire acondicionado	288	169	2.211	246	294	215	325	321	270	300	420	332	379
00	Total nacional	SECCIÓN E	Distribución de agua; evacuación y tratamiento de aguas residuales; gestión	206	179	1.439	242	278	177	240	184	165	162	226	164	187
00	Total nacional	SECCIÓN F	Construcción	4.604	3.139	33.031	5.216	5.368	4.195	5.400	4.671	4.999	4.367	5.573	4.323	4.541
00	Total nacional	SECCIÓN G	Comercio al por mayor y al por menor; reparación de vehículos automotores	15.370	12.280	112.957	16.914	18.223	16.038	19.373	16.966	18.378	16.031	18.718	16.317	15.994
00	Total nacional	SECCIÓN H	Transporte y almacenamiento	4.671	4.258	30.505	4.678	5.057	4.491	5.245	4.313	4.713	3.880	4.566	4.353	4.226
00	Total nacional	SECCIÓN I	Alojamiento y servicios de comida	4.233	3.738	25.052	4.379	4.184	3.858	5.069	4.182	4.325	4.212	5.025	5.205	5.304
00	Total nacional	SECCIÓN J	Información y comunicaciones	8.067	5.980	60.689	7.189	8.606	7.423	9.989	9.681	9.939	9.002	11.844	10.498	11.112
00	Total nacional	SECCIÓN K	Actividades financieras y de seguros	4.038	3.438	40.445	4.355	5.091	4.016	5.420	4.700	5.052	4.725	5.317	4.827	4.303
00	Total nacional	SECCIÓN L	Actividades inmobiliarias	921	566	6.299	1.098	1.213	1.046	1.089	991	843	787	862	798	702
00	Total nacional	SECCIÓN M	Actividades profesionales, científicas y técnicas	16.809	12.813	109.951	14.309	15.445	13.177	17.723	15.315	15.139	13.546	15.265	14.447	14.187
00	Total nacional	SECCIÓN N	Actividades de servicios administrativos y de apoyo	11.604	9.421	72.308	12.002	11.770	10.181	12.434	10.876	12.022	10.570	13.077	12.225	11.394
00	Total nacional	SECCIÓN O	Administración pública y defensa; planes de seguridad social de afiliación	2.827	2.378	22.840	2.743	2.603	2.269	3.080	3.159	2.876	2.291	2.530	2.506	
00	Total nacional	SECCIÓN P	Educación	4.009	3.611	28.045	4.719	3.892	3.219	4.062	3.321	4.173	3.983	4.196	3.579	3.418
00	Total nacional	SECCIÓN Q	Actividades de atención de la salud humana y de asistencia social	5.044	4.909	39.344	5.617	5.656	5.290	6.217	5.815	6.144	4.868	5.960	5.223	4.934

Nota. Adaptado de Observatorio del Servicio Público de Empleo, vacantes registradas por año, por Servicio público de empleo, estadísticas.

Una vez se han seleccionado las bases de datos se continúa estandarizando la información para su análisis, estas acciones se realizan con la eliminación de las columnas con información no relevante, y la preparación de la información de acuerdo con el estándar seleccionado por la aplicación utilizada.

### Modelamiento y evaluación

Para el modelamiento y la evaluación de la información se seleccionó el software RapidMiner, para iniciar se limpia y preparan las bases de datos con el fin de unificar su presentación. En lo referente al archivo de demanda laboral se procede a filtrar las ocupaciones para presentar solamente lo relacionado con la información y las comunicaciones y se presenta la información de forma vertical y en un periodo de mes a mes.

Ilustración 12. Estructura intermedia base de datos demanda laboral.

Desagregación		2015												Total		
Código DIVIPOL	Departamento	Sección	Nombre Sección	En	Fel	Ma	Ab	Ma	Jun	Jul	Ag	Se	Oc		No	Dic
00	Total nacional	SECCIÓN J	Información y comunicaciones	4.461	4.487	4.690	4.285	3.925	3.556	4.073	4.268	4.531	8.467	8.067	5.980	60.689
01	Amazonas	SECCIÓN J	Información y comunicaciones	2	4	0	5	0	1	0	1	0	0	3	0	16
05	Antioquia	SECCIÓN J	Información y comunicaciones	242	272	247	264	299	260	323	299	391	768	835	583	4784
08	Arauca	SECCIÓN J	Información y comunicaciones	3	4	1	2	4	0	2	2	5	3	5	3	33
08	Atlántico	SECCIÓN J	Información y comunicaciones	28	61	54	49	51	51	72	71	64	175	160	108	944
11	Bogotá D.C.	SECCIÓN J	Información y comunicaciones	3.691	3.508	3.679	3.311	3.004	2.789	3.088	3.369	3.381	6.376	6.001	4.577	46.773
13	Bolívar	SECCIÓN J	Información y comunicaciones	27	36	28	38	26	22	17	28	43	28	42	34	371
15	Bolívar	SECCIÓN J	Información y comunicaciones	24	24	11	20	23	13	19	44	41	36	28	12	246

Fuente: Elaboración propia.

En la ilustración No. 11 estructura intermedia de la base de datos para la demanda laboral se presenta la estructura del archivo original que presenta el servicio nacional de empleo.

Continuando, la información se totaliza por año y se presenta en forma vertical por departamento, tal como aparece en la ilustración No. 12 estructura final de la información para la demanda laboral.

Ilustración 13. Estructura final demanda laboral 2015-2020.

DEPARTAMENTO	VACANTES	FECHA
Amazonas	16	20/12/2015
Antioquia	4784	20/12/2015
Arauca	33	20/12/2015
Atlántico	944	20/12/2015
Bogotá D.C.	46.773	20/12/2015
Bolívar	371	20/12/2015
Boyacá	296	20/12/2015
Caldas	373	20/12/2015
Caquetá	24	20/12/2015

Fuente: Elaboración propia.

Con relación a la base de datos de la población mayor de 15 años asistente a cursos

TIC se debe filtrar la información para los cursos correspondientes a ingenierías e informática.

Ilustración 14. Estructura intermedia base de datos asistente cursos TIC.

**TOTAL NACIONAL**

Concepto	abr - jun 13	abr - jun 14	abr - jun 15
Informática	297	276	288
Ingeniería y profesiones afines	123	105	89

Fuente: Elaboración propia.

Una vez en este estado se procede a verticalizar la información y acomodarla al estándar buscado para la aplicación elegida, para esto se suman las dos casillas presentadas y se presentan como un único ítem, con lo cual finalmente la estructura es presentada en la ilustración No. 14 final base de datos asistente a cursos TIC.

Ilustración 15. Estructura final base de datos asistente cursos TIC.

NACIONAL	POBLACION	FECHA
Total Nacional	421152	01/06/2013
Total Nacional	381155	01/06/2014
Total Nacional	377512	01/06/2015
Total Nacional	375256	01/06/2016
Total Nacional	392561	01/06/2017
Total Nacional	399987	01/06/2018
Total Nacional	425632	01/06/2019

Fuente: Elaboración propia.

Para lo concerniente a los trabajadores del campo TIC la base de datos se convierte en dos hojas, donde se evidencian las personas que laboran en los diferentes cargos de la industria según su formación académica, primero se presentan los profesionales y tecnólogos de la industria y posteriormente las personas que no tiene ningún título formal, pero hacen parte de las empresas del subsector.

Ilustración 16. Estructura profesionales y técnicos trabajando en áreas TIC.

Descripción	Total	Fecha
Profesional y tecnólogo TIC	228562	15/12/2010
Profesional y tecnólogo TIC	271056	15/12/2011
Profesional y tecnólogo TIC	328952	15/12/2012
Profesional y tecnólogo TIC	389745	15/12/2013
Profesional y tecnólogo TIC	449862	15/12/2014
Profesional y tecnólogo TIC	474121	15/12/2015
Profesional y tecnólogo TIC	457633	15/12/2016
Profesional y tecnólogo TIC	462501	15/12/2017
Profesional y tecnólogo TIC	424511	15/12/2018
Profesional y tecnólogo TIC	498526	15/12/2019

Fuente: Elaboración propia.

Ilustración 17. Estructura empleados sin título formal trabajando en áreas TIC.

Descripción	Total	Fecha
Sin titulo	79526	15/12/2010
Sin titulo	69852	15/12/2011
Sin titulo	90015	15/12/2012
Sin titulo	91568	15/12/2013
Sin titulo	74598	15/12/2014
Sin titulo	73159	15/12/2015
Sin titulo	49526	15/12/2016
Sin titulo	50120	15/12/2017
Sin titulo	52458	15/12/2018
Sin titulo	41256	15/12/2019

Fuente: Elaboración propia.

En las ilustraciones No. 15 y 16 se presenta la estructura final de los archivos utilizados para el análisis del personal de la industria TIC según su nivel de estudios y su comportamiento a través de los años.

Con el objetivo de presentar la información correspondiente a la base de datos de los egresados en áreas TIC por anualidad, se debe filtrar cada carrera haciendo referencia a todos los campos de conocimiento de las telecomunicaciones e informática, filtrar la información por año seleccionado, al igual que por el departamento de ubicación de la institución de educación superior (IES) y realizar la sumatoria de todos los graduados en las diferentes disciplinas, para finalizar sintetizando el resultado de forma único por cada año anterior.

Ilustración 18. Base de datos inicial Graduados TIC.

Código de la Institución	IES PADRE	Institución de Educación Superior (IES)	Principal o Seccional	ID Sector IES	Sector IES	Caracte	Caracter IES	Código del departamento (IES)	Departamento de domicilio de la IES	Código del Municipio (IES)	Municipio de domicilio de la IES	Código SNIES del programa	Programa Académico
1101	1101	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	26	INGENIERIA DE SISTEMAS	
1101	1101	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	26	INGENIERIA DE SISTEMAS	
1101	1101	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	26	INGENIERIA DE SISTEMAS	
1101	1101	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	26	INGENIERIA DE SISTEMAS	
1101	1101	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	94	MAESTRIA EN INGENIERIA - INGENIERIA DE SISTEMAS Y COMPUTACION	
1101	1101	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	94	MAESTRIA EN INGENIERIA - INGENIERIA DE SISTEMAS Y COMPUTACION	
1101	1101	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	94	MAESTRIA EN INGENIERIA - INGENIERIA DE SISTEMAS Y COMPUTACION	
1103	1101	UNIVERSIDAD SECCIONAL	1	OFICIAL	4	UNIVERSIDAD	17	CALDAS	17001	MANIZALES	4132	ESPECIALIZACION EN ADMINISTRACION DE SISTEMAS INFORMATICOS	
1105	1105	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	147	LICENCIATURA EN DISEÑO TECNOLOGICO CON ENFASIS EN SISTEMAS MECANICOS	
1105	1105	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	170	MAESTRIA EN TECNOLOGIAS DE LA INFORMACION APLICADAS A LA EDUCACION	
1105	1105	UNIVERSIDAD PRINCIPAL	1	OFICIAL	4	UNIVERSIDAD	11	BOGOTA DC	11001	BOGOTA DC	147	LICENCIATURA EN DISEÑO TECNOLOGICO CON ENFASIS EN SISTEMAS MECANICOS	

Fuente: Elaboración propia.

El resultado final del trabajo presenta el siguiente aspecto, ilustración No. 18 “estructura graduados TIC”, un archivo en el cual cabe destacar que se pasaron de 388.771 filas a solo 28.878 registros repartidos en tan solo veintisiete (27) departamentos donde se ubican las diferentes instituciones que prestan el servicio de educación superior, durante el transcurso del horizonte analizado se observa como aparecen y desaparecen departamentos en razón a la creación y eliminación de sus correspondientes instituciones de educación superior, en este caso el valor para el periodo será igual a cero (0).

Ilustración 19. Estructura graduados TIC.

<b>Graduados TIC</b>	<b>Total</b>	<b>Fecha</b>
Antioquia	724	15/12/2001
Atlantico	567	15/12/2001
Bogotá D.C.	5380	15/12/2001
Bolivar	141	15/12/2001
Boyaca	70	15/12/2001
Caldas	110	15/12/2001
Cordoba	97	15/12/2001

Fuente: Elaboración propia.

Teniendo en cuenta que los modelos de información que se presentan tienen una línea de tiempo establecida, se aplicará el modelo de Forecasting pues el objetivo del proyecto es pronosticar el comportamiento de la demanda y oferta de las vacantes en el sector TIC, para determinar las nuevas fortalezas que se deben buscar en el proceso académico para la industria.

Para la elaboración de estos pronósticos se seleccionó el algoritmo de holt winters, esto en razón a que la regresión lineal necesita que el histórico presentado tenga forma de

línea recta, situación que no se presenta en ninguno de los conjuntos de datos analizados y por otro lado el algoritmo seleccionado permite pronosticar diferentes valores numéricos en un periodo de tiempo determinado por medio de la técnica de suavizamiento exponencial con tres variables, teniendo en cuenta tanto el crecimiento como el descenso de los datos, lo que aumenta la precisión del pronóstico y disminuye el margen de error para el proyecto en particular.

Ilustración 20. Formato establecido para RapidMiner

Open in  Turbo Prep  Auto Model

Row No.	DEPARTAME...	VACANTES	FECHA
1	Amazonas	16	Dec 20, 2015
2	Antioquia	4784	Dec 20, 2015
3	Arauca	33	Dec 20, 2015
4	Atlántico	944	Dec 20, 2015
5	Bogotá	46773	Dec 20, 2015

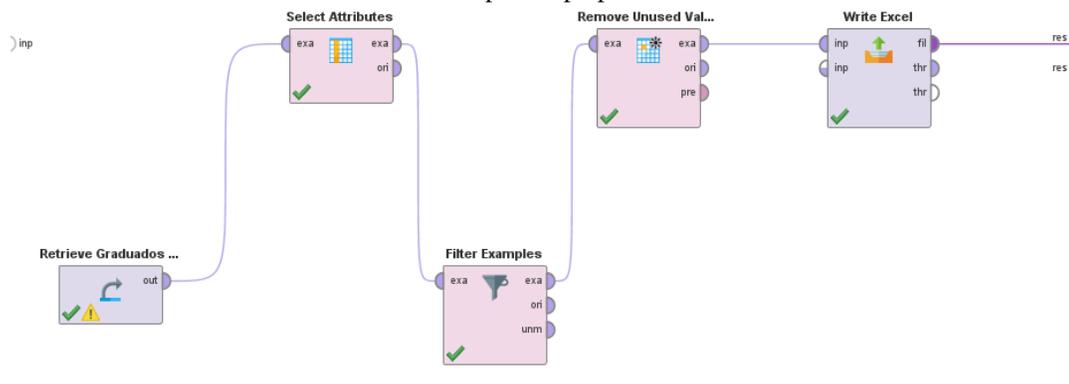
Fuente: Elaboración propia.

Cuando los archivos son correctamente recibidos por la aplicación seleccionada para la minería se presenta su información como aparece en la ilustración No. 19 “Formato establecido para RapidMiner”

Para el preprocesamiento de los datos se utilizan los módulos correspondientes para el alistamiento de la información y la selección de la información importante para la investigación, este proceso de preprocesamiento difiere para cada archivo analizado, pero a

continuación se presenta el proceso utilizado en general, pues sus cambios se dan en la configuración de cada uno de los modelos.

Ilustración 21. Modelo creado para el preprocesamiento de los datos.



Fuente: Elaboración propia.

Como se observa en la ilustración No. 20 “modelo creado para el preprocesamiento de los datos” el modelo inicia con el módulo de Retrieve, este módulo se crea cuando el archivo que contiene la información hace parte de los repositorios locales de RapidMiner, cuando esto se realiza se debe imponer diferentes atributos a los datos del archivo, para el caso las variables numéricas tienen el atributo prediction mientras las fechas tendrán el atributo id con el objetivo de informar a los modelos que la información que se busca predecir son los valores numéricos y el identificador son los valores de tipo date, una vez configurados este archivo se arrastra al panel de proceso y se crea automáticamente el módulo.

Con el módulo Select Attributes, como su nombre lo indica se seleccionan los atributos importantes del archivo, es decir, se está en capacidad de eliminar las columnas innecesarias para la investigación sin importar el tipo de atributo que sea la columna.

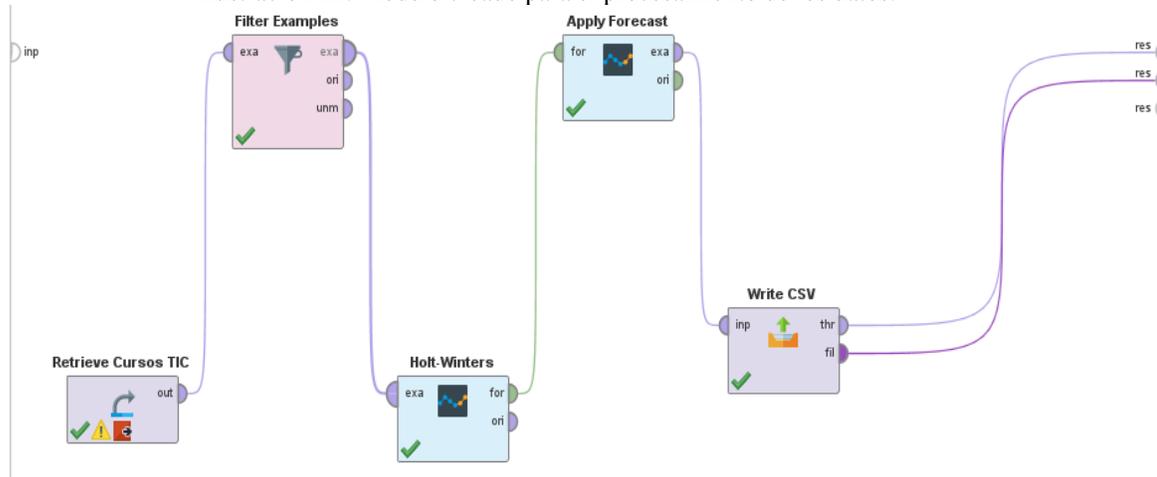
Filter Examples, permite filtrar las celdas del archivo por medio de diferentes filtros, este módulo permite seleccionar los valores adecuados para cada atributo, se utiliza para establecer valores que sean acordes a los rangos de cada característica.

Remove Unused Values, busca los valores que tienen errores en todo el archivo y de acuerdo con la configuración los ubica como un numero promedio, un valor predeterminado o una formula específica, se utiliza para tener un archivo consistente y completo.

Por último se crea un archivo de Excel con las especificaciones finales para ser procesado en búsqueda de un pronóstico certero, cabe resaltar que cada tipo de archivo necesita un número de módulos distintos, pero el modelo es eficiente con todos ellos, cuando el archivo tiene una característica en específico completa el módulo simplemente se deja con los valores específicos para no afectar su integridad y se continua con el siguiente, este modelo busca crear archivos que tengan la información de interés para la investigación y produzcan los mejores resultados de predicción.

Una vez cada base de información esta procesada en la aplicación seleccionada es posible implementar el modelo presentado a continuación para obtener un pronóstico para cada una de las variables de interés.

Ilustración 22. Modelo creado para el procesamiento de los datos.



Fuente: Elaboración propia.

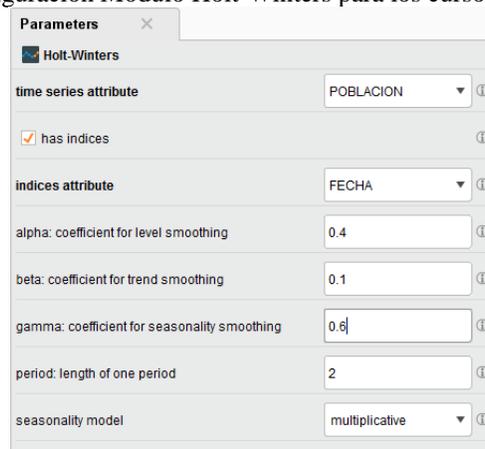
En lo relacionado con el procesamiento de los datos la ilustración No. 21 “modelo para el procesamiento de los datos” exterioriza como se inicia con la inclusión de los archivos en los repositorios locales de RapidMiner, esto con el objetivo de arrástralos al área de trabajo y formar el módulo de Retrieve este módulo provee el repositorio entrante para el archivo que será utilizado, estos archivos son los creados en la etapa de procesamiento presentada anteriormente.

A continuación, se presenta el módulo de Filter Examples con el objetivo de filtrar las celdas de acuerdo con parámetros que busquen aumentar la exactitud del modelo, en este módulo de acuerdo con las necesidades del archivo se pueden eliminar celdas que estén fuera de un rango razonable para cualquier atributo y buscar el mejor pronóstico de la aplicación.

El módulo de la técnica de suavizamiento exponencial triple Holt-Winters tiene la ventaja de tener en cuenta los componentes de nivel, tendencia y estacionalidad en sus

pronósticos lo que aumenta la efectividad de sus resultados para los tipos de datos que se manejan en la investigación, para entender su funcionamiento la variable Alpha especifica el coeficiente para el suavizado de nivel, la variable beta contiene el coeficiente de suavizado para las tendencias y Gamma maneja el coeficiente del suavizado para la estacionalidad, este operador se selecciona pues si bien existen operadores similares este módulo en particular está diseñado específicamente para trabajar con series de tiempo, su mayor ventaja es que el modelo de pronóstico se aplica a los mismos datos que se procesaron, los coeficientes utilizados para la configuración del archivo de cursos de capacitación en educación TIC son los siguientes:

Ilustración 23. Configuración Módulo Holt-Winters para los cursos de capacitación TIC.



The image shows a software window titled "Parameters" for the "Holt-Winters" model. It contains several configuration fields:

- time series attribute:** A dropdown menu set to "POBLACION".
- has indices:** A checked checkbox.
- indices attribute:** A dropdown menu set to "FECHA".
- alpha: coefficient for level smoothing:** A text input field containing "0.4".
- beta: coefficient for trend smoothing:** A text input field containing "0.1".
- gamma: coefficient for seasonality smoothing:** A text input field containing "0.6".
- period: length of one period:** A text input field containing "2".
- seasonality model:** A dropdown menu set to "multiplicative".

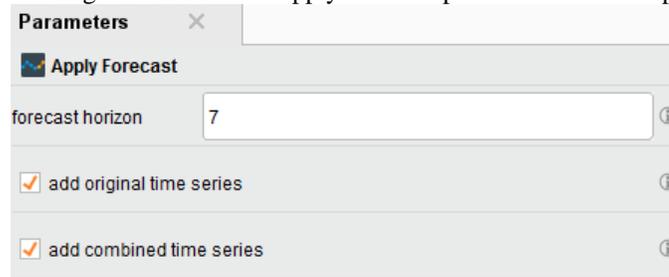
Fuente: Elaboración propia.

La selección del método Holt-winters tiene su fundamento en el comportamiento de las variables de interés en los periodos anteriores, donde las gráficas demostraban que se presentaba una tendencia lineal positiva pero con una estacionalidad prolongada en el tiempo, teniendo en cuenta lo anterior se seleccionó este método pero además se seleccionó el modelo multiplicativo porque estas estacionalidades cambian en función a las diferentes series de tiempo, con relación a la selección de los coeficientes seleccionados es natural que

se busque la minimización la medida global de error para el pronóstico, es por esto que para el Alpha obtiene un valor de 0.4 en relación a que la tendencia positiva en los comportamientos se mantiene, beta es igual 0.1 pues la estacionalidad se presenta en periodos de tiempo que son conocidos y su generación es un proceso casi automático, y gamma con 0.6 pues es necesario que el porcentaje de suavizamiento sea superior al 50% por la necesidad de realiza la aproximación más certera posible, estos indicadores se conocen por método del método de mínimos cuadrados. Por medio de la minimización de la suma de los errores cuadrados.

El Operador Apply Forecast se utiliza para presentar un horizonte de pronóstico específico, en la ilustración No. 22 “configuración módulo Holt-Winters” se presentan los diferentes atributos que se deben configurar para obtener los resultados deseados, su función es usar un modelo ya definido con el Holt-Winters para predecir los valores siguientes de una serie de tiempo específica, para esta investigación los horizontes serán diferentes con el objetivo de establecer pronósticos hasta el periodo 2026, esto con el objetivo de establecer conclusiones con un mismo horizonte pronosticado, en la ilustración No. 23 “configuración modulo apply forecast” se presenta la configuración para los datos de los cursos de capacitación TIC.

Ilustración 24. Configuración Modulo Apply Forecast para los cursos de capacitación TIC.



Fuente: Elaboración propia.

Por último, con el operador Write CSV se exportan los resultados a una ubicación específica en un archivo de extensión csv para facilitar su análisis.

Con este modelo propuesto, ya es posible continuar con la evaluación de la información, como insumo se obtienen los pronósticos para las demandas y ofertas en el sector TIC, así como las personas mayores de 18 años que asisten a cursos cortos sobre la industria y el número de profesionales, tecnólogos y técnicos al igual que las personas sin formación académica superior que se encuentran laborando en el sector TIC hasta el año 2026, para determinar las necesidades de la academia e industria de este subsector a futuro y las competencias necesarias para adecuada satisfacción de los nuevos puestos de trabajo.

Se resalta que para el caso de la demanda y la oferta del sector laboral TIC se obtienen resultados de acuerdo con cada departamento nacional, pero para el caso de los cursos cortos y el número de personas que laboran en el sector discriminados por su nivel de formación los resultados son de forma anual, y todas las proyecciones obtenidas siguen la misma estructura.

Ilustración 25. Estructura de la proyección Asistentes cursos TIC.

Row No.	FECHA	forecast of ...	POBLACION	POBLACION ...
1	Jun 1, 2013 1...	?	421152	421152
2	Jun 1, 2014 1...	?	381155	381155
3	Jun 1, 2015 1...	?	377512	377512
4	Jun 1, 2016 1...	?	375256	375256
5	Jun 1, 2017 1...	?	392561	392561
6	Jun 1, 2018 1...	?	399987	399987
7	Jun 1, 2019 1...	?	425632	425632
8	May 31, 2020 ...	348627.408	?	348627.408
9	May 31, 2021 ...	370265.015	?	370265.015
10	May 31, 2022 ...	308564.860	?	308564.860
11	May 31, 2023 ...	325122.184	?	325122.184
12	May 30, 2024 ...	268502.313	?	268502.313
13	May 30, 2025 ...	279979.353	?	279979.353
14	May 30, 2026 ...	228439.765	?	228439.765

Fuente: Elaboración propia.

Ilustración 26. Estructura de la proyección de profesionales, tecnólogos y técnicos trabajando en el sector TIC.

Row No.	Fecha	forecast of ...	Total	Total and for...
1	Dec 15, 2010...	?	228562	228562
2	Dec 15, 2011...	?	271056	271056
3	Dec 15, 2012...	?	328952	328952
4	Dec 15, 2013...	?	389745	389745
5	Dec 15, 2014...	?	449862	449862
6	Dec 15, 2015...	?	474121	474121
7	Dec 15, 2016...	?	457633	457633
8	Dec 15, 2017...	?	462501	462501
9	Dec 15, 2018...	?	424511	424511
10	Dec 15, 2019...	?	498526	498526
11	Dec 14, 2020...	525630.098	?	525630.098
12	Dec 14, 2021...	584954.232	?	584954.232
13	Dec 14, 2022...	603678.871	?	603678.871
14	Dec 14, 2023...	665808.917	?	665808.917
15	Dec 13, 2024...	681727.645	?	681727.645
16	Dec 13, 2025...	746663.602	?	746663.602
17	Dec 13, 2026...	759776.419	?	759776.419

Fuente: Elaboración propia.

Ilustración 27. Estructura de la proyección de personal sin educación superior trabajando en el sector TIC.

Row No.	Fecha	forecast of ...	Total	Total and for...
1	Dec 15, 2010...	?	79526	79526
2	Dec 15, 2011...	?	69852	69852
3	Dec 15, 2012...	?	90015	90015
4	Dec 15, 2013...	?	91568	91568
5	Dec 15, 2014...	?	74598	74598
6	Dec 15, 2015...	?	73159	73159
7	Dec 15, 2016...	?	49526	49526
8	Dec 15, 2017...	?	50120	50120
9	Dec 15, 2018...	?	52458	52458
10	Dec 15, 2019...	?	41256	41256
11	Dec 14, 2020...	51227.133	?	51227.133
12	Dec 14, 2021...	47747.830	?	47747.830
13	Dec 14, 2022...	54215.694	?	54215.694
14	Dec 14, 2023...	50454.458	?	50454.458
15	Dec 13, 2024...	57204.254	?	57204.254
16	Dec 13, 2025...	53161.087	?	53161.087
17	Dec 13, 2026...	60192.815	?	60192.815

Fuente: Elaboración propia.

Una vez se corre la aplicación se obtienen los resultados del pronóstico como se observa en las ilustraciones 25 a la 27 correspondientes a las estructuras de las proyecciones para vacantes, graduados y personal TIC hasta 2026.

## **Despliegue del proyecto**

Para el despliegue del proyecto es necesario el uso de una aplicación que permita presentar los resultados de forma interactiva, fácil de operar por parte de cualquier tipo de usuario y que este en capacidad de presentar modelos fiables en relación a una variable específica, para lo anterior se selecciona la herramienta Power BI en razón a su facilidad para conectar y presentar los datos mediante una plataforma escalable y unificada lo que beneficia a los usuarios en su observación y obtención de conocimientos por medio de los datos, además la interactividad en la presentación de información por medio de reportes, informes, paneles y conjuntos de datos de forma ordenada y fácil de entender, lo que facilita no solo la presentación de los resultados sino la facilidad de obtención de la información necesaria para nuevos investigadores que pretendan iniciar su generación de conocimiento desde los resultados de la presente investigación.

## **10. ANÁLISIS DE LOS RESULTADOS**

Teniendo como fundamento los resultados obtenidos de acuerdo a los pronósticos arrojados por el software utilizado en lo referente al mercado laboral TIC, el comportamiento de los empleados de la industria según sus estudios la investigación permite determinar el panorama del país en estos ámbitos y ofrecer posiciones que permitan identificar las necesidades actuales y futuras para el personal que se encuentra en preparación con el objetivo de enfrentar los cambios y retos de un subsector económico que tiene un amplio futuro en todos los campos laborales del país.

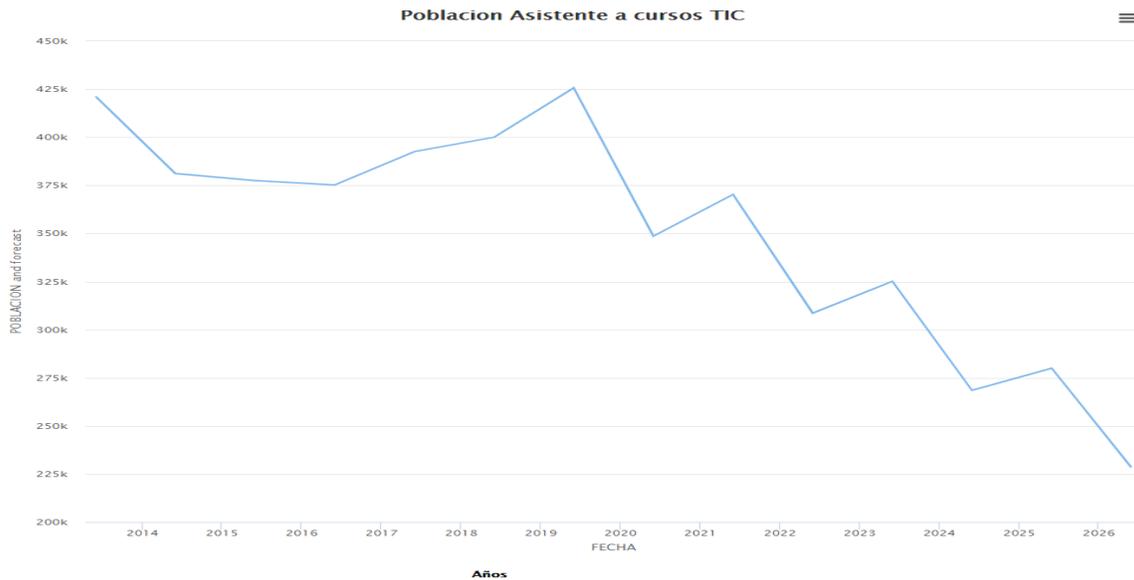
Como primera medida se establecen los vacíos que provee los graduados que tiene anualmente el país contra las nuevas vacantes que se generan en una industria de amplio crecimiento, pero también se evidencia la falta de colaboración por parte de las industrias con el sector educativo en la búsqueda de creación de programas académicos que se encuentren en capacidad de brindar profesionales competentes para estas nuevas labores.

También se evidencia el progreso periódico de la contratación de diferente tipo de personal donde se dividen las personas que trabajan en la industria TIC de acuerdo con sus conocimientos adquiridos, entre profesionales, tecnólogos y técnicos, así como el personal que cuenta con diferentes cursos enfocados en herramientas específicas de programación y otros modelos de educación individual.

### **10.1. Proyecciones nacionales**

Con el objetivo de presentar de manera inicial un panorama general de los resultados, a continuación, se presentan las proyecciones a nivel nacional de las diferentes variables analizadas, estudiadas en lapsos de tiempo anual con el objetivo de entender su comportamiento a través los años y establecer su futuro más próximo.

Ilustración 28. Comportamiento personas mayores de 18 años que asisten a cursos de educación TIC.

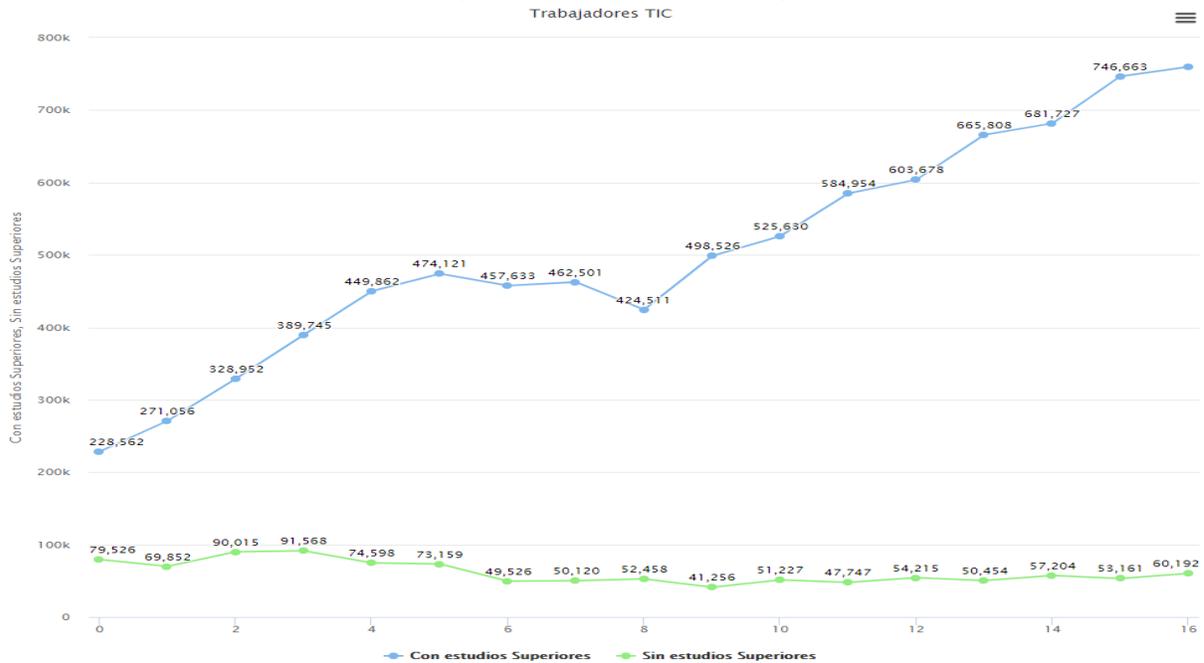


Fuente: Elaboración propia.

En la ilustración No. 28 “comportamiento de las personas mayores de 18 años asistentes a cursos de educación TIC” se presenta claramente como si bien en periodos pasados los cursos en temas TIC tenían un buen auge, este impacto cambio radicalmente y tiene como tendencia su comportamiento a la baja desde 2019, y es que para los años 2014 a 2019 las campañas publicitarias de las empresas encargadas de este tipo de cursos tenían amplia circulación a nivel departamental y nacional, lo que impulsaba a la juventud nacional a buscar su aplicación en razón a la rapidez de finalización y la posibilidad de mejorar su ingreso al empleo formal de una forma rápida y con mejores condiciones salariales, teniendo en cuenta que la industria de la tecnología y sus diferentes variantes ofrecen una amplia gama de necesidades de recurso humano, pero una vez los interesados son contratados se debe observar diferentes falencias específicas en términos de razonamiento crítico y operación sobre algoritmos, pues sus conocimientos se fundamentan en la operación de diferentes lenguajes de programación y sus entornos de trabajo, así como

de modelos específicos de conocimiento que no tienen un sustento adecuado para la aplicación en la vida laboral actual, en este orden de ideas las empresas pueden evaluar si la contratación de este tipo de personal no solo genera un desgaste administrativo sino económico, pues una vez se ha contratado el personal también se está en la necesidad de capacitarlo, generando costos adicionales pues su capacitación se ofrece desde las bases mismas de los diferentes paradigmas y teorías correspondientes.

Ilustración 29. Comportamiento nivel de estudios personal TIC.



Fuente: Elaboración propia.

De acuerdo a la ilustración No. 29 “comportamiento nivel de estudios personal TIC” se observa como dentro de todos los periodos el personal correspondiente al sector TIC cada día está más enfocado a las personas con diferentes estudios superiores completos (técnicos, tecnólogos, profesionales y superiores), lo que indica que la mano de obra es muy calificada en la industria, pero además se observa claramente que la tendencia de los profesionales aumenta año tras año mientras el personal sin este tipo de estudio tiende a la

baja, esto en razón a la especialización que el subsector muestra cada día, lo que indica que el personal humano de la industria es altamente capacitado y se debe amoldar constantemente a las necesidades del mercado, esta formación se logra con una base fuerte en temas de algoritmia, historia de la programación y fuentes en temas como base de datos y programación en la nube, la cual no es posible abarcar en cursos cortos y por lo tanto se crean grandes vacíos en las personas que optan por este tipo de capacitación corta.

Si a todo lo anterior, se le suma la innovación y la constante creación de nuevas formas de trabajo, así como de paradigmas de diseño y producción que traen consigo las nuevas tecnologías de la industria, es evidente que las personas que trabajan en la industria necesitan obligatoriamente excelentes bases para entender los cambios que se producen constantemente en la industria.

Frente a este panorama la única oportunidad que deben tener las personas interesadas en ingresar a la industria TIC es formarse profesionalmente en las instituciones de educación superior debidamente registradas ante el Ministerio de Educación Nacional, en este orden de ideas y en aras de combatir el factor económico que se resalta como el de mayor impacto para la no presentación de todos los interesados en este tipo de instituciones, sería recomendable implementar campañas que permitan mediante la unión entre industria y academia, el ingreso de toda la población interesada a este tipo de formación, por medio de una inclusión inmediata a las empresas del subsector y un descuento obligatorio de nómina que les permita garantizar sus estudios hasta la culminación, con lo cual no solo se acrecentaría el número de aspirantes a los cargos vacantes, sino que se aumentaría la calidad de educación impartida en las aulas nacionales.

Lo anterior, se puede confirmar cuando se observa el promedio de cada una de estas variables, mientras el promedio de profesionales trabajando en el sector se encuentra en 503.000 el personal sin este tipo de formación solo alcanza los 61.500(pero se debe tener en cuenta que dentro de este rango están otro tipo de profesiones que realizan funciones específicas de su conocimiento) por lo que este valor seria mucho menor, pero además el porcentaje promedio de aumento de profesionales del subsector se encuentra en un 8.8% mientras el personal con otro tipo de formación disminuye en razón de 4.4% con cada año que pasa.

Para sintetizar gráficamente esta relación entre personal con estudios superiores dentro del subsector se presenta la siguiente ilustración donde se expone la situación para el periodo 2019.

Ilustración 300. Distribución empleados TIC 2019.

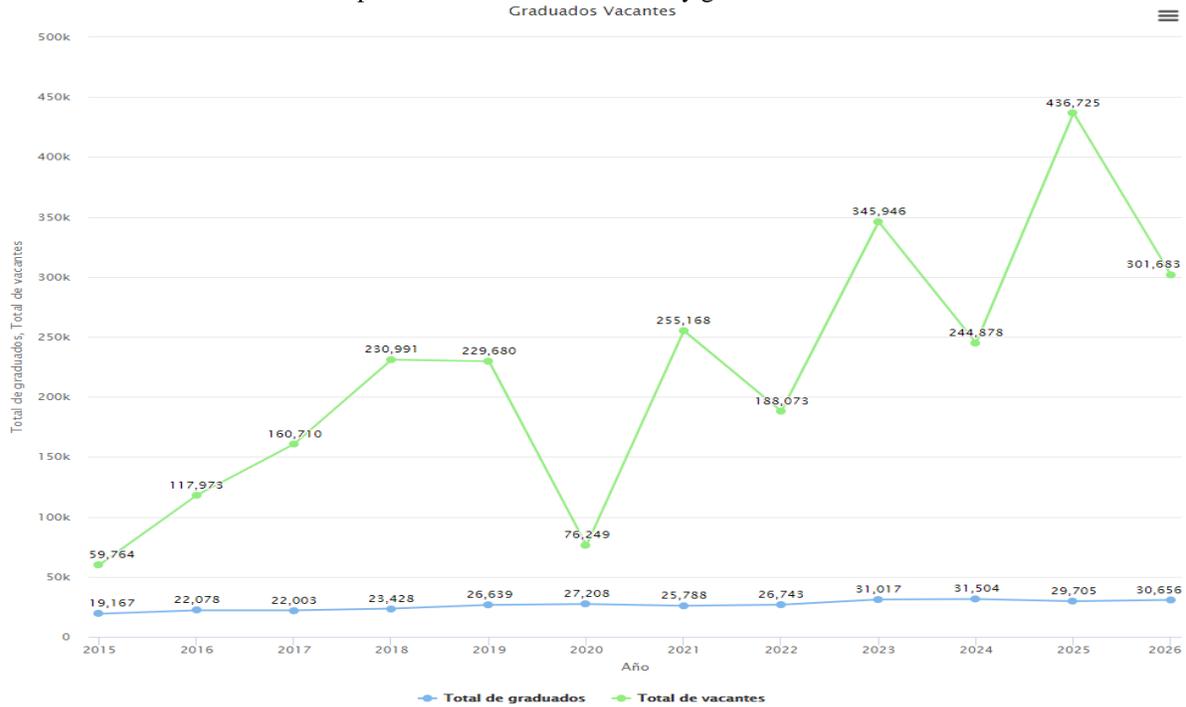


Fuente: Elaboración propia.

De la anterior ilustración No. 30 “distribución empleados TIC 2019” cabe resaltar como el 92% de la participación del personal TIC tiene estudios superiores desde términos de iniciados hasta finalizados, en sus diferentes niveles y solo un 8% de la población que labora en la industria no presenta ningún tipo estudio superior, lo que obligatoriamente los

enmarca en el sector de los empleados que ingresaron al subsector con el conocimiento de alguna herramienta en particular a través de educación de cursos cortos y específicos.

Ilustración 311. Comportamiento nivel de vacantes y graduados TIC a nivel nacional.



Fuente: Elaboración propia

Con esta ilustración No. 31 “comportamiento nivel de vacantes y graduados TIC a nivel nacional” queda claro que aunque el país tiene un incremento muy leve año tras año en el número de graduados en áreas TIC, estos valores no son suficientes para atender la demanda de vacantes para la industria, este es uno de los factores que puede afectar la calidad de los productos de la industria, puesto que las empresas se ven obligadas a recibir personal que no tiene un conocimiento previo coherente para los puestos que van a desempeñar y por lo tanto el tiempo de prueba o acomodación al cargo es muy amplio, siendo este el mejor escenario, porque también se puede dar que el porcentaje de rotación del personal en las empresas TIC sea muy alto, pues las personas contratadas solo buscan

satisfacer su necesidad básica de remuneración y una vez encuentran los puestos de trabajo adecuados amplían este porcentaje de rotación.

Ahora bien, es importante establecer que la política educativa del país tiene enfoques muy diferentes al de las necesidades actuales de puestos de trabajo, por lo que se enfocan a los estudiantes desde edades cortas en otros tipos de conocimiento que aunque necesarios en la actualidad tienen dificultades para el desarrollo de su ejercicio laboral, en este sentido es importante inculcar desde muy pequeños a los estudiantes colombianos prácticas enfocadas al uso y el entendimiento de las bases tecnológicas con el objetivo de ejercer presión para que sean alcanzados un mínimo de créditos o logros informáticos y tecnológicos, pues si bien todas las áreas del saber son necesarias y de obligatorio conocimiento, es indudable que cada una de ellas se ve afectada positivamente por la aparición de la tecnología en sus actividades diarias, lo anterior para concluir que en la actualidad para impulsar no solo el sector TIC, sino la economía nacional en general es importante ampliar la cátedra en áreas como la informática, tecnología y bases de programación pues son habilidades que en la actualidad pueden ser caracterizadas como generales.

Todo lo anterior, con base en la ilustración que presenta la diferencia del aumento entre los graduados de programas TIC con las vacantes que ofrece el clúster de esta industria, pues como se observa mientras el promedio de las vacantes oscila entre las 220.000, las instituciones de educación superior están graduando un promedio de 26.000 estudiante por año lo que genera no solo una brecha gigante de personal TIC, sino una

reducción creciente de la tecnología que es aplicada en las diferentes empresas del país, generando un rezago tecnológico que es cada vez mayor con los países de la región.

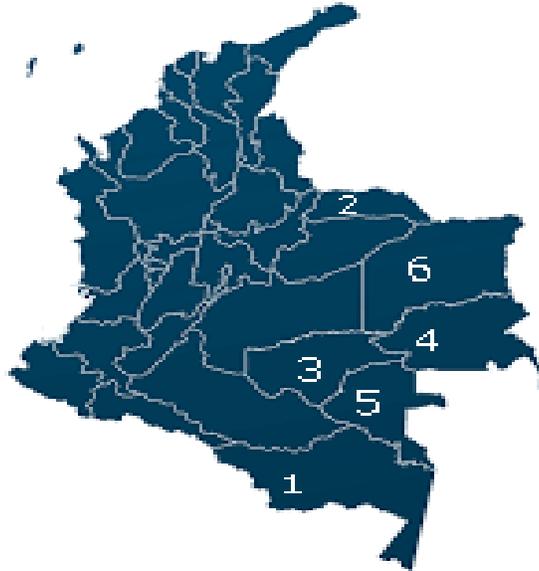
Ahora bien, ante este panorama se hace evidente que se establezcan políticas que permitan flexibilizar el ingreso a la educación superior, porque de no ser así y aunque no sea la solución es una de las razones por las cuales las empresas del subsector deben acudir a personal especializado en otras áreas para atender sus necesidades, lo que va en contravía de la calidad de sus productos y servicios, pero además dificulta la competencia de las empresas nacionales en el ámbito internacional.

Para concretar este análisis se hace inevitable observar cómo desde 2019 a 2025 los graduados aumentan tan solo en 3.400 mientras las vacantes lo hacen en 225.000 un fenómeno que debe ser atendido por todas las entidades públicas y privadas que deben velar por la modernización de todos los procesos de la sociedad pues es evidente que la falta de profesionales y personal TIC es una grave desventaja del subsector, no en vano esta situación ya es advertida por los diferentes medios de comunicación nacionales como es el caso del diario La República quien en su artículo de 11 de abril de 2019 por medio de entrevista al gerente de IBM Colombia ya mencionaba que el déficit de profesionales rondaba los 45.000 (la entrevista se puede observar en el siguiente enlace: <https://www.larepublica.co/empresas/sector-de-tecnologias-tiene-deficit-de-45000-empleados-en-colombia-2850190>) y según esta investigación este valor aumentaría para los próximos años.

## **10.2. Proyecciones departamentales**

Para continuar se presentan las diferencias entre el número de graduados a nivel nacional en el sector TIC contra las vacantes a nivel departamental, para iniciar el análisis es fundamental expresar que existen seis (06) departamentos que en la actualidad no cuentan con instituciones de educación superior lo que para el año 2021 es impensado para un país que tiene la economía naranja como su pilar fundamental, estos departamentos son Amazonas, Arauca, Guainía, Guaviare, Vaupés y Vichada como se observa en la ilustración No. 32 “departamentos sin instituciones de educación superior”, para estos casos se presentara el pronóstico hasta el 2026 de sus vacantes en este subsector, pues si bien no cuentan con programas presenciales TIC si ofrecen un número constante de vacantes en esta área de conocimiento.

Ilustración 322. Departamentos sin instituciones de educación superior.

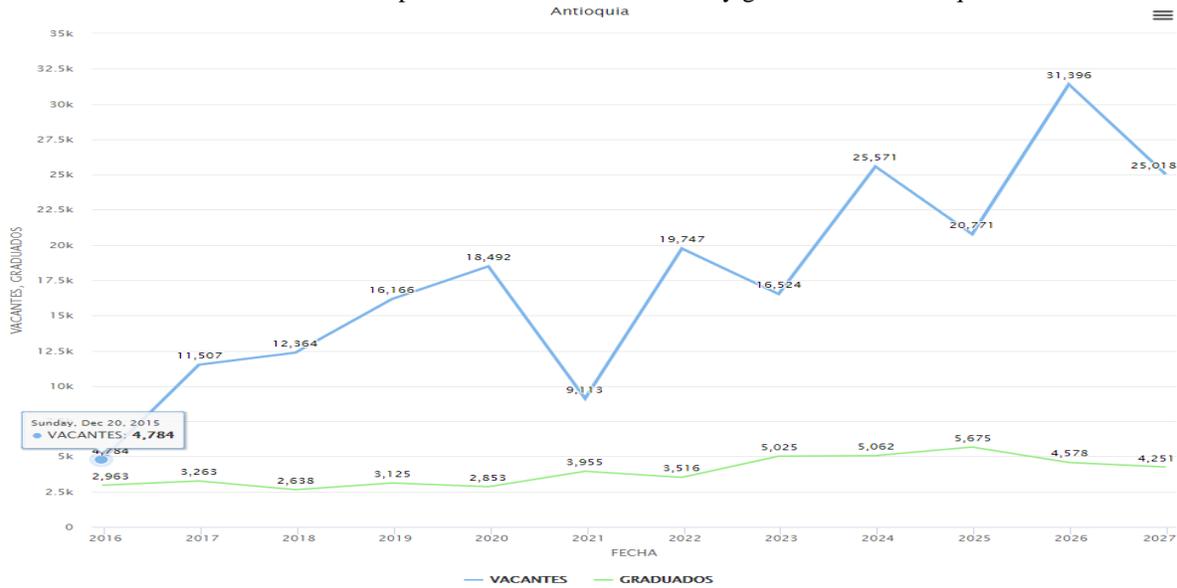


Fuente: Elaboración propia.

A continuación, se presentan los resultados de los departamentos más importantes y las regiones que presentan condiciones fuera de la generalidad del país, en la que se demuestra que con el paso del tiempo la brecha entre graduados y vacantes se amplía año

tras año, es decir, que con el paso del tiempo existe un mayor número de vacantes que graduados en el sector TIC.

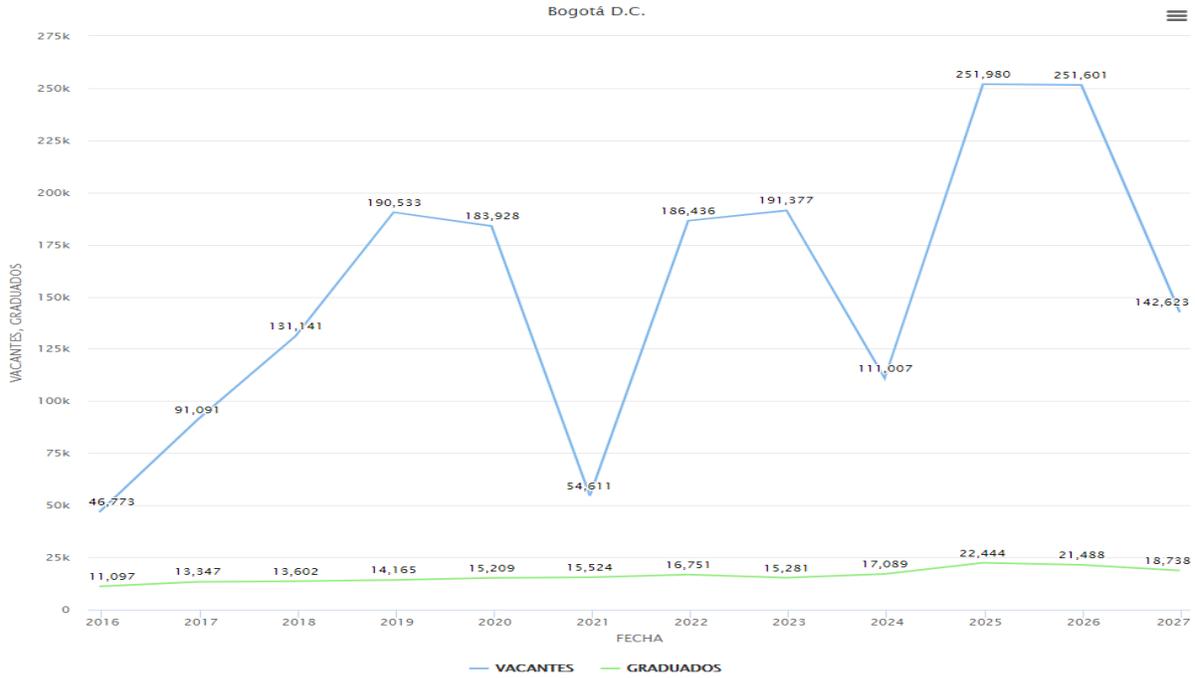
Ilustración 333. Comportamiento nivel de vacantes y graduados en Antioquia.



Fuente: Elaboración propia.

La ilustración No. 33 “comportamiento de vacantes y graduados en Antioquia” demuestra como en el departamento de Antioquia el incremento de las vacantes es muy alto para el comportamiento de los graduados de las universidades del departamento, si se tienen en cuenta los datos promedio, mientras las vacantes son 17.000 puestos los graduados solo llegan a 3.900 esta situación provee anualmente un promedio de 13.000 vacantes que no son suplidas por profesionales del departamento, cabe destacar que el departamento de Antioquia en la actualidad provee amplios puestos de trabajo en este sector por sus altos niveles de innovación y empresas con actividades relacionadas con el sector TIC.

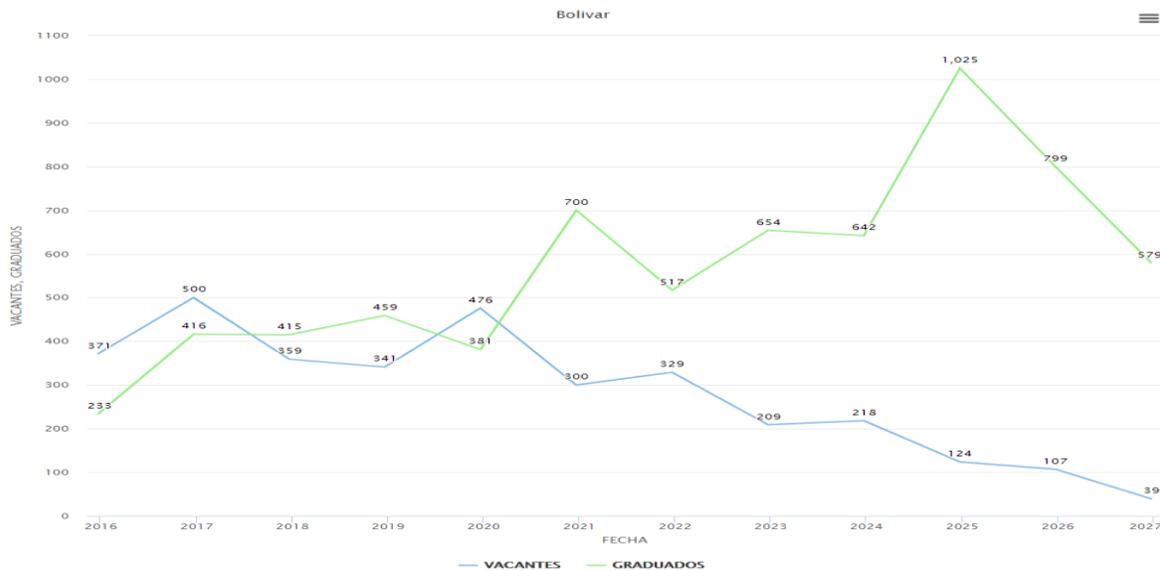
Ilustración 344. Comportamiento nivel de vacantes y graduados en Bogotá D.C.



Fuente: Elaboración propia.

Según la ilustración No. 34 “comportamiento de vacantes y graduados en Bogotá D.C.” es claro que en la ciudad capital siempre las vacantes han sido mínimas en referencia al número de vacantes que se presentan en la ciudad, los graduados que son oficialmente titulados por las IES no alcanzan para suplir estas necesidades, pero este factor se debe observar desde la calidad de las vacantes porque si bien es cierto se analiza el número de graduados muchas de estas vacantes son suplidas por los estudiantes universitarios, pero igualmente este fenómeno contribuye a la rotación de personal en las diferentes vacantes de la industria lo que reduce en muy poco la brecha entre vacantes y personal capacitado para la labor.

Ilustración 35. Comportamiento nivel de vacantes en Bolívar.



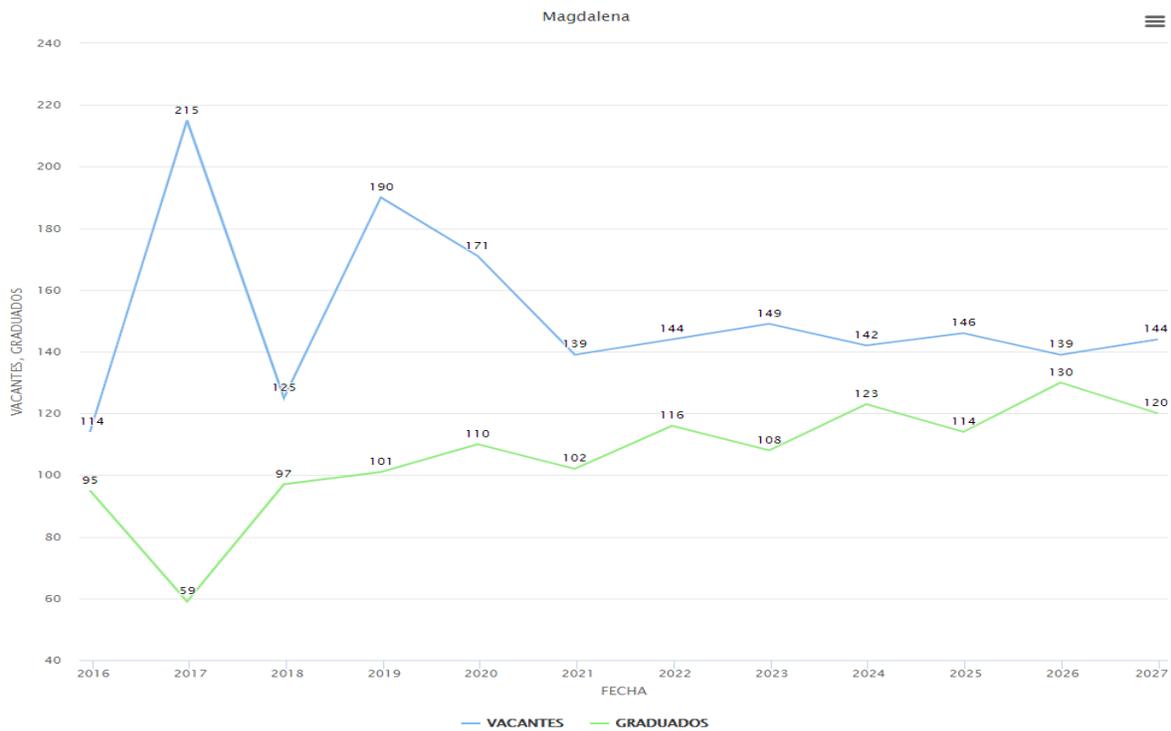
Fuente: Elaboración propia.

Como se presenta en la ilustración No. 35 “comportamiento de vacantes y graduados en Bolívar” Este departamento tiene una particularidad en el comportamiento de las dos variables analizadas, tal vez en los periodos pasados era el departamento con mejores resultados con relación a las vacantes que creaba y el número de graduados producidos, pero según el pronóstico obtenido bajo la herramienta utilizada este número de graduados será reducido a futuro mientras que las vacantes para el sector TIC en la región seguirán en aumento, un resultado que crea una oportunidad para las instituciones de educación superior que deben atender la reducción de este número de graduados y estudiantes de educación superior.

Una respuesta a este fenómeno debería ser el aumento de estudiantes en instituciones de educación de la misma región dentro de otro departamento lo que sería un llamado de atención para las universidades del departamento pues esta reducción puede

estar relacionada con la calidad de la educación impartida, no por nada mientras el promedio de graduados dentro del periodo analizado es de 238 estudiantes, el promedio de las vacantes oscila en 568 puestos de trabajo, mostrando la alta diferencia en la tendencia entre las dos curvas presentadas.

Ilustración 366. Comportamiento nivel de vacantes y graduados en Magdalena.

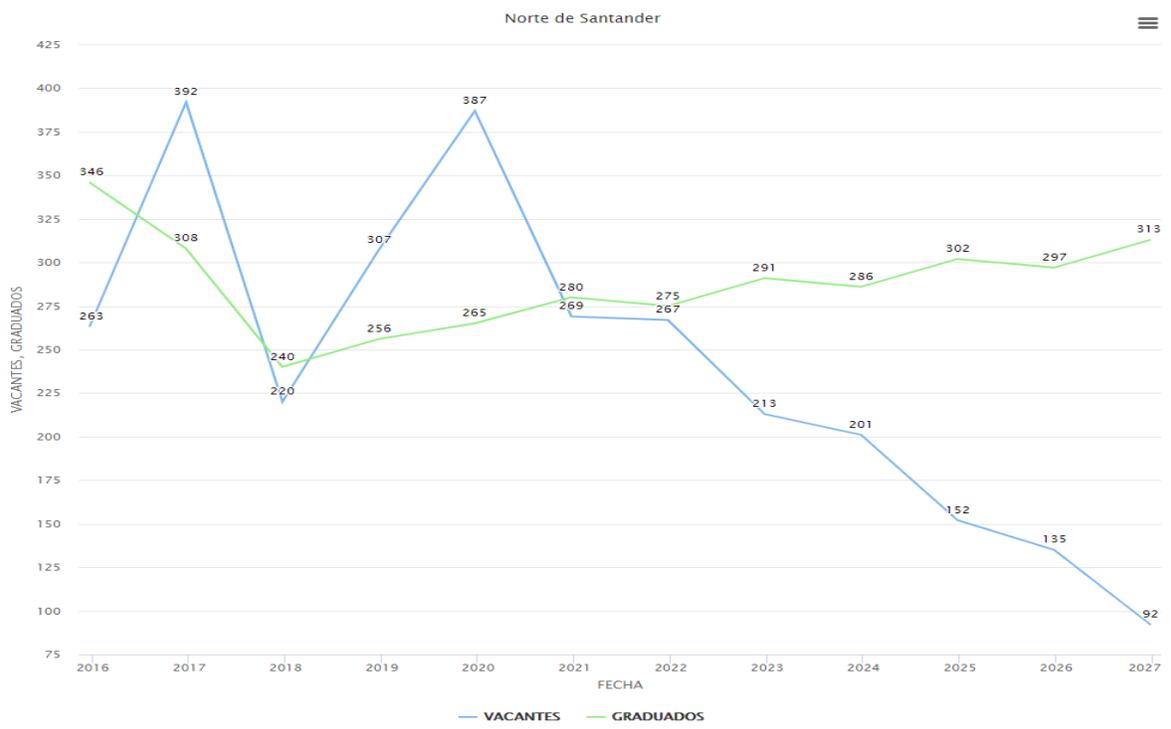


Fuente: Elaboración propia.

Para el departamento de Magdalena en la ilustración No. 36 “comportamiento nivel de vacantes y graduados en Magdalena” es evidente que las curvas son completamente inversas, lo que produce que en diferentes años la diferencia entre las dos variables se aumente y se acerque respectivamente, aunque igual que la mayoría de los departamentos la diferencia entre vacantes y graduados es abrumadora, en Magdalena el

promedio de graduados es de 130 mientras que de vacantes alcanza los 215, razón por la cual entre más pasa el tiempo esta brecha se cierra para el departamento.

Ilustración 37. Comportamiento nivel de vacantes y graduados en Norte de Santander.

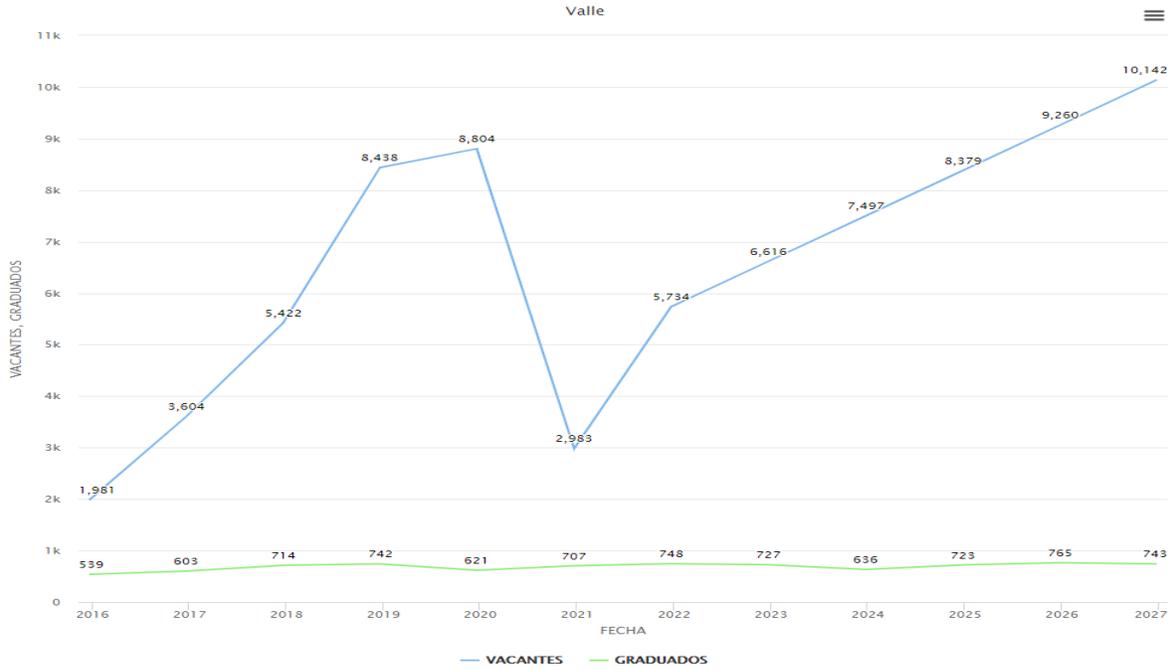


Fuente: Elaboración propia.

Según la ilustración No. 37 “comportamiento nivel de vacantes y graduados en Norte de Santander”, para este departamento se observan curvas muy distantes a los demás, en 2016 el número de vacantes es inferior a los graduados del departamento, pero al pasar el tiempo estos puestos de trabajo serán reducidos, mientras los graduados aumentaran pues su curva tiene una tendencia levemente positiva, esto quiere decir, que el departamento no cuenta con departamentos de tecnología e innovación en sus organizaciones, pero además

que los estudiantes de estas instituciones prestaran sus servicios a empresas de otros departamentos.

Ilustración 37. Comportamiento nivel de vacantes y graduados en Valle.



Fuente: Elaboración propia.

El análisis más importante en relación con la ilustración No. 38 “comportamiento nivel de vacantes y graduados en Valle” para el departamento es su ubicación como la región que mayor índice de puestos de trabajo crea en el país, su crecimiento es evidente únicamente opacado por el fenómeno de salud pública iniciado en 2020, pero se ubica como la región de mayor creación de empleos en el sector TIC, estos números contrastan con los graduados de la región pues si bien los números crecen no lo hacen de una manera similar a las vacantes lo que conlleva a pensar que la industria TIC del departamento es muy estable y experimentada, pero necesita profesionales de otras regiones para su correcto funcionamiento.

A nivel general del análisis realizado se aprecia como el país, que en la actualidad le apuesta al sector TIC como pilar para el desarrollo de la economía tienen grandes retrasos en términos de preparación de jóvenes para los retos de la industria, frente a este panorama el horizonte presenta un déficit grande entre personal humano idóneo y las vacantes generadas por el subsector, así las cosas es claro que el gobierno de turno debe implementar políticas que busquen la preparación de nuestros estudiantes en áreas TIC, no por nada las políticas nacionales buscan generar nuevos conocimientos sobre esta materia, en este ámbito es de destacar la aparición del programa misión TIC 2022 un programa del gobierno nacional que busca preparar a 100.000 nacionales en diferentes áreas de programación, tal vez el esfuerzo más grande de un gobierno nacional para preparar su población para los nuevos retos empresariales, lo anterior seguramente como respuesta a la gran brecha que se presenta en la investigación.

Para entender los análisis es claro que las vacantes que se crean por departamento están enfocadas para los diferentes niveles de estudio, por lo tanto, no solo los profesionales recién graduados deben suplirlas, pero si se observa el nivel de la brecha creada es claro que el país tiene un déficit de personal humano TIC, en este orden de ideas es importante recomendar que las instituciones técnicas, tecnológicas y profesionales del país deben buscar canales que permitan ofrecer una comunicación directa con las empresas líderes en el mercado, con el objetivo de presentar los pónsum de los diferentes programas académicos amoldados a las necesidades del país, y que estos no sean alineados con las necesidades mundiales o de la región, pues es claro que la gran mayoría de los graduados nacionales buscan aplicar sus conocimientos en las empresas del país, por lo anterior, es

clave implementar estos canales para establecer los programas académicos que permitan afrontar los retos de las empresas nacionales. Las empresas TIC del país tienen en su personal asociado un 92% con alto nivel educativo iniciando por técnicos, tecnólogos y profesionales y solo un 8% sin esta formación, asociando estos resultados a los presentados por el estudio de identificación de brechas de capital humano para el sector TIC presentado por el Ministerio TIC los datos son muy similares pues en este se concluye que en el sector existe un 28% de técnicos, 51% de profesionales y 15% de postgrado y superior, lo que concluye en la especialización del personal humano requerido por la industria.

Como competencias necesarias en la actualidad se evidencia dentro del proceso investigativo la necesidad de contar con conocimientos en lenguajes de programación como SQL, PHP y Angular, análisis de sistemas de información y en especial pensamiento algorítmico y como competencias generales es necesario intensificar la educación en orientación al cliente y liderazgo.

## **11. VISUALIZADOR DE RESULTADOS**

La herramienta para la visualización de datos que hace parte de los objetivos de la investigación se presenta como un informe interactivo en Power BI, pues por medio de esta es posible presentar de forma ordenada y clara los resultados obtenidos del proceso de minería de datos a partir de la información recolectada de diferentes fuentes reconocidas.

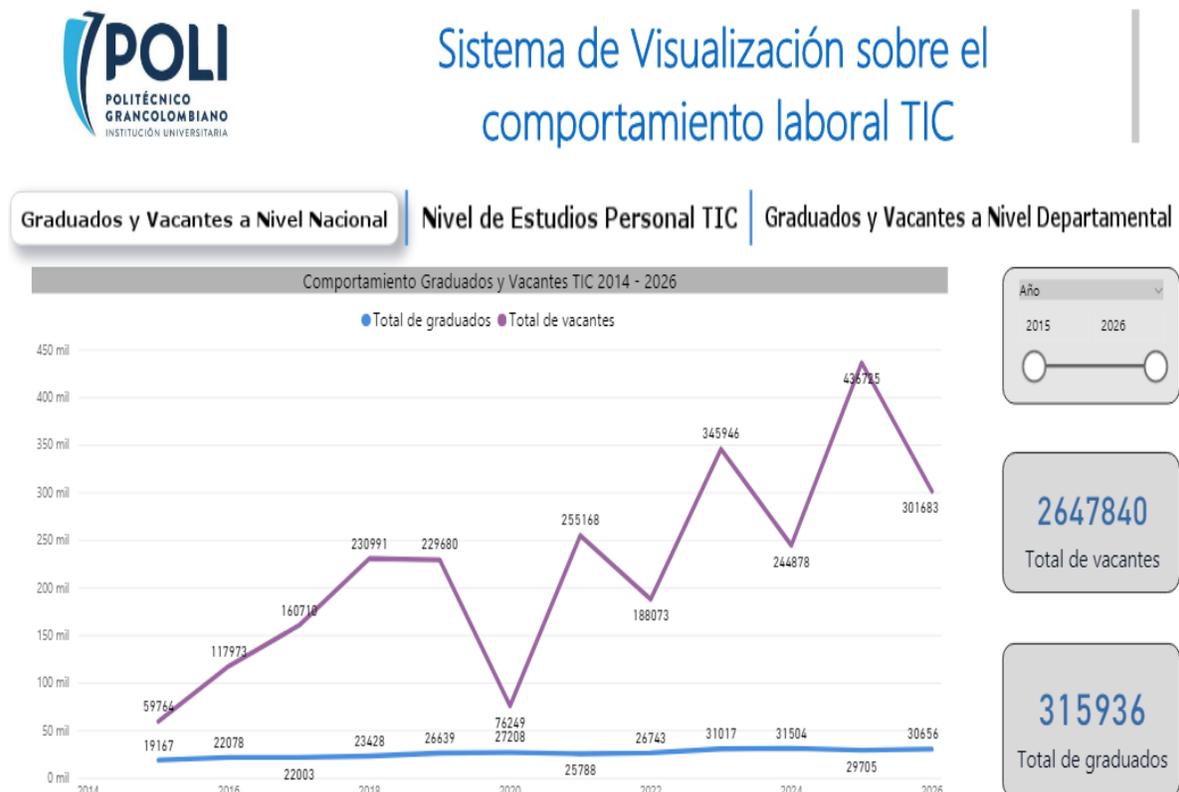
Esta herramienta informática posibilita la presentación de los datos de tal forma que pueda ser compartida para cualquier usuario que tenga un dispositivo con acceso a internet y un navegador web, además a lo anterior se suma la posibilidad de ser incluido en

diferentes sitios web con la facilidad de proveerlo como un elemento HTML nativo para esta función.

El enlace para observar esta herramienta es el siguiente:

<https://app.powerbi.com/reportEmbed?reportId=32ae48ea-388c-4d95-9c3e-567c850502da&autoAuth=true&ctid=dd505be5-ec69-47f5-92df-caa55febf5fa&config=eyJjbHVzdGVyVXJsIjoiaHR0cHM6Ly93YWJpLXNvdXRoLWNIbnRyYWwtdXMtcmVkaXJlY3QuYW5hbHlzaXMud2luZG93cy5uZXQvIn0%3D>

Ilustración 383. Interfaz inicial herramienta de visualización



Fuente: elaboración propia.

Ilustración 394. Interfaz comportamiento graduados y vacantes a nivel departamental.



Fuente: elaboración propia.

## 12. CONCLUSIONES

A lo largo de la investigación se observa que no existe una alianza entre academia e industria para la búsqueda de la formación acertada de la población en edad para realizar sus estudios superiores en el país, pues cada periodo en los diferentes departamentos la brecha entre graduados y vacantes generadas es mucho mayor, esto quiere decir que las necesidades de la industria no son totalmente satisfechas por los profesionales que suministra la academia, con virtud en lo anterior es necesario, que se instalen grupos de

acompañamiento en los dos organismos que busquen la identificación de las competencias necesarias para afrontar los retos actuales de la industria y de esta forma disminuir la brecha de estos dos indicadores al igual que el aumento en la calidad de los productos y servicios de la industria.

Es importante que las políticas económicas del gobierno tengan coherencia con las educativas a nivel nacional, por esta razón, es fundamental que se aumenten las horas de enseñanza en los colegios de las materias enfocadas a las ciencias TIC, así como la obligatoriedad de los fundamentos de programación en los niveles primarios y secundarios de la población nacional, pues no solo el sector TIC es un pilar de las políticas económicas del gobierno, sino una obligación para todos los seres humanos en la actualidad, puesto que todos los campos del saber se han visto influenciados positivamente por la implementación de la tecnología y programación en sus procesos, con esto no solo se prepara a los nacionales para el futuro, sino que se aumenta el número de estudiantes que adquieren un mayor gusto y preferencia hacia estas carreras profesionales.

De acuerdo a la investigación se evidencia que la tecnología ha llegado a todas las regiones nacionales, pues si bien es cierto existen departamentos que no cuentan con Institución de Educación Superior en todas las regiones nacionales se presentan vacantes para cargos TIC, esto concluye que las regiones están en capacidad de implementar diferentes procedimientos tecnológicos para sus procesos y se convierten en un nicho de mercado para las diferentes organización de la industria, en este mismo sentido se hace necesario apropiarse del conocimiento adecuado a todas las comunidades nacionales en la búsqueda de la interconexión nacional, que permita a todos los colombianos por igual las

mismas capacidades de educación y seguramente el aumento de personal TIC con amplio talento.

El grado de especialización del sector TIC es evidente, por esta razón la mayor parte de sus vacantes se presentan con algún tipo de formación superior, en este sentido es importante preparar al futuro personal TIC desde los fundamentos de las diferentes materias que son la base del conocimiento, lo anterior en razón a que los pronósticos muestran como el personal que solo cuenta con cursos cortos de una herramienta o lenguaje de programación en particular disminuye con el paso del tiempo, lo que hace evidente la necesidad de las empresas nacionales de contratar personal con un tiempo mayor de preparación, lo que significa que el talento humano TIC con un alto grado de especialización tiene una mayor probabilidad de disminuir la rotación en las diferentes vacantes y aumenta la calidad de productos y servicios de la industria.

La investigación está en capacidad de ser complementada con el estudio del aporte de los diferentes programas universitarios a la oferta general de graduados, así como la búsqueda del patrón de rotación de los diferentes cargos de la industria, también es necesario buscar la individualización de los cargos ofrecidos por las diferentes empresas con el objetivo de identificar los cargos con mayor número de demanda para ser priorizados por las instituciones de educación superior nacionales.

### 13. REFERENCIAS

- Aldana Galvis, J., Nieves Lombo, M., & Sánchez Herrán, P. (12 de 2018). *Universidad Piloto de Colombia*. Obtenido de <http://repository.unipiloto.edu.co>: <http://repository.unipiloto.edu.co/handle/20.500.12277/5999>
- Camana, R. (2016). Potenciales Aplicaciones de la Minería de Datos en Ecuador. *Revista Tecnológica ESPOL*, 29(1), 170-183. Obtenido de <http://www.rte.espol.edu.ec/index.php/tecnologica/article/view/464/338>
- Departamento Nacional de Estadística. (11 de 12 de 2019). *DANE*. Obtenido de <https://www.dane.gov.co>: <https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion/formacion-para-el-trabajo>
- Departamento Nacional de Estadística. (12 de 06 de 2020). *DANE*. Obtenido de <https://www.dane.gov.co>: <https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion/fuerza-laboral-y-educacion>
- Educación Revista Semana. (03 de 03 de 2019). *Revista Semana*. Obtenido de <https://www.semana.com>: <https://www.semana.com/educacion/articulo/por-que-estudiar-tecnologia/607932/>
- Espinosa Zuñiga, J. (2020). Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. *Ingeniería, investigación y tecnología*, 21(1), 22-29. Obtenido de [http://www.scielo.org.mx/scielo.php?pid=S1405-77432020000100008&script=sci\\_arttext](http://www.scielo.org.mx/scielo.php?pid=S1405-77432020000100008&script=sci_arttext)
- Espinosa, P. (11 de 04 de 2019). *Diario La República*. Obtenido de <https://www.larepublica.co>: <https://www.larepublica.co/empresas/sector-de-tecnologias-tiene-deficit-de-45000-empleados-en-colombia-2850190>
- Flores Lagla, G. A., Cadena Moreano, J. A., & Villa Quisphe, M. W. (15 de 01 de 2019). Minería de datos como herramienta estratégica. *Revista Científica Mundo de la Investigación y el Conocimiento*, 3(1), 955-970. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=6796766>
- Giraldo Mejía, J. C., & Vargas Agudelo, F. A. (2019). Aplicación de la técnica regresión logística de la minería de datos en el proceso de escubrimiento de conocimiento (KDD) en bases de datos operativas o transaccionales. *Perspectivas Revista de*

- Tecnología e Información*, 51-55. Obtenido de <http://revistas.uigv.edu.pe/index.php/perspectiva/article/view/568/493>
- Gonzalez Farroñan, L. A. (12 de 2018). *Universidad Católica Santo Toribio de Mogrovejo*. Obtenido de Repositorio de Tesis USAT: <http://54.165.197.99/handle/20.500.12423/2192>
- Gorbea Portal, S., & Madera Jaramillo, M. (2017). Diseño de un data warehouse para medir el desarrollo disciplinar en instituciones académicas. *Investigación bibliotecológica*, 31(72), 24-38. Obtenido de [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0187-358X2017000200161](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-358X2017000200161)
- Grande, M., Cañon , R., & Cantón , I. (2016). Tecnologías de la información y la comunicación: evolución del concepto y características. *Revsita Internacional de investigación e innovación educativa*, 218-230. Obtenido de <https://www.upo.es/revistas/index.php/IJERI/article/view/1703/1559>
- Gutiérrez O., J., & Molina , B. (2018). Identificación de herramientas básicas para la aplicación de la minería de datos en la solución de problemas empresariales. *Dirección de Gestión de Conocimiento de la Universidad CES*, 33-52.
- Informatica UV. (12 de 05 de 2001). *Informatica UV España*. Obtenido de <http://informatica.uv.es>: <http://informatica.uv.es/iiguia/AED/oldwww/EDS/9900/PR6/>
- Martinez Heras, J. (10 de 02 de 2020). *Iartificial*. Obtenido de <https://www.iartificial.net>: <https://www.iartificial.net/regresion-lineal-con-ejemplos-en-python/>
- Marulanda Echeverry, C., López Trujillo, M., & Mejía Salazar, M. (06 de 2017). *Universidad Católica del Norte*. Obtenido de <https://revistavirtual.ucn.edu.co>: <https://revistavirtual.ucn.edu.co/index.php/RevistaUCN/article/view/821>
- Ministerio de las Tecnologías de información y la telecomunicaciones. (12 de 11 de 2019). *MinTIC*. Obtenido de <https://www.mineduacion.gov.co>: <https://www.mineduacion.gov.co/sistemasinfo/Informacion-a-la-mano/212400:Estadisticas>
- Ministerio de las Tecnologías de la Información y las Comunicaciones. (07 de 11 de 2018). *MinTIC*. Obtenido de <https://colombiatic.mintic.gov.co>: <https://colombiatic.mintic.gov.co/679/w3-propertyvalue-36671.html>
- Ministerio de Tecnologías de la Información y las Comunicaciones . (2018). *MINTIC*. Obtenido de [www.mintic.org](http://www.mintic.org): [https://estrategia.gobiernoenlinea.gov.co/623/articles-81473\\_recurso\\_1.pdf](https://estrategia.gobiernoenlinea.gov.co/623/articles-81473_recurso_1.pdf)

- Ministerio de Tecnologías de la Información y las Comunicaciones. (19 de 08 de 2020). *MinTIC*. Obtenido de <https://mintic.gov.co>: <https://mintic.gov.co/portal/inicio/Sala-de-Prensa/MinTIC-en-los-Medios/150116:Gobierno-lanza-estrategia-para-formar-100-mil-programadores-en-2022>
- Ministerio Nacional del Trabajo. (03 de 10 de 2019). *Fuente de Información Laboral de Colombia*. Obtenido de <http://filco.mintrabajo.gov.co>: <http://filco.mintrabajo.gov.co/FILCO/faces/indicadores.jsf?nombre=Numero+de+graduados+en+educaci%F3n+superior&ind=46>
- MinTIC. (20 de 07 de 2020). *Gobierno Nacional de Colombia*. Obtenido de <https://www.mintic.gov.co>: <https://www.mintic.gov.co/portal/inicio/5306:Gobierno-en-L-nea-GEL>
- MinTIC. (12 de 06 de 2020). *Mision TIC 2022*. Obtenido de <https://www.misontic2022.gov.co>: <https://www.misontic2022.gov.co/portal/>
- Miranda, M., & Gúzman, J. (2017). Análisis de la Deserción de Estudiantes Universitarios usando Técnicas de Minería de Datos. *Formación universitaria*, 10(3), 61-68. Obtenido de [https://scielo.conicyt.cl/scielo.php?pid=S0718-50062017000300007&script=sci\\_arttext&tlng=n](https://scielo.conicyt.cl/scielo.php?pid=S0718-50062017000300007&script=sci_arttext&tlng=n)
- Ordoñez, Y. (2011). Herramienta de Minería de Uso de la Web Aplicado a los Registros del Proxy. *Research Gate*, 18/26. Obtenido de [https://www.researchgate.net/publication/233426470\\_HERMINWEB\\_Herramienta\\_de\\_Mineria\\_de\\_Uso\\_de\\_la\\_Web\\_Aplicado\\_a\\_los\\_Registros\\_del\\_Proxy](https://www.researchgate.net/publication/233426470_HERMINWEB_Herramienta_de_Mineria_de_Uso_de_la_Web_Aplicado_a_los_Registros_del_Proxy)
- Ordoñez, Y., & Darian, H. (2011). HERMINWEB Herramienta de Minería de Uso de la Web Aplicado a los Registros del Proxy. *Research Gate*, 65-72.
- Pincay Ponce, J., Angulo Murillo, N., Herrera Tapia, J., & Delgado Muentes, W. (2020). TÉCNICAS DE MINERÍA DE DATOS COMO SOPORTE PARA LA GESTIÓN DE UN SISTEMA DE COMERCIALIZACIÓN DE ENERGÍA ELÉCTRICA. *Revista Científica Multidisciplinaria*, 38-49. Obtenido de [https://www.researchgate.net/profile/Jorge-Pincay-Ponce/publication/347712864\\_Tecnicas\\_de\\_mineria\\_de\\_datos\\_como\\_sopORTE\\_para\\_la\\_gestion\\_de\\_un\\_sistema\\_de\\_comercializacion\\_de\\_energia\\_electrica/links/5fe35ba6a6fdccdb8f58526/Tecnicas-de-mineria-de-datos-como](https://www.researchgate.net/profile/Jorge-Pincay-Ponce/publication/347712864_Tecnicas_de_mineria_de_datos_como_sopORTE_para_la_gestion_de_un_sistema_de_comercializacion_de_energia_electrica/links/5fe35ba6a6fdccdb8f58526/Tecnicas-de-mineria-de-datos-como)
- Portal Público de Datos Abiertos. (21 de 01 de 2020). *Datos Públicos*. Obtenido de <https://www.datos.gov.co>: [https://www.datos.gov.co/Educacion/MEN\\_ESTADISTICAS-MATRICULA-POR-MUNICIPIOS\\_ES/y9ga-zwzy](https://www.datos.gov.co/Educacion/MEN_ESTADISTICAS-MATRICULA-POR-MUNICIPIOS_ES/y9ga-zwzy)
- RapidMiner. (01 de 2020). *RapidMiner*. Obtenido de <https://docs.rapidminer.com/>: <https://docs.rapidminer.com/>

- Revista Semana Tecnología. (09 de 01 de 2020). *Revista Semana*. Obtenido de semana.com: <https://www.semana.com/educacion/articulo/el-diploma-universitario-se-esta-quedando-obsoleto--noticias-bogota-hoy/621848/>
- Romero Romero, C. A. (28 de 05 de 2018). *UDEC Universidad de Cundinamarca*. Obtenido de <http://repositorio.ucundinamarca.edu.co>: <http://repositorio.ucundinamarca.edu.co/handle/20.500.12558/1086>
- Ruiz, S., & Myriam, H. (07 de 2019). *Universidad Nacional de la Plata*. Obtenido de <http://sedici.unlp.edu.ar>: <http://sedici.unlp.edu.ar/handle/10915/90628>
- Silva Montes , C. (2018). Perfil de egreso y empleo en el contexto del avance tecnológico. *Universidad Pedagógica Nacional Pedagogía y Saberes*, 83-96. Obtenido de <http://www.scielo.org.co/pdf/pys/n48/0121-2494-pys-48-00083.pdf>
- SITEAL/TIC. (11 de 2019). *Sistema de información de tendencias educativas en America Latina*. Obtenido de Organización de las Naciones Unidas: <http://www.publicaciones.siteal.iipe.unesco.org/eventos-del-siteal/31/formacion-tic-y-empleo-para-los-jovenes-desafios-y-posibles-cursos-de-accion>
- Unidad del Servicio de Empleo. (25 de 11 de 2019). *Servicio de Empleo*. Obtenido de <https://www.serviciodeempleo.gov.co>: <https://www.serviciodeempleo.gov.co/estudios-e-investigacion/oferta-y-demanda-laboral/anexo-estadistico-de-oferta-laboral-buscadores>
- Uvidia Fassler, M., Cisneros Barahona, A., Méndez Naranjo, P., & Villa Yáñez , H. (2017). Minería de datos para la toma de decisiones en la unidad de nivelación y admisión universitaria ecuatoriana. *Revista Cumbres*, 55-67.