



**ESCUELA DE NEGOCIOS**

**MAESTRÍA EN INTELIGENCIA DE NEGOCIOS Y CIENCIA DE DATOS**

**APLICACIÓN DE ESTRATEGIAS DE TRANSFORMACIÓN DIGITAL E  
INNOVACIÓN EN EL SECTOR EMPRESARIAL EN ECUADOR**

**Profesor**

**Mario Salvador González**

**Autores**

**Guido Yangón Pintag  
Alejandro Reyes Guzmán**

**2023**

## RESUMEN

La investigación se basa en una muestra representativa de empresas de diversos tamaños, que recopila información sobre variables relacionadas con la inversión, el uso y la conexión a internet, el comercio electrónico, el hardware y software, los medios de comunicación y la capacitación, el cual tiene como objetivo analizar la inversión en Tecnologías de la Información y Comunicación (TIC) y su impacto en diversos aspectos empresariales.

Para iniciar se realiza un análisis descriptivo con el fin de comprender la situación actual referente a la temática abordada, donde se revela una brecha significativa en la inversión en TIC entre empresas de diferentes tamaños. Adicional se identifican patrones de inversión y se observa una falta de adopción de tecnologías emergentes en ciertas herramientas.

Los modelos predictivos por utilizar son, Clustering, logístico, random forest y lineal, los mismos que permiten predecir la probabilidad de inversión en TIC y analizar su impacto en variables clave, como el nivel de madurez tecnológica y el comercio electrónico, además de entender la importación de estas herramientas al momento de la toma de decisiones a nivel directivo.

En base a los resultados obtenidos, se propone una estrategia prescriptiva que busca abordar la problemática encontrada. Se sugiere promover la adopción de tecnologías emergentes a través de inversiones focalizadas y programas de capacitación. Asimismo, se plantea fomentar la colaboración entre empresas y partners para desarrollar soluciones integrales que busque avanzar hacia una sólida transformación e innovación digital.

Palabras clave: transformación digital, innovación, inversión, usos, conexión, gestión, hardware, software, medios de comunicación, obstáculos, ventas, compras, capacitación, especialistas.

## **ABSTRACT**

The research is based on a representative sample of companies of various sizes, which collects information on variables related to investment, use and connection to the Internet, electronic commerce, hardware and software, the media and training, the which aims to analyze investment in Information and Communication Technologies (ICT) and its impact on various business aspects.

To start, a descriptive analysis is carried out in order to understand the current situation regarding the topic addressed, where a significant gap in ICT investment between companies of different sizes is revealed. Additionally, investment patterns are identified and a lack of adoption of emerging technologies in certain tools is observed.

The predictive models to be used are clustering, logistic, random forest and linear, the same ones that allow predicting the probability of investment in ICT and analyzing its impact on key variables, such as the level of technological maturity and electronic commerce, in addition to understanding the import of these tools at the time of decision-making at the management level.

Based on the results obtained, a prescriptive strategy is proposed that seeks to address the problems encountered. It is suggested to promote the adoption of emerging technologies through targeted investments and training programs. Likewise, it is proposed to promote collaboration between companies and partners to develop comprehensive solutions that seek to move towards a solid transformation and digital innovation.

Keywords: digital transformation, innovation, investment, uses, connection, management, hardware, software, media, obstacles, sales, purchases, training, specialists.

## ÍNDICE DEL CONTENIDO

RESUMEN .....	2
ABSTRACT .....	3
1. INTRODUCCIÓN .....	1
2. REVISIÓN DE LITERATURA .....	2
3. IDENTIFICACIÓN DEL OBJETO DE ESTUDIO .....	9
4. PLANTEAMIENTO DEL PROBLEMA .....	11
5. OBJETIVO GENERAL .....	13
6. OBJETIVOS ESPECÍFICOS .....	13
7. JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA.....	14
7.1 SELECCIÓN DE LA BASE DE DATOS .....	14
7.2 LIMPIEZA, PREPROCESAMIENTO Y/O TRANSFORMACIÓN DE DATOS.....	15
7.3 IDENTIFICACIÓN Y DESCRIPCIÓN DE VARIABLES.....	18
7.4 SELECCIÓN DEL MODELO ESTADÍSTICO.....	32
8. RESULTADOS .....	38
8.1 ANÁLISIS DE MODELO ESTADÍSTICO E INTERPRETACIÓN DE RESULTADOS.....	38
8.1.1 MODELOS DE REGRESIÓN .....	38
8.1.2 ANÁLISIS MEDIANTE REGRESION LINEAL LOG LIN .....	38
8.1.3 ANÁLISIS MEDIANTE EL ALGORITMO K-MEANS .....	42
8.1.4 ANÁLISIS MEDIANTE REGRESIÓN LOGÍSTICA.....	47
8.1.5 ANÁLISIS MEDIANTE RANDOM FOREST FEATURE IMPORTANCE.....	54
8.1.6 ANÁLISIS MEDIANTE RANDOM FOREST REGRESSOR.....	58
9. PROPUESTA DE SOLUCIÓN .....	62
9.1 IMPLICACIONES PARA LA ORGANIZACIÓN .....	62
9.2 ESTRATEGIA ORGANIZACIONAL .....	64
9.3 ESTRATEGIA ORGANIZACIONAL, INNOVACIÓN Y COMPETITIVIDAD EMPRESARIAL.....	68
10. CONCLUSIONES Y RECOMENDACIONES.....	70
10.1 CONCLUSIONES .....	70
10.2 RECOMENDACIONES.....	72
11. REFERENCIAS .....	74

## ÍNDICE DE TABLAS

Tabla 1 <i>Diccionario de variables</i> .....	18
Tabla 2 <i>Resultados modelos de aprendizaje supervisado</i> .....	38
Tabla 3 <i>Grupos de variables numéricas</i> .....	38
Tabla 4 <i>Coefficientes regresión lineal</i> .....	39
Tabla 5 <i>Porcentajes de Inversión TIC</i> .....	42
Tabla 6 <i>Análisis con 3 clústeres con la media.</i> .....	44
Tabla 7 <i>Coefficientes regresión ventas por internet</i> .....	47
Tabla 8 <i>Matriz de confusión ventas por internet</i> .....	49
Tabla 9 <i>Coefficientes compras por internet</i> .....	51
Tabla 10 <i>Matriz de confusión compras por internet</i> .....	53

## ÍNDICE DE FIGURAS

Figura 1	<i>Conformación de base de datos</i>	14
Figura 2	<i>Representación de variables con valores perdidos</i>	16
Figura 3	<i>Unión y suma de variables personal y especialistas</i>	17
Figura 4	<i>Promedio de inversión TIC</i>	19
Figura 5	<i>Porcentaje de inversión de acuerdo con las áreas</i>	20
Figura 6	<i>Porcentaje de tipo de conexión</i>	21
Figura 7	<i>Usos de internet</i>	22
Figura 8	<i>Cronológico de Usos de internet</i>	23
Figura 9	<i>Porcentaje de uso de páginas web</i>	24
Figura 10	<i>Cronológico de Ecommerce</i>	25
Figura 11	<i>Porcentaje de Ecommerce</i>	25
Figura 12	<i>Cronológico de uso de dispositivos</i>	26
Figura 13	<i>Cronológico de uso de software</i>	27
Figura 14	<i>Gestión de TICs</i>	28
Figura 15	<i>Cronológico de gestión del conocimiento</i>	29
Figura 16	<i>Cronológico de uso de medios de comunicación</i>	30
Figura 17	<i>Cronológico de personal vs especialistas</i>	31
Figura 18	<i>Cronológico de capacitación</i>	32
Figura 19	<i>Representación de Agrupación mediante K-Means</i>	34
Figura 20	<i>Modelo Log – Lin Valores observados vs Valores predichos</i>	40
Figura 21	<i>Método del codo - Escalado</i>	42
Figura 22	<i>Agrupación por 3 clústeres</i>	43
Figura 23	<i>Agrupación por 3 clústeres</i>	44
Figura 24	<i>Visualización de clúster vs porcentajes de inversión</i>	45
Figura 25	<i>Características más importante por Feature Importance</i>	54
Figura 26	<i>Relación entre Inversión TIC y Personal TIC</i>	55
Figura 27	<i>Relación entre Inversión TIC y Capacitación Especialistas TIC</i>	56
Figura 28	<i>Relación Inversión TIC y No Capacitación TIC</i>	57
Figura 29	<i>Relación entre Inversión en TIC y SA Servidores Web</i>	58
Figura 30	<i>Análisis valores observados vs. valores predichos</i>	59
Figura 31	<i>Gráfico de dispersión Observados vs Predichos</i>	60

Figura 32 *Ajuste para Error de Predicción* ..... 61

## 1. INTRODUCCIÓN

El desarrollo económico y el auge de nuevas tecnologías es muy a menudo el producto de varios cambios económicos y sociales. La transformación digital es una de las manifestaciones más recientes de dichos cambios, y esta se refiere al proceso de utilizar tecnologías digitales para mejorar la eficiencia y la eficacia de una empresa en la realización de sus actividades. Este proceso implica que la organización desarrolle una amplia gama de capacidades, cuya importancia variará según el contexto comercial y las especificaciones que la organización necesite, a la par se requiere la integración de tecnologías digitales en todos los aspectos del negocio, incluyendo la automatización de procesos, la recopilación y análisis de datos, la colaboración en línea y la adopción de nuevas tecnologías, convirtiéndose así en un elemento central el modo de operación de un negocio.

Es así como, a raíz de la gran cantidad de información que se produce diariamente, se requiere una herramienta que permita convertir dichos datos en información, información en conocimiento y estrategias para liderar de forma eficiente las distintas actividades que intervienen en los negocios (Cano, 2007). Ese término paraguas se denomina inteligencia de negocios (BI, por sus siglas en inglés), y se refiere a las herramientas tecnológicas utilizadas para recopilar, analizar y presentar información útil y relevante a los líderes empresariales y a los tomadores de decisiones, con el fin de ayudarles a comprender mejor el rendimiento de la empresa en base al conocimiento adquirido.

Por lo tanto, diremos que la innovación de nuevas ideas, tecnologías, procesos y productos en el Ecuador, sufrieron un antes y un después, considerando el nivel creciente de madurez tecnológica empresarial. Por tal razón es importante que las organizaciones no solo se centren en implementar herramientas tecnológicas para mejorar la experiencia del consumidor, sino que, a la par se debe establecer los objetivos que se desea conseguir mediante la transformación digital, tanto en la cultura organizacional como en los modelos de toma de decisiones basados en datos.

## 2. REVISIÓN DE LITERATURA

Hasta hace pocos años, las empresas tradicionales solían tener un establecimiento físico para brindar sus servicios a los clientes. Sin embargo, con la irrupción de internet, las empresas, tanto pequeñas como grandes, se ven en la necesidad de actualizarse y adoptar la digitalización como un nuevo modelo de negocio. Esta transformación implica no solo tener presencia en línea, sino también utilizar nuevas tecnologías y herramientas para mejorar los procesos empresariales y la experiencia del cliente (Vásquez & Orozco, 2018). En particular, la digitalización ofrece oportunidades de crecimiento y expansión para las Pymes, permitiéndoles competir con las grandes empresas y llegar a nuevos mercados de manera más eficiente y económica. Por esto, la digitalización es una oportunidad para que las empresas tradicionales evolucionen y prosperen en la era digital actual.

La universalización y la expansión de Internet han creado una infraestructura informática cada vez más sólida en la economía global. Esta infraestructura permite a las empresas conectarse y colaborar de manera más rápida y efectiva, lo que puede resultar en modelos de negocio más eficientes y competitivos. La creciente adopción de la tecnología en todas las áreas de la economía también ha llevado a la creación de nuevos modelos de negocio que aprovechan las ventajas de la Internet; estos modelos a menudo involucran la utilización de plataformas digitales para conectar a diferentes actores del mercado, lo que permite una mayor eficiencia y generación de valor (Fernández & Puig, 2020). Es fundamental destacar que las organizaciones deben mejorar su capacidad tecnológica para ser más innovadoras y competitivas en los mercados abiertos actuales. De esta manera, podrán aumentar su participación en el mercado, lo que representa el primer paso para lograr una estrategia competitiva efectiva: la estrategia tecnológica, por lo cual es preciso indicar que: “Las empresas con vocación innovadora siempre están orientadas a crear nuevos clientes, a abrir nuevos mercados, aprovechando sistemáticamente su capacidad tecnológica acumulada de aprender haciendo durante varios años e identificando tendencias y oportunidades del mercado” (Ortega, 2000).

En este sentido, debemos responder a la interrogante, ¿Qué es transformación digital?, por lo cual, primero es importante mencionar que la digitalización es el proceso que nos permite transformar los procesos analógicos y la información en un formato físico a uno digital. La transformación digital se refiere al uso de habilidades y herramientas digitales para optimizar los procesos, productos y recursos de una empresa con el fin de aumentar su eficiencia. El objetivo de la transformación digital es aprovechar las oportunidades que surgen gracias a la aparición de nuevas tecnologías y estrategias de negocio. Principalmente, se trata de una reestructuración en la forma de hacer negocios, aprovechando las nuevas posibilidades y herramientas digitales disponibles para mejorar la eficacia y rentabilidad de la empresa (Waltermán, 2021).

De esta forma, resulta necesario mencionar que la transformación digital se ha convertido en una necesidad para las empresas que buscan mejorar la calidad de sus soluciones para los clientes. Este proceso implica la adopción de herramientas tecnológicas avanzadas como Big Data, la nube y la inteligencia artificial, y a su vez, implica cambios en el modelo de negocio y la cultura de la organización. Por lo cual, la transformación digital conlleva una reestructuración de la forma en que se hacen negocios, promoviendo una nueva mentalidad en la empresa y en la manera de gestionarla (Vásquez & Orozco, 2018).

El entorno empresarial actual requiere que las organizaciones reconsideren sus estrategias empresariales, para adaptarse a un mercado cada vez más global y competitivo. En este contexto, la innovación y las tecnologías de la información y comunicación (TIC) son factores clave para el desarrollo y deben ser considerados en el diseño e implementación de las estrategias empresariales, ya que pueden mejorar las organizaciones (Maldonado, Mojica, & Molina, 2013).

Según Escorza & Valls (2001), definen a la innovación como la generación de nuevos productos, servicios, tecnologías y sistemas de gestión empresarial, lo cual puede tener un impacto positivo significativo tanto en el rendimiento de las organizaciones como en la mejora del entorno empresarial que promueve la productividad de las empresas. Además, el uso de tecnologías puede ayudar a las organizaciones a ser más ágiles y flexibles en un entorno empresarial cada

vez más competitivo, permitiendo un mayor nivel de automatización en los procesos y una mayor conectividad entre diferentes áreas y sistemas.

Los avances en la tecnología han permitido que cualquier organización, sin importar su tamaño, pueda implementar un sistema de información organizacional en la actualidad. En efecto, en los últimos años, cada vez más pequeñas y medianas organizaciones han comenzado a revisar e implementar sistemas de información en sus procesos administrativos y operacionales, comenzando por lo general con aplicaciones contables como cuentas por cobrar, nómina y libro mayor general (Mergel, Edelmann & Huag, 2019).

Considerando este contexto previo, es crucial que las pequeñas y medianas empresas se adapten a los avances tecnológicos para mantenerse en el mercado y ser más competitivas a nivel global, por lo que la tecnología dura es fundamental e indispensable (Oliva, 2019). Bajo este orden de ideas, es importante analizar los diferentes aspectos que intervienen en el proceso de transformación digital e innovación.

A tal efecto, la inversión en Tecnologías de la Información y Comunicación (TIC) ha tenido un gran impacto en el desarrollo empresarial a nivel mundial. Según una investigación realizada por la Organización para la Cooperación y el Desarrollo Económicos (OCDE), la inversión en TIC ha demostrado ser uno de los principales impulsores del crecimiento empresarial, ya que permite una mayor eficiencia en la producción y mejora la capacidad de innovación y la capacidad de adaptación de las empresas a los cambios del mercado (Organisation for Economic Co-operation and Development OECD, 2019).

En la misma línea y de acuerdo con el Banco Interamericano de Desarrollo (BID) en 2019, las empresas latinoamericanas destinan alrededor del 2% de sus ingresos a la inversión en tecnología. Entre los principales destinos de la inversión se encuentran el hardware y software, la consultoría y servicios de telecomunicaciones. Además, la inversión en TIC también ha llevado la expansión de las empresas a nivel global, ya que facilita el acceso a nuevos mercados y clientes. Según la consultora PwC, las empresas que invierten en

tecnología digital han experimentado un aumento del 16% en su margen de beneficio, en comparación con aquellas que no invierten (PwC, 2017).

En América latina, un ejemplo de inversión en TICs es la plataforma de comercio llamada Mercado Libre, los cuales han aumentado sus ventas en un 20%. (Forbes, 2020). En el contexto ecuatoriano, la empresa Cervecería Nacional ha invertido en tecnología y automatización de procesos, permitiendo una reducción del consumo de energía eléctrica en un 17 % y uso de agua en un 15 %. (El Universo, 2019). En cuanto a la inversión para la consultoría en TIC, un estudio realizado por la consultora Gartner en 2020, encontró que la consultoría de estrategia es la principal área de inversión en este rubro. La consultoría de estrategia ayuda a las empresas a definir y planificar sus estrategias de transformación digital, lo que es fundamental para el éxito de cualquier iniciativa tecnológica

Evidentemente, los servicios de telecomunicaciones son importantes para la transformación digital de las empresas, ya que permiten la conectividad y la comunicación entre diferentes dispositivos y sistemas. En una investigación de Frost & Sullivan en 2020, encontró que los principales destinos de la inversión en servicios de telecomunicaciones son la conectividad de red y la seguridad de la información.

Entre 2014 y 2021, la penetración de Internet en Ecuador ha aumentado significativamente, de acuerdo con los datos proporcionados por la Agencia de Regulación y Control de las Telecomunicaciones (ARCOTEL). En 2014, la penetración de Internet en el país era del 31,8%, mientras que, en 2021, la penetración de Internet alcanzó el 63,1%. Esto representa un aumento del 98% en la cantidad de personas que tienen acceso a Internet en el país. La conexión DSL ha sido una de las tecnologías de conexión a Internet más utilizadas en Ecuador durante los últimos años. En 2014, el 75,2% de las conexiones a Internet en el país eran de tipo DSL, según datos de la ARCOTEL. Sin embargo, desde entonces, esta tecnología ha perdido popularidad, en parte debido a la creciente disponibilidad de opciones de conexión por fibra óptica y banda ancha

móvil. En 2021, solo el 11,2% de las conexiones a Internet en el país eran de tipo DSL.

Por otro lado, la conexión por fibra óptica ha ganado terreno en Ecuador durante los últimos años. En 2014, solo el 2,8% de las conexiones a Internet en el país eran de tipo fibra óptica. Sin embargo, en 2021, este tipo de conexión se ha convertido en una de las opciones más populares, con el 37,7% de las conexiones a Internet en el país siendo de tipo fibra óptica. Otro aspecto importante es la conexión de banda ancha móvil, que en los últimos 9 años ha crecido de manera significativa en el país, con una evolución constante tanto en términos de cobertura como calidad de la señal, en donde es importante destacar que las tecnologías con mayor auge en este aspecto están basadas en la tercera y cuarta generación (3G/4G), mismas que han permitido a los usuarios tener acceso a Internet de alta velocidad en sus dispositivos móviles en cualquier momento y lugar.

Considerando lo descrito en el párrafo anterior, una pregunta que se debería hacer es como el medio de conexión repercute de manera significativa la productividad, la eficiencia y la calidad del trabajo en general. Por lo cual, es importante considerar que la velocidad y calidad en la conectividad que brinda la fibra óptica que va desde los 50 a 100 megabits por segundo (Mbps), ha permitido establecer nuevas modalidades de trabajo como la tele presencialidad, además de servicios, plataformas e infraestructura montada en la nube o también denominada cloud computing, donde no existe limitación de recursos, el acceso o manipulación de los datos y aplicaciones se lo puede hacer desde cualquier lugar y a cualquier hora, y por sobre todas las cosas, está dirigida para las pequeñas y medianas empresas (PYMES) que no tienen el presupuesto ni la capacidad de inversión para comprar y mantener su propia infraestructura informática.

En el caso de Ecuador, para el año 2018 se puede observar que las empresas medianas de Ecuador están haciendo un uso significativo del Cloud Computing en términos avanzados. Sin embargo, las tecnologías de gestión como el ERP, CRM y otras, así como el uso de servidores o intranet, tienen una baja tasa de

adopción. Por lo que es importante destacar que, exceptuando el uso de intranet, no se aprecian diferencias significativas entre los dos tipos de empresas medianas en cuanto a la adopción y uso de estas tecnologías (Dini, Gligo y Patiño, 2021).

Otro ejemplo importante fue la Encuesta de Acceso y Uso de TIC en Empresas, cuyo año de referencia es 2018, realizada por el Ministerio de Economía, Fomento y Turismo en colaboración con el Instituto Nacional de Estadísticas de Chile (2020). Esta encuesta se enfoca en las empresas formales que llevan a cabo su actividad económica principal en alguno de los sectores productivos establecidos según la Clasificación Internacional Industrial Uniforme de Naciones Unidas. En este caso los indicadores de uso avanzado de tecnologías permiten evidenciar que las aplicaciones ERP (22%) y Cloud Computing (18%) son las más utilizadas, mientras que el uso de herramientas de gestión empresarial como SCM, CRM o RFI es mucho menor. Estos resultados sugieren que, aunque la mayoría de las PYMES chilenas están conectadas, no están aprovechando estas tecnologías de manera intensiva. Además, la encuesta también proporciona información relevante sobre el nivel de uso de una de las tecnologías de frontera, Big Data, que solo es utilizada por el 1,7% de las pymes chilenas, lo que indica que su adopción es prácticamente nula en este tipo de empresas.

Con el análisis precedente es imposible dejar de mencionar la nueva forma de comercio electrónico o también denominado e-commerce de acuerdo con el Ministerio de Telecomunicaciones y de la Sociedad de la Información (s.f.) de Ecuador, este se refiere a la compra y venta de productos o servicios a través de redes informáticas y, principalmente, de internet. En este sentido es importante comprender que, debido a la amplia difusión del uso del internet en Ecuador, que abarca a más de la mitad de la población, se ha producido una expansión en este tipo de comercio, el cual utiliza las Tecnologías de la Información y Comunicación (TIC). En otras palabras, el fomento del comercio electrónico contribuye a dinamizar y diversificar la economía, ya que permite una mayor interacción entre vendedores y usuarios. Asimismo, el uso de las TIC para el comercio electrónico contribuye al cambio de la matriz productiva,

al permitir el desarrollo y creación de nuevas plataformas de venta para pequeñas y medianas empresas, tanto a nivel nacional como internacional.

Considerando la información proporcionada y de acuerdo con los datos disponibles para el año 2019, se observan algunas situaciones particulares en relación con el uso del comercio electrónico por parte de las pymes en distintos países, más allá de las diferencias esperadas en función del tamaño de las empresas. México se destaca por presentar niveles elevados de uso de esta herramienta, con un 74% de las pymes utilizándolo para compras y un 71% para ventas de sus productos o servicios. En Brasil, el uso del comercio electrónico para compras en empresas medianas y grandes registra niveles similares a los de México; sin embargo, en la venta, la diferencia con México es significativa. Por otro lado, aunque los datos de Chile muestran niveles bajos de este indicador con el 10% de comercio electrónico para ventas en PYMES, y el 17% en grandes empresas, es importante destacar que esto se debe en gran medida a que la encuesta utilizada cuenta con una definición más estricta del comercio electrónico que en otros países (Dini, Gligo y Patiño, 2021).

Después de todo, es importante tener en cuenta que la tecnología por sí sola no garantiza una transformación digital a través del éxito empresarial, sino que es necesario que las organizaciones tengan la capacidad y la cultura adecuadas para aprovechar al máximo las TIC y utilizarlas de manera efectiva para lograr sus objetivos estratégicos (Dominik, Modrák & Zsifkovits, 2020). Esto implica una inversión significativa en infraestructura, innovación de servicios y productos, talento humano con capacitación y desarrollo de habilidades continua, así como una estrategia clara y bien definida para la implementación y uso de la tecnología de la información. Considerando este contexto previo, es crucial que las pequeñas y medianas empresas se adapten a los avances y características descritas en este estudio, reconociendo la trascendencia de su implementación en el sector productivo.

### 3. IDENTIFICACIÓN DEL OBJETO DE ESTUDIO

Análisis comparativo del nivel de implementación tecnológica y de innovación en empresas de diferentes tamaños de Ecuador, mediante información actualizada y oportuna sobre estadísticas empresariales en el país, emitido por el INEC, como ente rector del Sistema Estadístico Nacional (INEC, 2022).

Esta investigación tiene como interés primordial analizar el nivel de implementación tecnológica y de innovación en empresas de diferentes sectores y comparar los resultados obtenidos. Para ello, se llevará a cabo un análisis de datos que nos permita incidir directamente sobre los alcances y limitaciones que han experimentado las empresas ecuatorianas pequeña, medianas y grandes desde el año 2016 hasta el año 2021, referente a la transformación digital empresarial. En este sentido, resulta importante mencionar que las pequeñas se caracterizan por contar con un personal ocupado entre 50 a 99; las medianas se caracterizan por tener un personal ocupado entre 100 a 99 personas; y, las grandes empresas con personal que va desde los 200 en adelante.

Se medirán diferentes aspectos relacionados con la implementación tecnológica, como la adopción de sistemas de información y comunicación (TIC), identificación de empresas que poseen conexión a internet, en función del uso que estas realizan con el servicio, la realización de actividades de comercio electrónico sea por medio de la compra o la venta de bienes y servicios a través de páginas web. Además, se evalúa la disponibilidad y uso de herramientas TIC, tanto a nivel de software como hardware dentro de las empresas, lo cual permite conocer el estado actual de estas herramientas en conjunción con el personal humano y el sector externo. De esta forma, a nivel de innovación se puede incidir en términos de la capacidad de generar nuevos protocolos empresariales, la implementación de procesos de investigación y desarrollo, y la capacitación al personal.

Los resultados obtenidos permitirán identificar las diferencias en el nivel de implementación tecnológica y de innovación entre empresas de diferentes tamaños, así como las barreras que enfrentan las empresas en la adopción de tecnologías y prácticas innovadoras. La información derivada de esta

investigación podrá ser utilizada por las empresas y los responsables de políticas públicas para mejorar sus estrategias de innovación y aumentar su competitividad en el mercado.

#### 4. PLANTEAMIENTO DEL PROBLEMA

La transformación digital es un proceso clave en la actualidad que busca mejorar la productividad y competitividad de las empresas mediante el uso de tecnologías digitales en todas sus áreas. En Ecuador, la adopción de tecnologías digitales en las empresas ha crecido en los últimos años, pero aún existen desafíos que dificultan su implementación efectiva. La falta de recursos, la resistencia al cambio, la falta de capacitación, entre otros factores, son barreras importantes que las empresas deben superar para lograr una transformación digital efectiva.

Además, la implementación de la transformación digital en las empresas de Ecuador se enfrenta a desafíos específicos, como la falta de acceso a tecnologías avanzadas, la brecha digital entre las empresas grandes y pequeñas, y la falta de políticas públicas y regulaciones que apoyen la transformación digital. Por lo tanto, es importante entender los desafíos en función de inversión en TIC, conexión y uso de internet, comercio electrónico, uso de herramientas TIC, medios de comunicación empleados y personal capacitado en las empresas ecuatorianas, para evaluar su desarrollo hacia la transformación digital y buscar soluciones efectivas que permitan a las empresas superar estas barreras y lograr una transformación efectiva que les permita competir en el mercado globalizado.

En consecuencia, el propósito de la presente investigación es evaluar los factores que pueden limitar la adopción de tecnologías para la transformación digital de las empresas en Ecuador. Por lo cual, las empresas que no estén preparadas para afrontar los cambios que implica la revolución digital, pueden quedar obsoletas y perder competitividad en el mercado. La implementación de tecnologías digitales, al igual que la digitalización de procesos incide en la empresa de manera integral, y aborda las fronteras de la empresa con potenciales clientes, competidores y proveedores; permite a las organizaciones mejorar la eficiencia en sus operaciones, reducir costos y aumentar la productividad. De esta forma, la transformación digital es clave para el éxito empresarial y actúa transversalmente en la sociedad, la economía y en el día a

día convirtiéndose en el acelerador de cambio eliminando las fronteras entre productos y servicios (Botto-Tobar, 2020).

## **5. OBJETIVO GENERAL**

Realizar un análisis integral del entorno empresarial en Ecuador, para identificar las necesidades, oportunidades, barreras y desafíos por tamaño de empresa en relación con la transformación digital, y promover la adopción de estrategias y tecnologías innovadoras que mejoren la eficiencia, productividad y competitividad de las empresas, fomentando el uso de tecnologías de la información y comunicación como herramientas habilitadoras para acceder a nuevos mercados de forma competitiva.

## **6. OBJETIVOS ESPECÍFICOS**

1. Aplicar técnicas de agrupación no supervisada para encontrar grupos de empresas en función de su campo de inversión de acuerdo con sus necesidades específicas y como estas contribuyen a la transformación digital y la innovación tecnológica.
2. Evaluar la relación entre las variables categóricas de compra y venta en línea, utilizando el modelo logístico para proporcionar insights y recomendaciones. Paralelamente se analizará las variables numéricas mediante el modelo lineal Log-Lin, donde se evaluará para ambos casos los coeficientes más significativos y como estos inciden en el ámbito de e-commerce y la inversión en TIC respectivamente.
3. Evaluar la precisión y robustez del modelo Random Forest en la predicción de la inversión en TIC, proporcionando recomendaciones y perspectivas para maximizar el retorno de la inversión y optimizar la estrategia de transformación digital.
4. Proponer estrategias específicas para que las empresas en Ecuador puedan superar las barreras identificadas y adopten procedimientos empresariales, a partir de herramientas tecnológicas en su proceso de transformación digital.

## 7. JUSTIFICACIÓN Y APLICACIÓN DE LA METODOLOGÍA

### 7.1 SELECCIÓN DE LA BASE DE DATOS

La base de datos escogida es una encuesta sobre el uso de Tics en las empresas a nivel nacional, la información recopilada puede ser desglosada según el tamaño de la empresa o la actividad económica utilizando la Clasificación Internacional Industrial Uniforme. Los datos se recaban mediante un formulario electrónico que es completado por los informantes, es decir, personas designadas por la empresa para que, con el asesoramiento continuo de los encuestadores del INEC, proporcionen los datos requeridos. El periodo de referencia está comprendido entre el año 2016 al año 2021, en donde constan 22591 registros y 68 campos, segmentados a partir de variables relevantes que incluyen la Inversión en TIC, Conexión y Uso de Internet, Comercio Electrónico, Herramientas TIC, Medios de Comunicación y Especialistas en TIC. Ver Figura 1.

**Figura 1**

*Conformación de base de datos*



Nota. Elaboración propia

## 7.2 LIMPIEZA, PREPROCESAMIENTO Y/O TRANSFORMACIÓN DE DATOS

Para el preprocesamiento y análisis de los datos, se lo realizará a través del programa Python, con su herramienta interactiva llamada Jupyter notebook, esta genera un entorno flexible para combinar texto explicativo, imágenes y código ejecutable en un único documento.

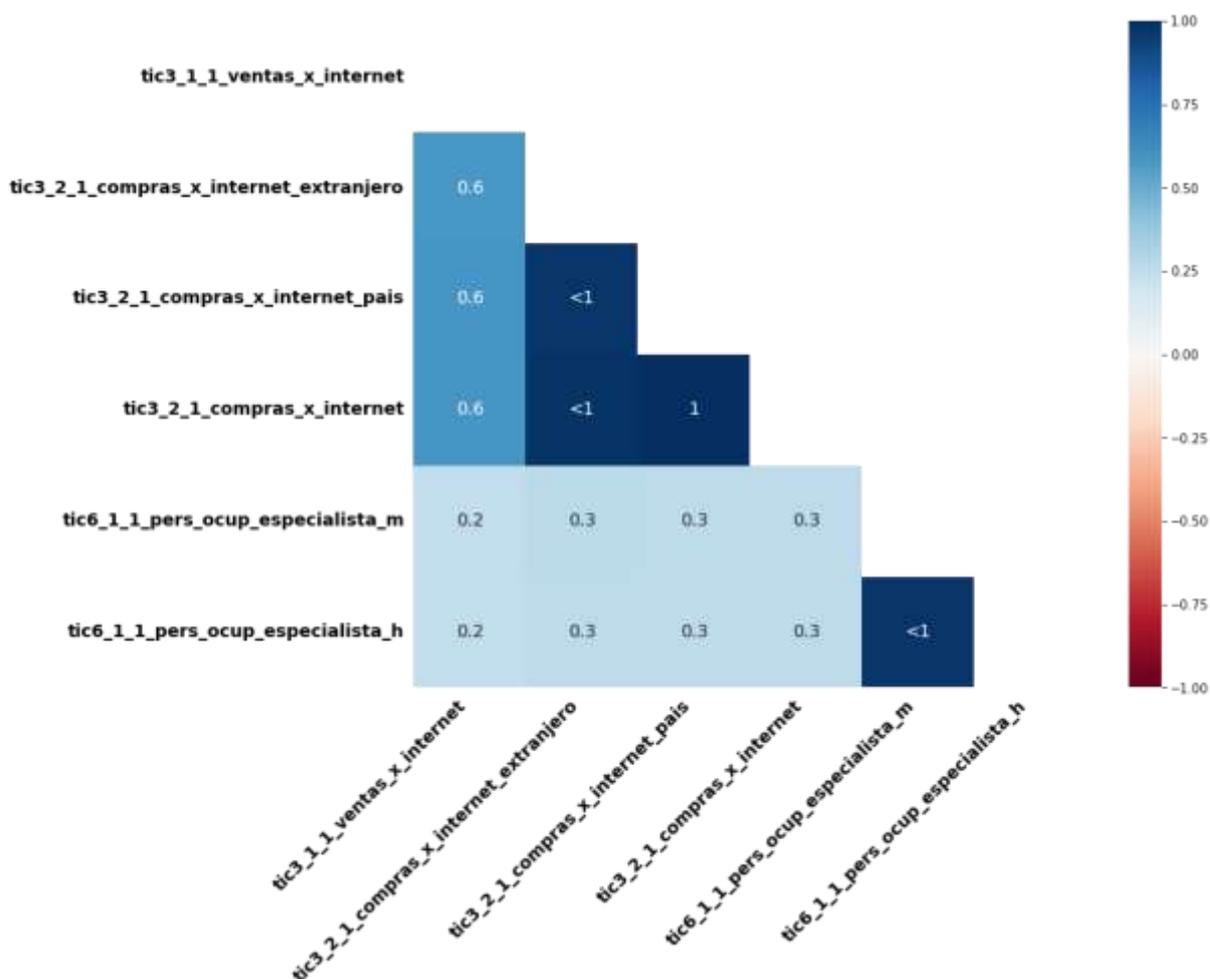
Una vez finalizado la recolección de los datos, es imprescindible preparar la matriz de datos de acuerdo con el algoritmo que se va a aplicar. En este caso, el estudio está basado en la aplicación de estrategias de transformación digital en las empresas que han invertido en TIC, por lo cual, se procederá a la eliminación de todos los registros de empresas que no han invertido en TIC desde la base original BDTIC.csv.

Para la exploración, limpieza y análisis de datos, se utilizarán algunas de las librerías, tales como: Pandas, la cual proporciona estructuras de datos tabulares eficientes, como DataFrames, NumPy, que ofrece potentes arreglos multidimensionales y funciones numéricas, Matplotlib, que permite crear visualizaciones estáticas y Seaborn, que facilita la creación de gráficos estadísticos atractivos. Estas herramientas son ampliamente utilizadas en el análisis de datos en Python (McKinsey, 2018).

Posteriormente se realizará un análisis del dataset, donde se resalta mediante un gráfico las variables que contiene la mayor cantidad de valores perdidos. Ver figura 2.

Figura 2

Representación de variables con valores perdidos



Nota. Elaboración propia

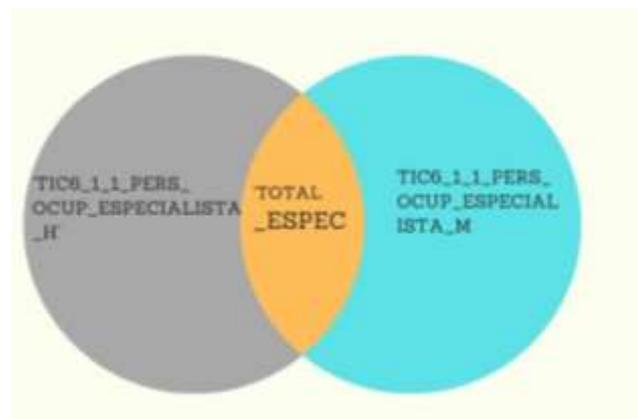
El gráfico anterior muestra las 6 variables con la mayor cantidad de valores perdidos, que rondan entre el 30% al 60%. Como regla general, se puede descartar variables si los datos faltantes corresponden a más del 60% de las observaciones, siempre teniendo en cuenta que la imputación es preferida sobre la eliminación de variables. Para el caso en particular las variables tanto de ventas como de compras numéricas serán separadas del dataset, ya que las mismas tienen un alto grado de valores perdidos los cuales sesgarían el análisis.

En lo referente a las demás variables numéricas que contiene información sobre el número de dispositivos y cantidad de personal, estas serán completadas con

valores de 0, ya que las mismas, de acuerdo con la encuesta corresponden a ausencia de estos. Adicional se unieron las columnas de personal y especialistas hombre/mujer para un mejor análisis, para posteriormente eliminar las columnas. Ver figura 3.

### Figura 3

*Unión y suma de variables personal y especialistas*



Nota. Elaboración propia

Para otro grupo de variables numéricas que representa el porcentaje de inversión, estas se las transformo a dólares, con el fin de estandarizar los datos que pertenecen al mismo grupo, y en particular a la variable Inversión en TIC (tic1\_2)

Posteriormente, se observa que las variables categóricas están directamente relacionadas con la variable conexión a internet (tic2\_1), por lo tanto, se la reemplazará con el valor de 2 que indica no aplicar al uso de la herramienta. De igual forma se eliminó algunas variables tales como Provincia y Año que para nuestro estudio no serán utilizadas.

Para el escalamiento de las variables categóricas se utilizó la librería LabelEncoder de sklearn.

Finalmente, se obtuvo la base de datos presentada a continuación:

#### **BD1: BDTICFINAL**

Está compuesta de 14029 filas × 58 columnas.

**Tabla 1***Diccionario de variables*

<b>Código</b>	<b>Descripción</b>	<b>Tipo</b>
des_tamano	Tamaño de la empresa	Categórica
tic_1	Inversión en Tic	Numérica
tic_2	Tipo de conexión a internet	Categórica
tic_2_6	Uso de Internet	Categórica
tic_2_8	Servicios de página web	Categórica
tic_3_1	Ventas por internet	Categórica
tic_3_1_2	Obstáculos de ventas por internet	Categórica
tic_3_2	Compras por internet	Categórica
tic_4_1	Tipos de hardware	Numérica
tic4_2	Tipos de software	Categórica
tic_4_4	Apoyo en uso de TIC	Categórica
tic_5	Medios de comunicación usado	Categórica
tic_6	Capacitación	Categórica
Total Personal	Número total de trabajadores	Numérica
Total_Espec	Número total de especialistas	Numérica

Nota. Elaboración propia

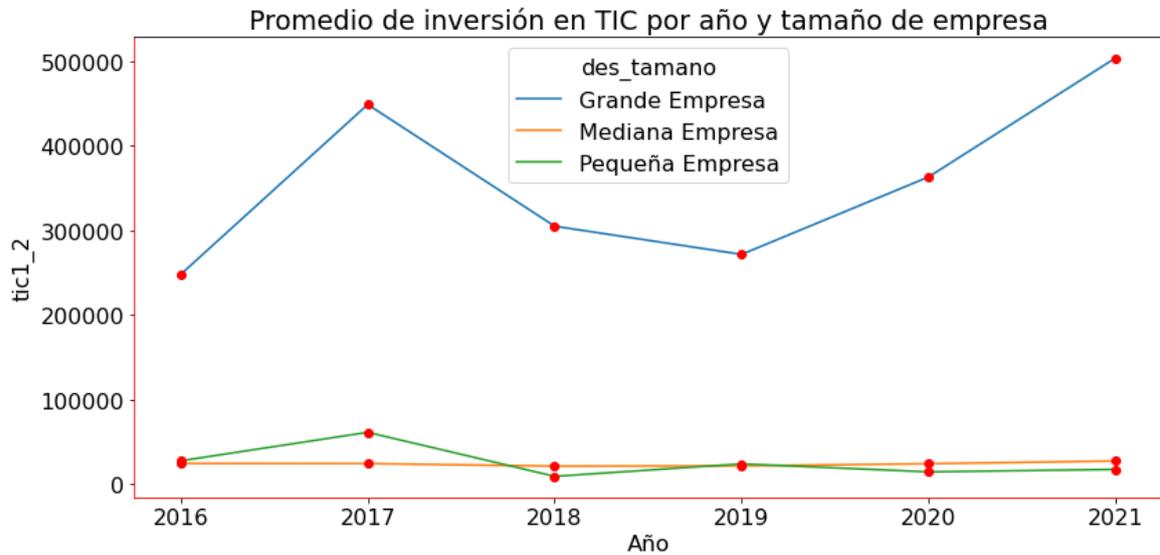
*Para una visualización más detallada ver anexo 1.*

### **7.3 IDENTIFICACIÓN Y DESCRIPCIÓN DE VARIABLES**

Para la identificación y descripción de las variables, estas se las elaborará basada en los grupos descritos en la sección anterior.

- **Análisis y descripción “Inversión TIC”**

En la figura 4 se observa una cronología de la inversión en TIC para los 3 tamaños de empresa. Donde se observa que, a largo de los 6 años, la inversión en TIC en términos económicos se ha mantenido en un rango determinado para los 3 tamaños de empresa.

**Figura 4***Promedio de inversión TIC*

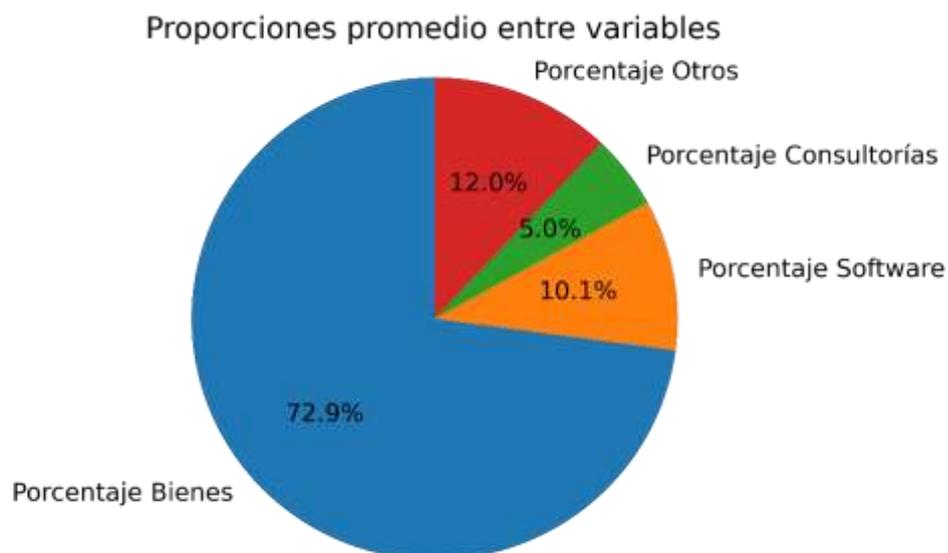
Nota. Elaboración propia

En términos generales, la grande empresa durante los 6 años observados aumento su porcentaje de inversión en TIC en un 103 %, la mediana empresa aumento un 11.94% y la pequeña empresa disminuyo la inversión en un 37% de acuerdo con la media indicada en el gráfico anterior.

Un parámetro necesario para la encuesta fue conocer en que se invierten la inversión proporcionada anteriormente. En el gráfico siguiente, se muestra dicho porcentaje y a las áreas a las que se destina:

**Figura 5**

*Porcentaje de inversión de acuerdo con las áreas*



Nota. Elaboración propia

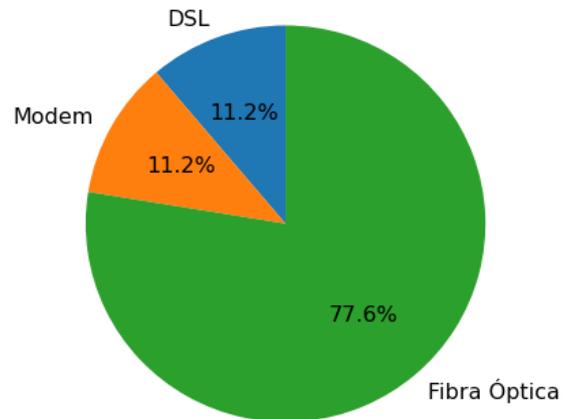
Este porcentaje se ha mantenido a lo largo de los años, donde siempre ha sido la constante la inversión en bienes para los tres tamaños de empresa.

- **Análisis y descripción “Conexión y Uso de Internet”**

El tipo de conexión a internet juega un papel predominante en el ámbito de la transformación e innovación digital, durante el último tiempo la conexión por fibra óptica se ha posicionado y mantenido dentro de la conexión que en su mayoría es preferida por las empresas a nivel nacional, abarcando el 77.6 % del total hasta 2021, marcando una tendencia positiva para este tipo de tecnología. En el siguiente gráfico se observa los 3 tipos de conexión que predominan en la actualidad.

**Figura 6***Porcentaje de tipo de conexión*

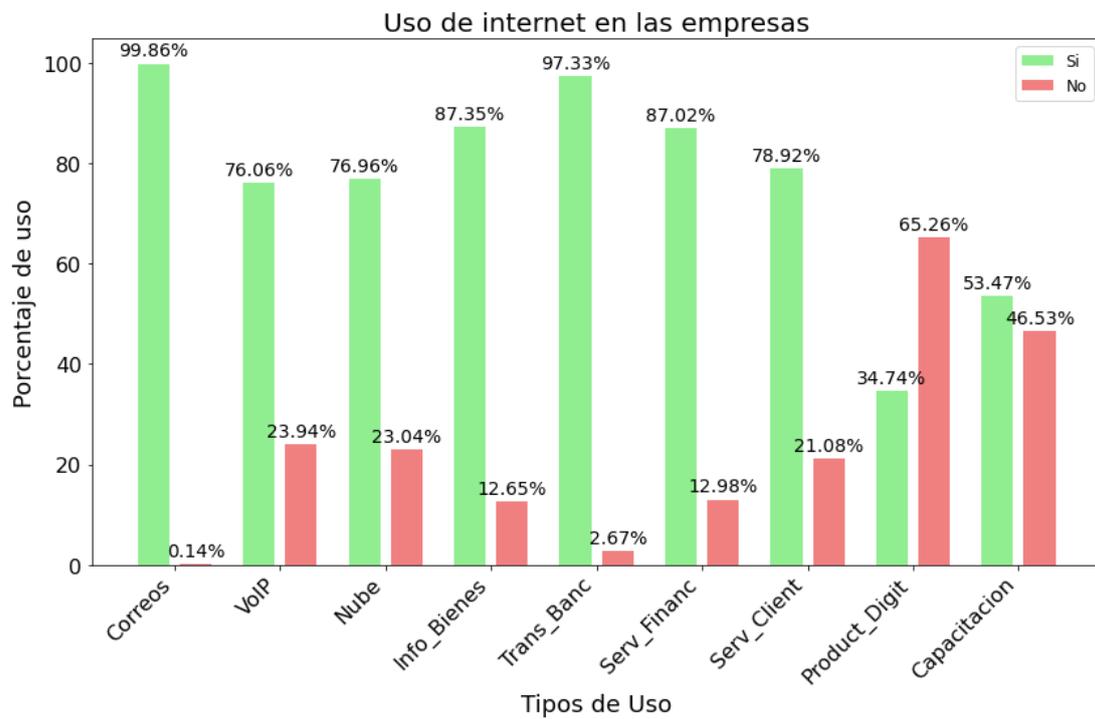
Proporción de empresas que usan internet por tipo de conexión



Nota. Elaboración propia

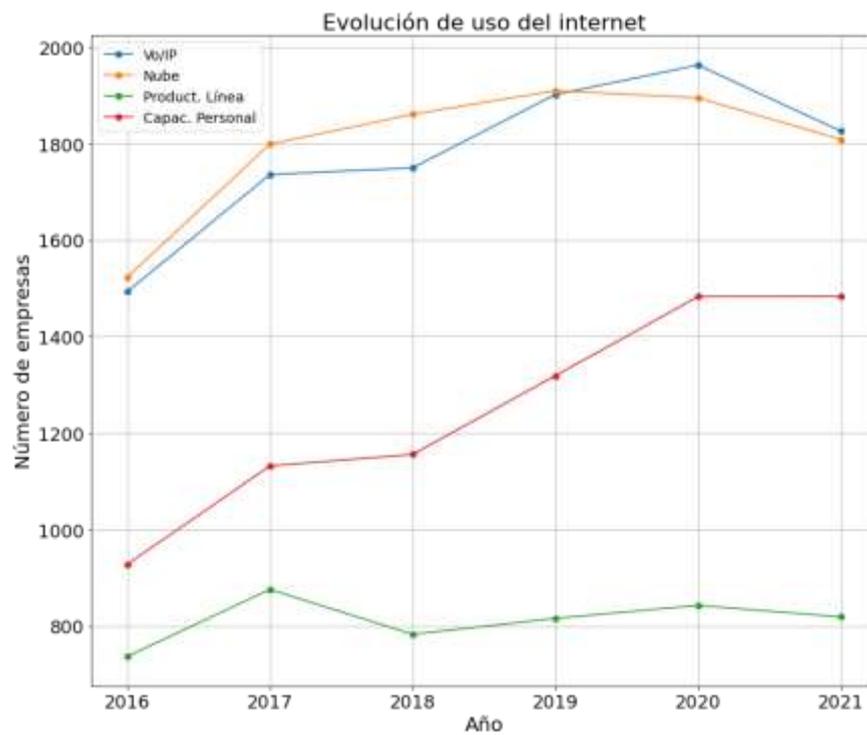
Como se observa, los 2 tipos de conexión restante son equivalentes en su uso, y esto se debe a que en su mayoría han sido reemplazadas por la rapidez y confiabilidad que ofrece la fibra óptica.

Un segmento importante dentro de este grupo es el uso del internet en las empresas y como este ha variado a lo largo de los años. Si se compara a nivel general, las variables de mayor uso son “correos, transacciones bancarias, información de bienes y servicios financieros”.

**Figura 7***Usos de internet*

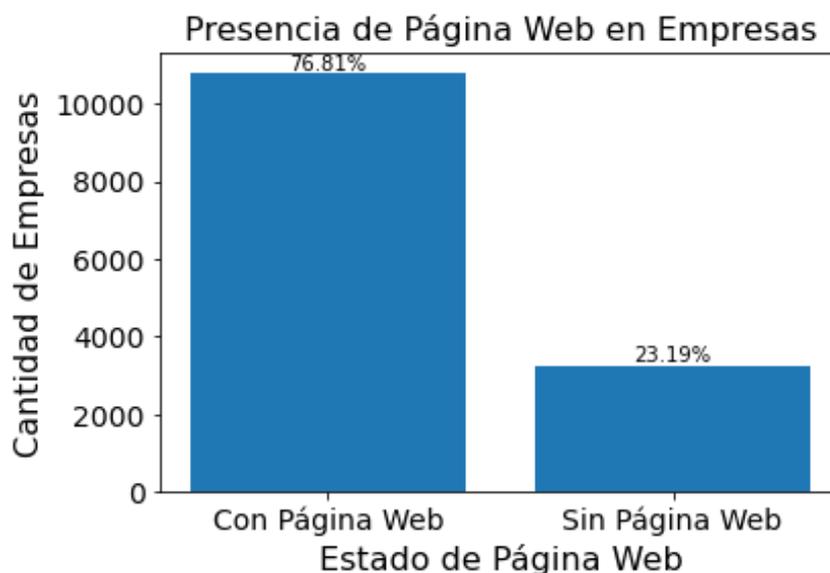
Nota. Elaboración propia

Si se lo analiza a través del tiempo, el comportamiento del uso de internet ha sido ascendente en los 6 años, resaltando las variables de capacitación con un 60%, entrega de productos en línea con 11%, nube con 18.8% y VoIP con 22.31%, por sobre los otros usos. Ver figura 8.

**Figura 8***Cronológico de Usos de internet*

Nota. Elaboración propia

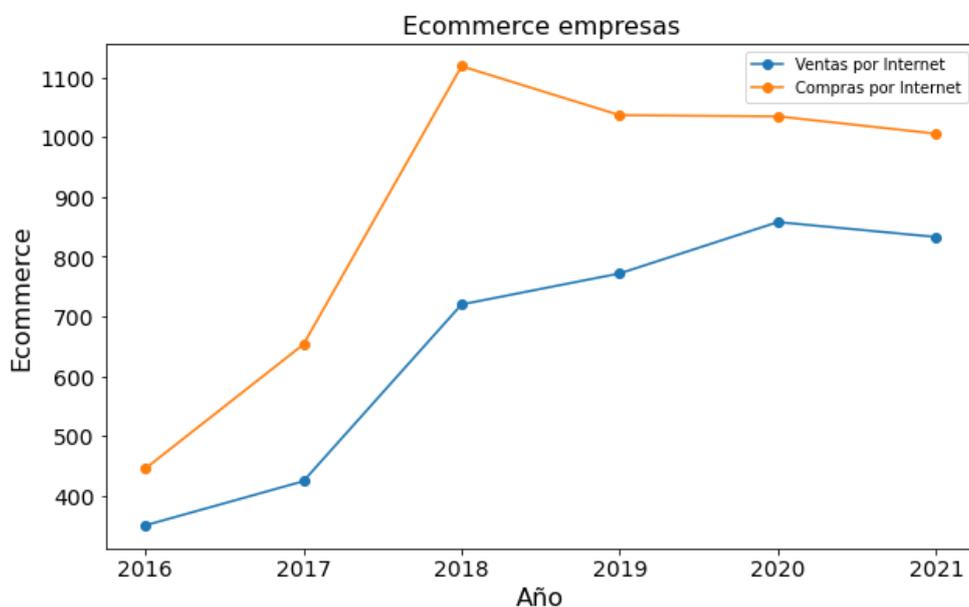
La siguiente variable dentro de este segmento, se trata del uso de páginas web en las empresas, la misma que se encuentra dentro del 70% para los tres tamaños de empresa, sin variaciones significantes a lo largo de los años, observándose un estancamiento para esta tecnología. A continuación, se presenta un gráfico del total de las empresas que utilizan esta herramienta innovadora dentro del mercado actual.

**Figura 9***Porcentaje de uso de páginas web*

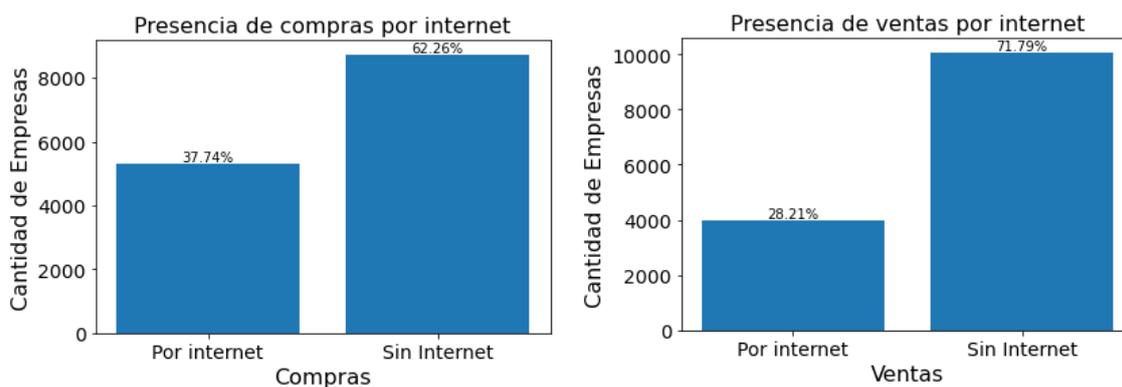
Nota. Elaboración propia

- **Análisis y descripción “Comercio electrónico”**

En el grupo 3 se va a realizar un análisis a nivel de ventas y compras por internet, las mismas que están representadas por variables categóricas de 1 y 2, Sí y No respectivamente. Como se observa en la figura 10, las compras en Ecuador son superiores a las ventas por internet de acuerdo con la encuesta realizada por el INEC, y esto se debe principalmente a factores sociales reflejados en el comportamiento y confianza del consumidor, o a la falta de plataformas digitales donde las empresas puedan ofertar su producto. Durante los 6 años la cantidad de empresas que han optado por el ecommerce se han duplicado, pasando de 445 en compras a 1006 y de 350 en ventas a 833.

**Figura 10***Cronológico de Ecommerce*

Nota. Elaboración propia

**Figura 11***Porcentaje de Ecommerce*

Nota. Elaboración propia

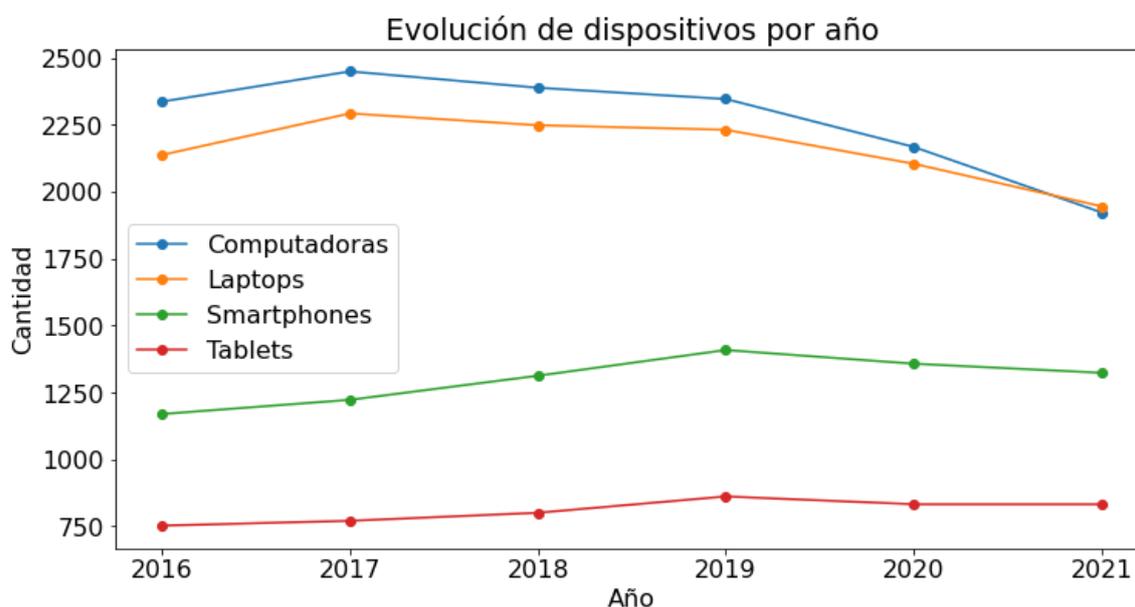
Mediante la figura 11, se corrobora en términos porcentuales que las ventas por internet son menores que las compras por internet a nivel empresarial. Como conclusión en este grupo se puede confirmar que la cultura de ecommerce en el país al 2021 es apenas de alrededor del 30 %.

- **Análisis y descripción “Herramientas TIC”**

Para el grupo 4 se analiza la infraestructura tanto de hardware como de software, en el gráfico 12 contiguo se observa como la tendencia hacia los equipos móviles está en alza en aproximadamente un 10% con respecto al 2016, mientras que las computadoras de escritorio y laptop han sufrido un pequeño descenso del 15 % en el uso a nivel empresarial.

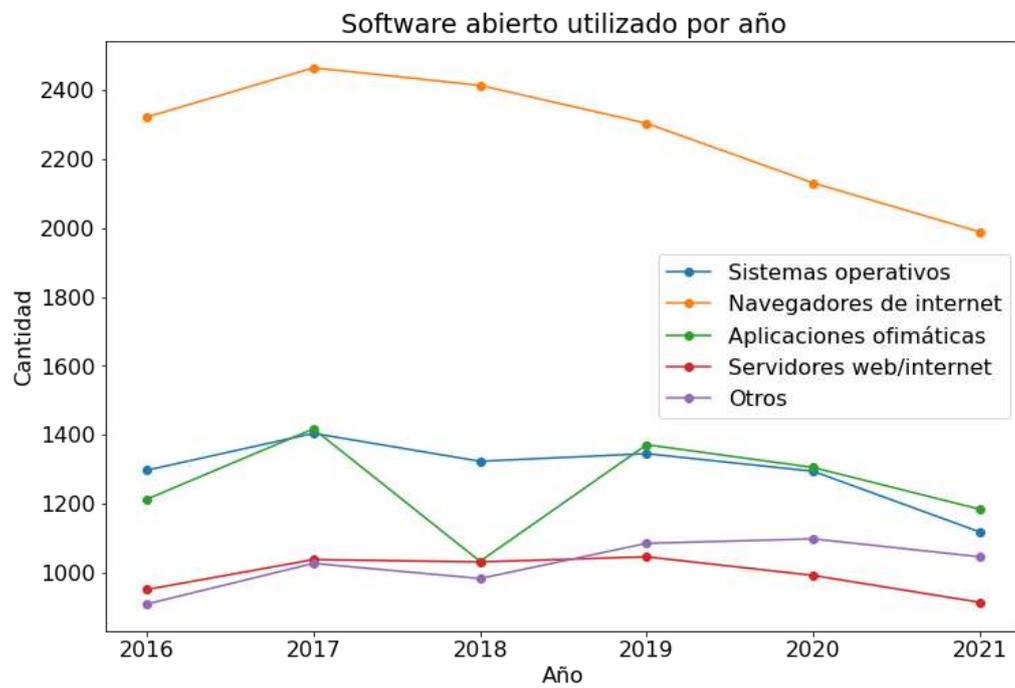
**Figura 12**

*Cronológico de uso de dispositivos*



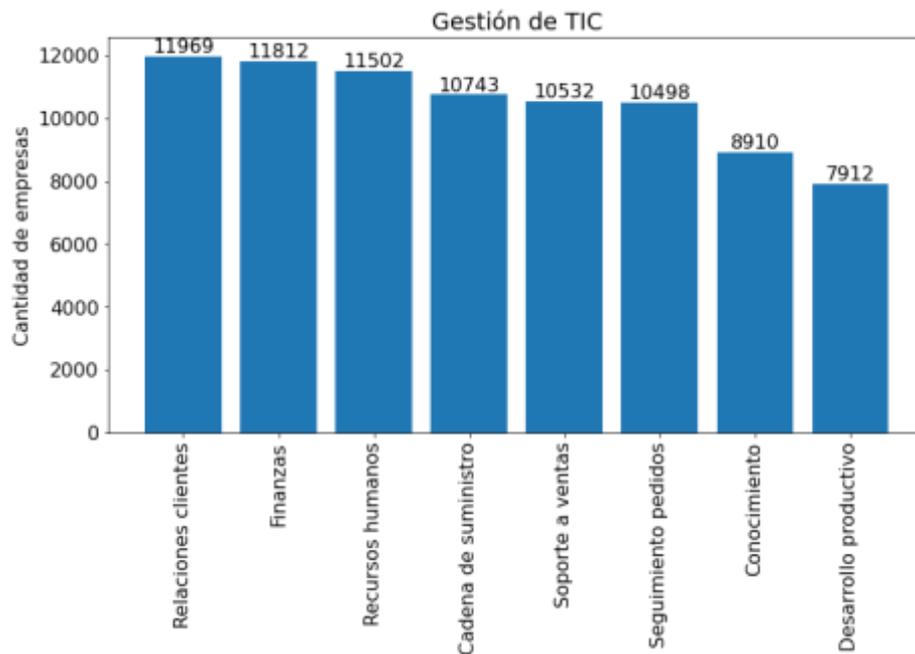
Nota. Elaboración propia

En lo que respecta al software, se observa en la figura 13 una tendencia positiva de la línea denominada “Otros”, esto indica que las empresas en la actualidad están invirtiendo en aproximadamente un 15 % más para aplicaciones tipo ERP o CRM, plataformas de aprendizaje, software de seguridad. Esto sin duda aporta al desarrollo e innovación tecnológica en las organizaciones.

**Figura 13***Cronológico de uso de software*

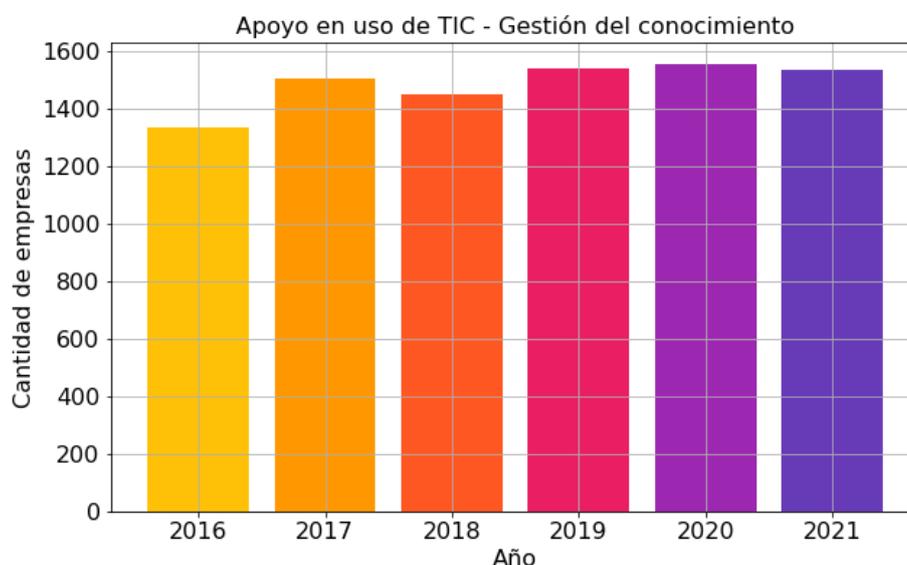
Nota. Elaboración propia

Para la gestión de las diferentes áreas en la empresa, se observa en la figura 14 que las empresas en su mayoría la usan para gestionar las relaciones con los clientes, finanzas y recursos humanos entre las 3 más representativas.

**Figura 14***Gestión de TICs*

Nota. Elaboración propia

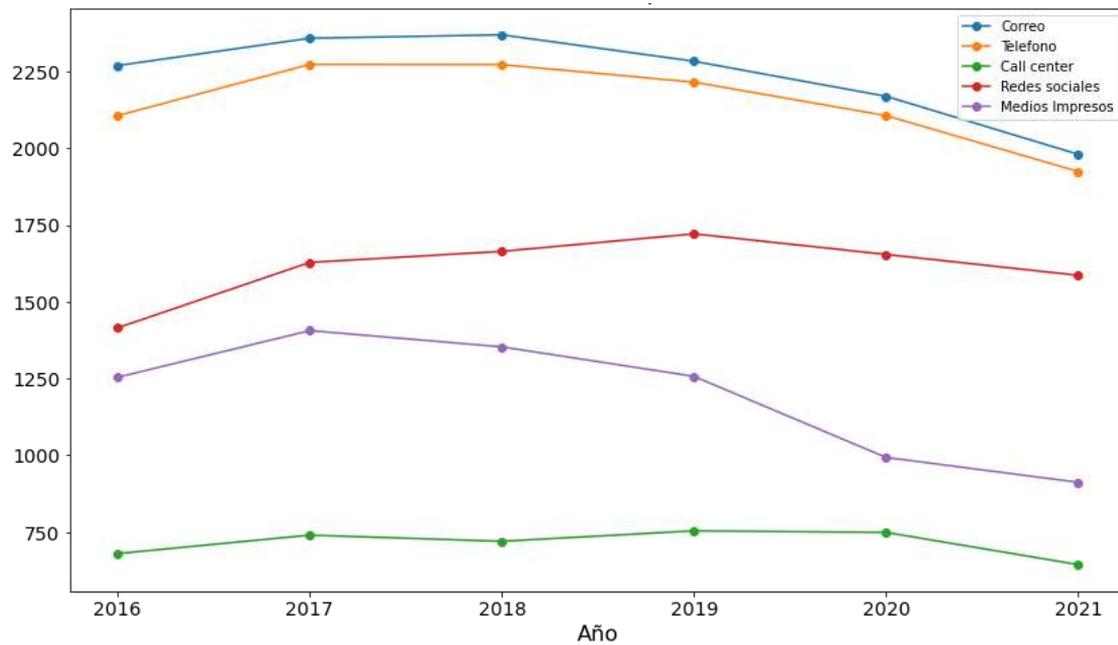
En este aspecto los diferentes tipos de gestiones se han mantenido a lo largo del tiempo, con una leve prioridad del 8% en la gestión del conocimiento, como se observa en la figura 15. De ahí que las demás variables se han mantenido en el orden de la imagen anterior.

**Figura 15***Cronológico de gestión del conocimiento*

Nota. Elaboración propia

- **Análisis y descripción “Medios de comunicación”**

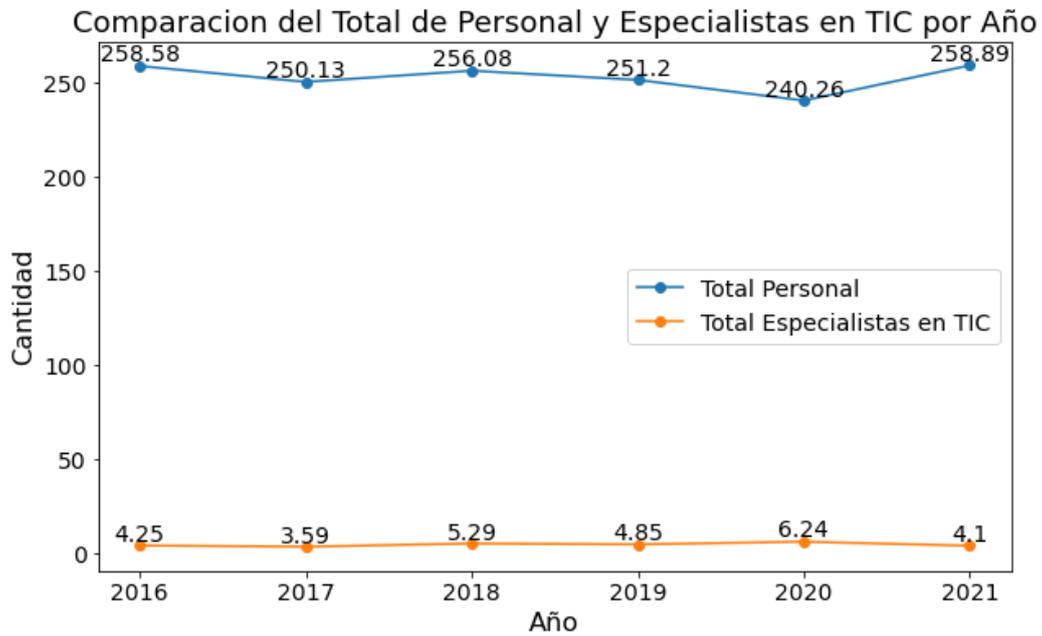
Para el segmento de los medios de comunicación, se observa en la figura 16 que las variables que resaltan son las redes sociales con un 12 % de tendencia positiva y los medios impresos con 27% como tendencia negativa, y esto es debido a que las redes sociales han ganado una gran popularidad y se han convertido en plataformas de comunicación masivas. Estas plataformas ofrecen a las empresas una serie de ventajas significativas, como una amplia audiencia global, la capacidad de interactuar directamente con los clientes, una mayor personalización y segmentación del contenido, la posibilidad de realizar campañas publicitarias altamente dirigidas, y un costo relativamente bajo en comparación con otros medios de comunicación tradicionales.

**Figura 16***Cronológico de uso de medios de comunicación*

Nota. Elaboración propia

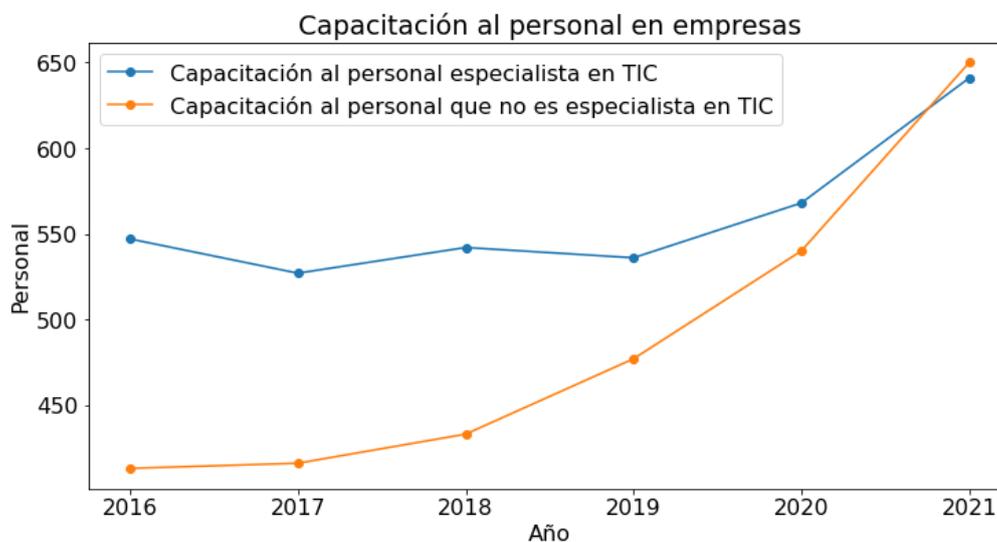
- **Análisis y descripción “Especialistas TIC”**

En el gráfico 17 se observa la media de personal con respecto a la media de especialistas por año, que como vemos en 2020 a raíz de la pandemia este aumento, sin embargo, al 2021 esta nuevamente regreso al promedio de 4.

**Figura 17***Cronológico de personal vs especialistas*

Nota. Elaboración propia

La capacitación en TIC se ha vuelto esencial para que los empleados adquieran estas habilidades y sean más adaptables en un entorno laboral cada vez más digitalizado, esto sin duda fomenta la innovación y la creatividad en las organizaciones. La falta de capacitación en TIC puede llevar a la obsolescencia y dificultar la capacidad de una organización para mantenerse competitiva en un entorno empresarial en constante cambio. Como se observa en la figura 18 esta ha ido creciendo a lo largo de los años.

**Figura 18***Cronológico de capacitación*

Nota. Elaboración propia

Algo particular en la gráfica anterior, es que existe un aumento en la capacitación a personal no especializado en TIC en aproximadamente un 57 % en 6 los años, mientras que la capacitación para personal especializado es de 17%, lo que se podría interpretar que muchas de las empresas buscan reducir costos al contratar personal especializado, lo que puede resultar más rentable a largo plazo.

#### **7.4 SELECCIÓN DEL MODELO ESTADÍSTICO**

En términos generales, el modelamiento de variables se refiere al proceso de desarrollar un modelo estadístico o matemático que capture la relación entre una variable dependiente y una o más variables independientes. El objetivo principal consiste en utilizar dicho modelo para predecir o comprender el comportamiento de la variable dependiente basándose en las variables independientes. De esta forma, el modelamiento de variables implica diversos enfoques y técnicas, los cuales varían en función de la naturaleza de los datos y los objetivos del análisis (Gerstner & Kugler, 2018). Algunos de los enfoques más comunes incluyen el modelado lineal, el modelado no lineal, el modelado basado en árboles de

decisión y el modelado basado en Clustering, los cuáles se ajustan a la investigación realizada, y sus objetivos específicos.

En este sentido, el clustering k-means es una técnica de aprendizaje no supervisado que agrupa los datos en diferentes categorías o clústeres según su similitud. Para esta investigación, el clustering k-means se utiliza para identificar patrones y segmentos de empresas con características similares en términos de su nivel de implementación tecnológica y de innovación. Esto permite comprender las diferencias y similitudes entre los grupos de empresas y obtener información valiosa sobre las barreras que enfrentan y las estrategias que han implementado para mejorar su competitividad.

Por otro lado, el random forest es un algoritmo de aprendizaje automático que combina múltiples árboles de decisión para realizar predicciones más precisas. En el contexto del objeto de estudio planteado, el random forest se utiliza para modelar y predecir variables complejas y relacionadas con la implementación tecnológica y de innovación. En este contexto, sirve para utilizar para predecir la probabilidad de que una empresa adopte tecnologías específicas o para evaluar la importancia relativa de diferentes factores en la implementación exitosa de prácticas innovadoras.

Finalmente, la regresión logística resulta adecuada para modelar y predecir variables categóricas, como la realización de actividades de comercio electrónico. Esta técnica permitiría identificar los factores que influyen en la probabilidad de que una empresa adopte estas tecnologías y prácticas innovadoras, y evaluar su importancia relativa en función de las características de las empresas. En este sentido, podemos analizar los modelos que serán usados para el análisis del trabajo propuesto:

### **1. Clustering. –**

El Clustering o segmentación se basa en la asignación de puntos de datos a grupos basados en la similitud de los puntos entre sí. En este sentido, el uso más común es mediante el algoritmo k-means, el cual toma un enfoque incremental, de manera que, pueda agregar de forma dinámica un centro de clúster de

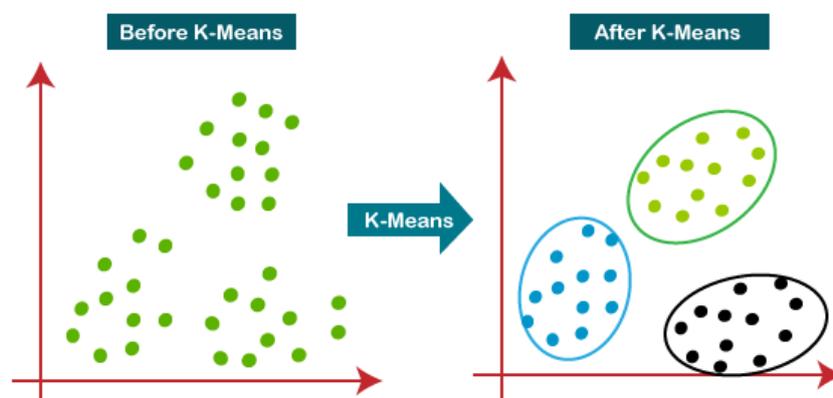
manera consecutiva, a propósito de llevar una búsqueda determinística dependiendo del tamaño de un conjunto de datos.

El algoritmo clásico de agrupamiento k-means se emplea para agrupar un conjunto de datos  $x^m$  ( $m = 1, \dots, M$ ) en  $k = 1, \dots, K$  grupos, considerando sus características mediante un proceso iterativo. Comienza con una suposición inicial de los centroides del clúster  $c^k$ , *para*  $k = 1, \dots, K$ , que generalmente se seleccionan al azar. Los centroides clasifican los datos asignándolos al centroide correspondiente si su distancia es mínima en relación con los centroides del grupo (Mafla, 2022). Así, los datos se agrupan basándose en la norma Euclidiana y los centroides se actualizan mediante el cálculo del promedio de los elementos del grupo al que pertenecen. El algoritmo se repite hasta que no haya cambios significativos en la asignación de los datos a los centroides. En resumen, k-means clasifica los datos en  $k$  grupos mediante la distancia entre los datos y los centroides, obteniendo una agrupación óptima donde la ubicación de los centroides refleja la posición promedio de los elementos del grupo respectivo.

El clustering de K-Means es el algoritmo de aprendizaje no supervisado más usado y conocido, donde se agrupa en función de patrones y a través de aprendizaje automático. Podemos representarlo gráficamente de la siguiente manera:

**Figura 19**

*Representación de Agrupación mediante K-Means*



Nota. Adaptado de K means clustering algorithm, por Prateek, 2022,  
<https://keytodatascience.com/k-means-clustering-algorithm/>

## 2. Regresión Logística. –

La regresión logística es una herramienta estadística utilizada para analizar datos con una variable dependiente dicotómica y un conjunto de variables predictoras. Se aplica tanto para explicar como para predecir. La variable dependiente tiene dos valores posibles (por ejemplo, para nuestro análisis, “sí” y “no”), mientras que las variables predictoras pueden ser tanto cuantitativas como categóricas. La idea fundamental es analizar por medio de una predicción, la probabilidad de un evento suceda, o no; de esta forma, seremos capaces de determinar cuáles son aquellas variables que tienen un mayor peso, en función de aumentar o disminuir la probabilidad de que suceda, o no, determinado propósito o evento (Chitarroni, 2022).

A partir de este modelo podemos trabajar con una variable dependiente categórica, particularmente con dos niveles; puesto que, a comparación de la regresión lineal, en este modelo encontramos una mejor función logística para describir una relación existente entre la variable dependiente e independiente. En este sentido, el modelo de regresión logística es capaz de predecir  $P(Y=1)$  en función de "X".

La función logística también conocida como Sigmoidea generalmente se denota por  $\sigma(x)$  o  $\text{sig}(x)$  y está dada por:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Considerando que la regresión lineal no siempre da una buena línea de ajuste en problemas que tiene únicamente dos niveles de la variable dependiente, por lo cual, existe menos precisión en el modelo, es necesario considerar un modelo diferente para la predicción a fin de cubrir nuestro conjunto de datos.

### 3. Random Forest. –

Los árboles de decisión son ampliamente utilizados como métodos de clasificación supervisada (Breiman et al., 1984). Se caracterizan por ser no paramétricos, robustos y de fácil interpretación. Estos árboles dividen sucesivamente el espacio de variables mediante particiones que buscan maximizar la homogeneidad de los grupos resultantes. Esto se logra al seleccionar la variable y el valor umbral que mejor separan las observaciones en cada partición. El proceso de partición se lleva a cabo hasta lograr la total homogeneidad en todas las particiones. El objetivo es eliminar las particiones más pequeñas que son influenciadas por el ruido en los datos de entrenamiento. Después de “podar el árbol”, cada partición en el espacio de variables se asigna a la clase más frecuente, lo que permite clasificar nuevos casos en función de su ubicación en dicho espacio. El algoritmo de Random Forest utiliza dos parámetros: la cantidad de árboles y el número de predictores utilizados en cada partición de cada árbol. A pesar de esto, una de las principales ventajas de este algoritmo es su baja sensibilidad a dichos parámetros, lo que significa que los valores predeterminados suelen generar resultados satisfactorios (Hastie et al., 2009).

En este sentido, cada árbol de decisión tiene una alta varianza, pero cuando los combinamos todos juntos en paralelo, la varianza resultante es baja, ya que cada árbol de decisión se entrena perfectamente en esos datos de muestra particulares, por lo que la salida no depende de un solo árbol de decisión, sino de varios árboles de decisión. En el caso de un problema de clasificación, la salida final se obtiene utilizando un clasificador de votación mayoritaria. En el caso de un problema de regresión, la salida final es el promedio de todas las salidas, por lo que, esta parte se llama *Agregación*.

En términos matemáticos, un Random Forest se puede representar como una función compuesta, donde cada árbol de decisión se considera una función individual y la salida final se obtiene mediante la combinación de las salidas de todos los árboles. Para el caso de clasificación se determina la siguiente ecuación:

$$\text{RandomForest}(X) = f_{1(X)} + f_{2(X)} + \dots + f_n(X)$$

Donde  $X$  es un vector de características de una nueva observación,  $f_i(X)$  es la predicción del  $i$ -ésimo árbol y  $\text{RandomForest}(X)$  es la predicción final del Random Forest.

En el caso de una regresión, podemos explicar la ecuación como:

$$\hat{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$$

En este caso, encontramos que  $\hat{Y}$  se trata de la predicción final, mientras que  $Y_i$  son las predicciones individuales de cada uno de los árboles y  $n$  es el número total de árboles en el Random Forest.

#### 4. Regresión Lineal Log-Lin.

Una de las ventajas de aplicar logaritmos sobre  $Y$ , es su habilidad para acotar el rango de la variable en una cantidad más pequeña que la original. Este efecto generalmente reduce la sensibilidad de las estimaciones a las observaciones extremas o atípicas, tanto de las variables independientes como las dependientes.

$$\log(Y) = \beta_0 + \beta_1 X + e$$

El modelo Log-Nivel se conoce también como la semielasticidad de "Y" respecto a "X", y se interpreta como un incremento de 1 unidad en "X", se asocia a un cambio en "Y" de  $(100 \cdot \beta_1)\%$ , ( 0.1%)

## 8. RESULTADOS

### 8.1 ANÁLISIS DE MODELO ESTADÍSTICO E INTERPRETACIÓN DE RESULTADOS

#### 8.1.1 MODELOS DE REGRESIÓN

En la siguiente tabla se encuentran representados los principales resultados obtenidos con los modelos propuestos:

**Tabla 2**

*Resultados modelos de aprendizaje supervisado*

Modelos	Error medio absoluto (MAE)	Coefficiente de determinación (r2_score)
Random Forest	0.132	0.89
Regresión Logística	0.135	0.68
Regresión Lineal	0.25	0.36

Nota. Elaboración propia

#### 8.1.2 ANÁLISIS MEDIANTE REGRESION LINEAL LOG LIN

Para el análisis del modelo se va a escoger todas las variables numéricas y se las va a agrupar como `numeric_data`. Este segmento se conforma de los siguientes grupos:

**Tabla 3**

*Grupos de variables numéricas*

<b>numeric_data</b>
<ul style="list-style-type: none"> <li>• Grupo 1 (Inversion Tic)</li> <li>• Grupo 4 (Número de dispositivos)</li> <li>• Grupo 6 (Total personal y especialistas)</li> </ul>

Nota. Elaboración propia

Este modelo alcanza a explicar en un 36 % la variabilidad de nuestra variable dependiente Inversión en TIC (tic1\_2), por medio de la inclusión de 10 variables. De igual forma, todas las variables con excepción de Total\_Espec pasan el 95 % de confianza. A continuación, se presentan los resultados de los coeficientes.

**Tabla 4**

*Coeficientes regresión lineal*

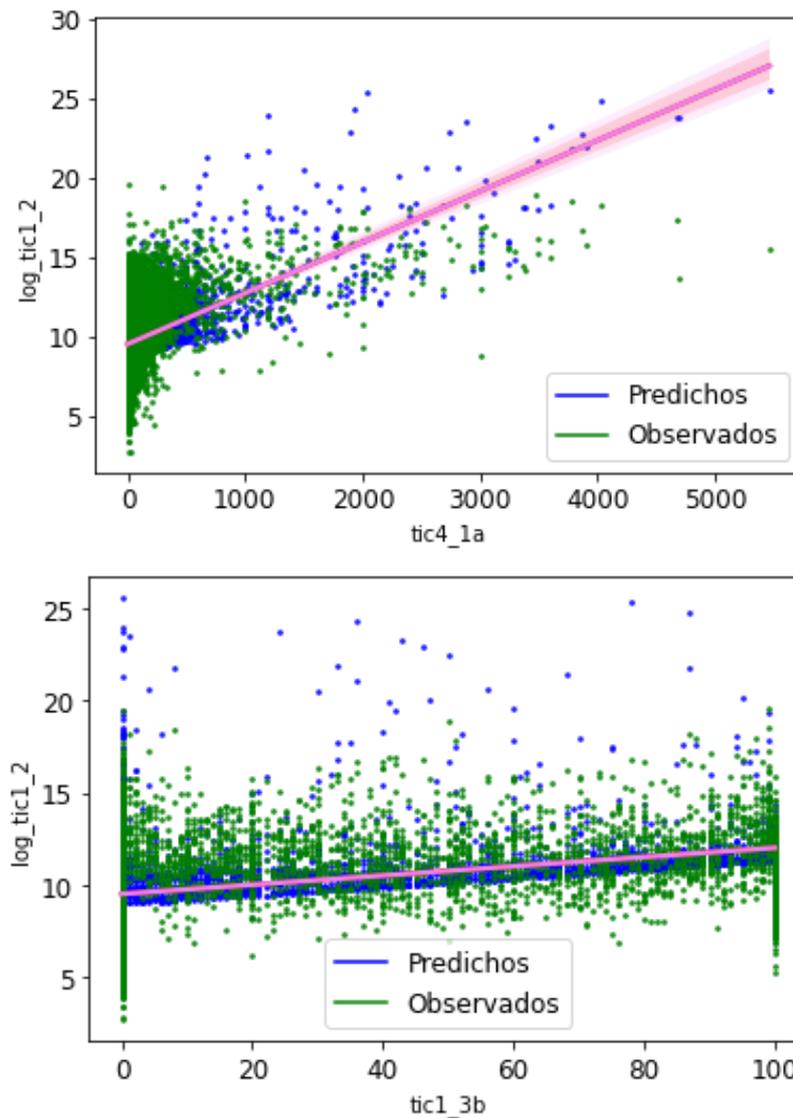
<b>Variable independiente</b>	<b>Coeficiente</b>	<b>Valor p</b>
<b>tic1_3a: Inversión - Bienes TIC</b>	0.0887	0.000
<b>tic1_3b: Inversión - Software</b>	0.1127	0.000
<b>tic1_3c: Inversión - Consultorías en TIC</b>	0.1020	0.000
<b>tic1_3d: Inversión – Otros (Telecomunicaciones, etc)</b>	0.1034	0.000
<b>tic4_1a: Número dispositivos - Computadoras de escritorio</b>	0.0010	0.000
<b>tic4_1b: Número dispositivos - Laptop o notebook</b>	0.0028	0.000
<b>tic4_1d: Número dispositivos - Smartphone</b>	0.0022	0.000
<b>tic4_1e: Número dispositivos - Tablet</b>	0.0036	0.000
<b>Total_Personal</b>	0.0005	0.000
<b>Total_Espec: Total especialistas</b>	-0.0005	0.125

Nota. Elaboración propia

Una vez obtenido los coeficientes se procederá a graficar los valores predichos con los observados. A continuación, se observa la imagen de la inversión en TIC sobre el número de computadores de escritorio y la inversión en software representada con porcentajes.

**Figura 20**

*Modelo Log – Lin Valores observados vs Valores predichos*



Nota. Elaboración propia

Para ejemplificar cómo sería la interpretación, se tomará como ejemplo 6 variables:

- Tic1\_3a: El incremento de 1 punto en la inversión de bienes implica un aumento en "Y" de 8.89 %.
- Tic1\_3b: El incremento de 1 punto en la inversión de software implica un aumento en "Y" de 11.27 %.

- Tic1\_3c: El incremento de 1 punto en la inversión de consultoría implica un aumento en "Y" de 10.2 %.
- Tic1\_3d: El incremento de 1 punto en la inversión de otros (telecomunicaciones y alquiler de equipos, etc.) implica un aumento en "Y" de 10.34 %.
- tic4\_1a: El incremento de 1 computador de escritorio, se asocia a una variación positiva de 0,1% en "Y".
- tic4\_1d: El incremento de 1 Tablet, se asocia a una variación positiva de 0,3 % en "Y".

Estos valores son determinantes al momento de tomar decisiones en lo referente a la innovación y la transformación digital. De los cuatro parámetros que tiene la encuesta sobre el porcentaje de inversión, se concluye que uno de los factores que en la actualidad generan mayor costo es el software, seguido de otros servicios, consultorías y bienes. De igual forma en lo que respecta a hardware contenida en el grupo 4, los computadores de escritorio generan un menor gasto para la inversión en tic, y por el contrario la Tablet es el dispositivo que mayor inversión requiere.

Con estos resultados, se observa que el modelo tiene un buen ajuste y las variables incluidas son significativas para predecir la inversión en TIC.

### 8.1.3 ANÁLISIS MEDIANTE EL ALGORITMO K-MEANS

Para la ejecución del modelo clúster por kmeans, se ha escogido las variables de porcentaje de inversión con respecto al total de Inversión en TIC, representada en la tabla siguiente:

**Tabla 5**

*Porcentajes de Inversión TIC*

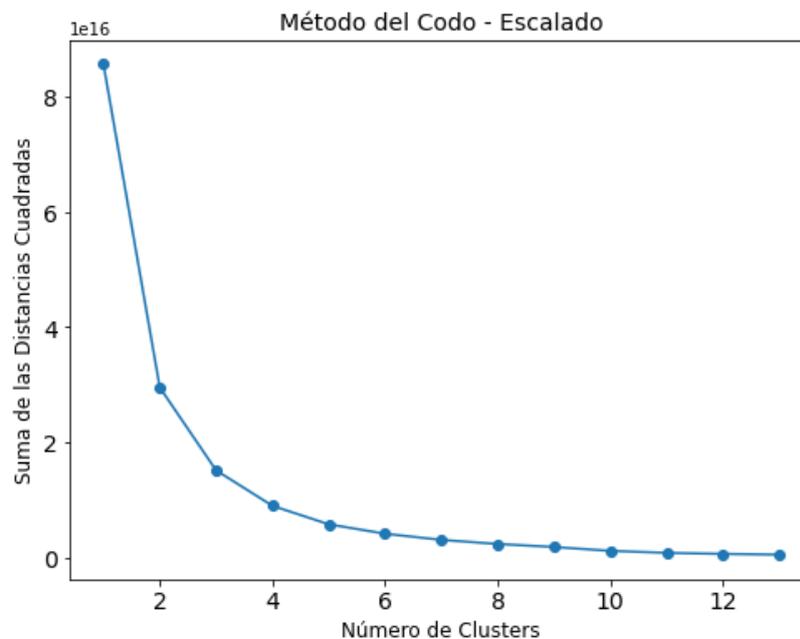
<b>tic1_3a: Inversión - Bienes TIC</b>
<b>tic1_3b: Inversión - Software</b>
<b>tic1_3c: Inversión - Consultorías en TIC</b>
<b>tic1_3d: Inversión – Otros (Telecomunicaciones, etc.)</b>

Nota. Elaboración propia

En primer lugar, se procederá al escalado los datos con el algoritmo StandardScaler, para posterior aplicar la técnica del codo comúnmente utilizada para determinar el número óptimo de clústeres.

**Figura 21**

*Método del codo - Escalado*



Nota. Elaboración propia

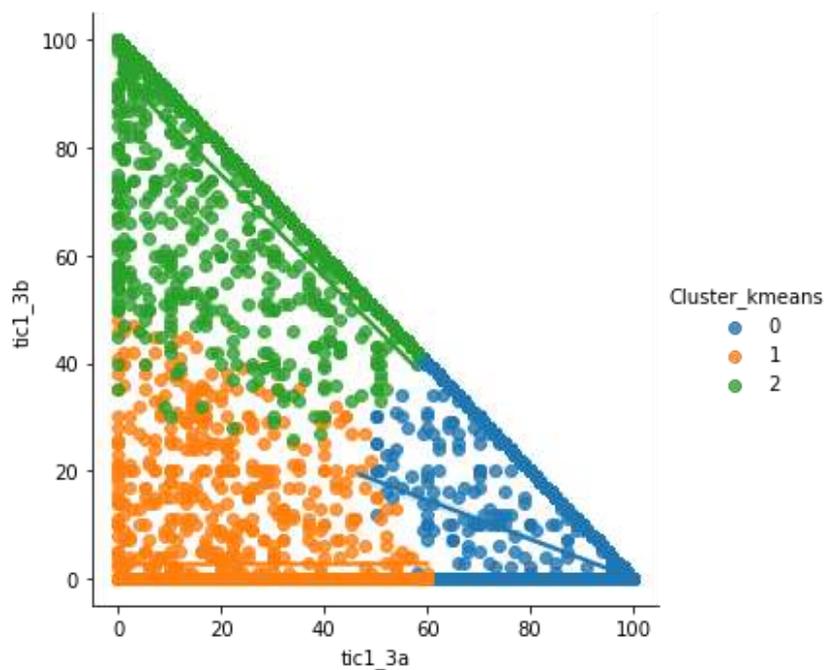
De acuerdo con el método de codo, el número de clúster podría colocarse entre 2 o 3, esto se verificará una vez modelado el segmento en cuestión.

- **Visualización y análisis mediante gráficos de dispersión**

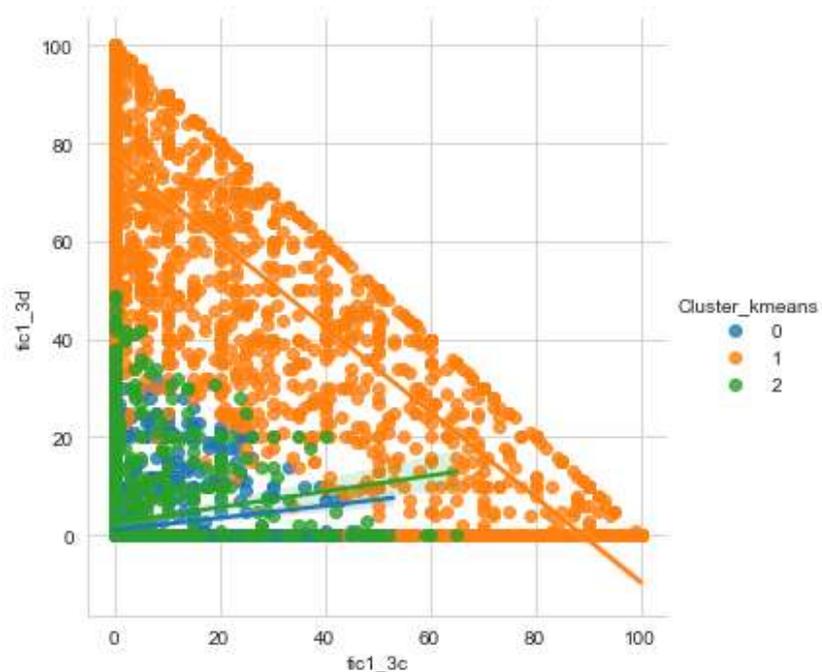
Al aplicar  $n\_cluster=3$ , se observa una dispersión uniforme en los datos y por ende una mejor formación de los grupos. En el caso particular se observa la relación entre el porcentaje de inversión en bienes (tic1\_3a) y el porcentaje de inversión en Software (tic1\_3b) y el porcentaje de inversión en consultorías (tic1\_3c) y el porcentaje de inversión en Otros (tic1\_3d).

### Figura 22

*Agrupación por 3 clústeres*



Nota. Elaboración propia

**Figura 23***Agrupación por 3 clústeres*

Nota. Elaboración propia

Como se observa en las imágenes anteriores, el gráfico de dispersión ofrece una mejor visualización de los grupos, por lo tanto, se puede concluir que el número de clúster adecuado es 3. Posteriormente, se procederá a realizar un análisis general de los resultados, mediante la siguiente tabla:

**Tabla 6***Análisis con 3 clústeres con la media.*

<b>Clúster</b>	<b>tic1_3a</b>	<b>tic1_3b</b>	<b>tic1_3c</b>	<b>tic1_3d</b>
<b>0</b>	96.8	1.2	0.8	1.1
<b>1</b>	17.1	2.6	21.9	58.3
<b>2</b>	17.6	77.2	2.3	2.9

Nota. Elaboración propia

En el clúster 0, la inversión en bienes bordea el 96% y la inversión en las otras variables es muy bajo promediando el 1%, lo que indica un nivel de influencia reducido.

En el clúster 1, la inversión en servicios de telecomunicaciones, alquiler de equipos etc., es mayor al 50 % del total de la inversión en TIC, seguido de bienes y consultorías, en este clúster la inversión en software es mínima.

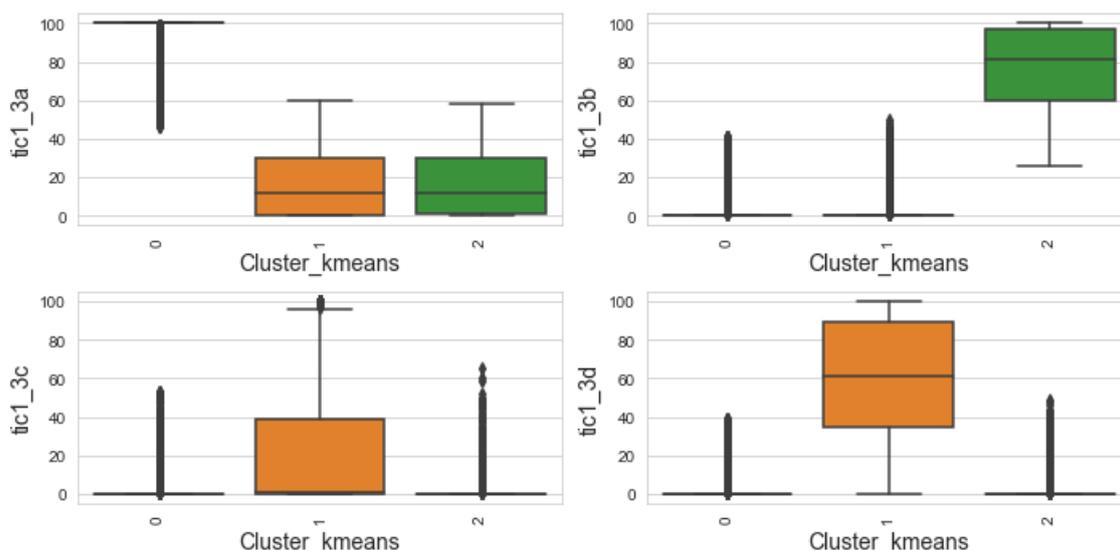
En el clúster 2, la inversión en software lidera el porcentaje de inversión, seguido de bienes en menor cantidad, y casi nula en software y otros.

- **Visualización y análisis mediante diagramas de cajas**

Con el fin de analizar de manera cuantitativa la distribución de los datos en diferentes grupos, se presentan los gráficos a continuación para el posterior análisis.

**Figura 24**

*Visualización de clúster vs porcentajes de inversión*



Para la variable tic1\_3a o Inversión en bienes el mínimo de inversión se da en el clúster 0, el cual es del 47%, posteriormente en todos los cuartiles se encuentra en el 100%. Para los 2 clústeres restantes el porcentaje máximo de inversión es 60%. Para el objeto de estudio significa que existe un grupo de empresas que invierte la mayor parte del presupuesto en TIC para la adquisición de bienes.

Para la variable tic1\_3b o Inversión en software el mínimo de inversión se da en el clúster 2, el cual es del 26 %, posteriormente en todos los cuartiles existe un aumento similar hasta llegar al 100 %. Para los 2 clústeres restantes el

porcentaje máximo de inversión es 50 %. Esto traducido al contexto del estudio significa que existe un grupo de empresas que invierte la mayor parte del presupuesto en TIC para la compra de aplicaciones o plataformas digitales. Mientras que los otros 2 grupos de empresas designan un máximo de 50%.

Para la variable tic1\_3c o Inversión en consultorías el mínimo de inversión se da en los tres clústeres, el cual es del 0 %, posteriormente solo el clúster 1 registra un valor de 38% en el tercer cuartil, en este mismo clúster existen empresas que incluso gastan el 100 % del presupuesto para TIC en las consultorías. Para los 2 clústeres restantes el porcentaje máximo de inversión está entre el 53% y 65%.

Para la variable tic1\_3d o Inversión en otros servicios, el mínimo de inversión se da en los tres clústeres, el cual es del 0 %, posteriormente solo el clúster 1 registra valores en todos los cuartiles, llegando incluso en algunos casos al 100 %. Para los 2 clústeres restantes el porcentaje máximo de inversión es esta entre el 39 % y 49 %.

Las prioridades de inversión siempre irán variando, dependiendo de los recursos y la visión de las empresas, las cuales son las principales causas al momento de adquirir A o B producto. De acuerdo con los resultados, las empresas que invierten más en bienes se encuentran en el clúster 0, que a la par son las empresas que invierten en menor cantidad en TIC (Mediana y Pequeña empresa), ya sea por falta de recursos o ausencia de conocimientos en las tecnologías actuales. Mientras que en el clúster 2 se encuentran las empresas con mayor poder adquisitivo, las cuales no escatiman recursos al momento de innovar sus diferentes servicios a través de las diferentes plataformas digitales. En el tema de consultorías, estas han ganado terreno en los últimos años y en la actualidad se invierte una cantidad significativa en este campo.

Sin embargo, el panorama está cambiando y cada vez más medianas y pequeñas empresas empieza a invertir en soluciones de software. A medida que la tecnología evoluciona y se vuelve más accesible, por lo que es probable que veamos un aumento en la adopción de soluciones de software por parte de este tipo de empresas en el futuro.

### 8.1.4 ANÁLISIS MEDIANTE REGRESIÓN LOGÍSTICA

Para la aplicación de este modelo escogimos 2 variables referentes a nuestro tercer modelo del segmento “comercio electrónico”, para lo cual, pudimos analizar 2 modelos de regresión logística distintos donde nuestra variable dependiente es tic3\_1 (Ventas por internet) en el primer caso, y tic 3\_2 (Compras por internet) para nuestro segundo modelo.

#### 1. Variable dependiente: Venta de bienes o servicios por internet (tic3\_1)

**Tabla 7**

*Coefficientes regresión ventas por internet*

<b>Variable independiente</b>	<b>Coefficiente</b>	<b>Valor p</b>
<b>Uso Internet – Brindar servicio al cliente (tic2_6h)</b>	0.2961	0.000
<b>Uso de internet - Entregar productos en línea (tic2_6i)</b>	0.259	0.000
<b>Disponibilidad de sitio o página web (tic2_8)</b>	0.2887	0.000
<b>Página Web - Acceso a catálogos de productos (tic2_8_2b)</b>	0.4569	0.000
<b>Página web – Recepción de pedidos (tic2_8_2c)</b>	0.969	0.000
<b>Obstáculos - Logística (tic3_1_2b)</b>	-0.3076	0.000
<b>Obstáculos - Seguridad informática y protección (tic3_1_2d)</b>	-0.3702	0.000
<b>Obstáculos - Costo de venta mayor a beneficio (tic3_1_2f)</b>	-0.2959	0.000
<b>Apoyo de TIC - Relaciones con los clientes (tic4_4a)</b>	0.4263	0.000
<b>Apoyo de TIC - Control y seguimiento de pedidos (tic4_4b)</b>	0.2947	0.000
<b>Apoyo de TIC - Gestión de recursos humanos (tic4_4e)</b>	-0.2733	0.000
<b>Apoyo de TIC - Servicio y soporte a ventas (tic4_4f)</b>	0.766	0.000
<b>Personal especialista en TIC (tic6_1)</b>	-0.1965	0.000

Nota. Elaboración propia

Estas variables representan diferentes características empresariales relacionadas con las ventas por internet; por lo cual, escogemos las más representativas en función de sus coeficientes y su p value, por lo cual, es importante identificar cuáles afectan de manera positiva y negativa a la variable dependiente.

Las variables independientes tic2\_6h, tic2\_6i, tic2\_8, tic2\_8\_2b, tic2\_8\_2c, tic4\_4a, tic4\_4b y tic4\_4f tienen un impacto positivo significativo en la variable dependiente tic3\_1, que se refiere a las ventas de bienes o servicios por internet. En este sentido, el uso de internet para brindar servicio al cliente (tic2\_6h) contribuye positivamente a las ventas por internet al permitir una interacción más cercana y efectiva con los clientes, lo que puede generar confianza y fidelidad. Asimismo, la entrega de productos digitales en línea (tic2\_6i) agiliza y facilita el proceso de compra, brindando una experiencia conveniente y satisfactoria para los consumidores. De igual manera, la disponibilidad de sitio o páginas web (tic2\_8) es fundamental para las ventas por internet, ya que un acceso constante y fluido a la plataforma en línea asegura que los clientes puedan explorar y adquirir productos o servicios sin contratiempos.

Los servicios de página web que ofrecen acceso a catálogos de productos (tic2\_8\_2b) y recepción de pedidos o reservas (tic2\_8\_2c) impulsan las ventas al proporcionar a los clientes información detallada sobre los productos, así como la posibilidad de realizar compras de manera rápida y segura. Al igual que, el apoyo en el uso de TIC para la gestión de relaciones con los clientes (tic4\_4a), el control y seguimiento de pedidos (tic4\_4b) y el servicio y soporte a ventas (tic4\_4f) fortalecen la interacción y la comunicación con los clientes, permitiendo una atención personalizada, eficiente y satisfactoria. Estas acciones contribuyen a construir relaciones sólidas con los clientes, lo que a su vez puede aumentar las ventas por internet.

Por el contrario, tenemos el caso de que las variables independientes tic3\_1\_2b, tic3\_1\_2d, tic3\_1\_2f, tic4\_4e y tic6\_1 tienen un efecto negativo significativo en la variable dependiente tic3\_1, que representa las ventas de bienes o servicios por internet. En este sentido, la presencia de obstáculos logísticos (tic3\_1\_2b) puede dificultar el proceso de entrega y distribución de los productos vendidos por internet, lo que puede generar retrasos y afectar negativamente la experiencia del cliente, disminuyendo así las ventas en línea. La preocupación por la seguridad informática y la protección de datos (tic3\_1\_2d) puede generar desconfianza en los clientes y reducir su disposición a realizar transacciones en línea, lo que puede limitar las ventas. Además, si el costo de venta por internet

es mayor que los beneficios obtenidos (tic3\_1\_2f), esto puede afectar negativamente la rentabilidad de las ventas en línea y desincentivar su desarrollo.

La gestión de recursos humanos en el ámbito de las TIC (tic4\_4e) es crucial para garantizar el buen funcionamiento de las operaciones en línea, ya que la falta de personal capacitado o recursos insuficientes puede afectar negativamente la calidad y eficiencia de los servicios en línea, lo que puede disminuir las ventas. Asimismo, la disponibilidad de personal especializado en TIC dentro del talento humano empresarial (tic6\_1) es esencial para implementar y mantener adecuadamente las estrategias de ventas en línea. La falta de personal TIC puede limitar las capacidades tecnológicas de la empresa y afectar la capacidad para impulsar las ventas en internet.

- **Accuracy**

Para la evaluación de nuestro modelo nuestro accuracy es de: 0.79; esto nos indica que, de las 14.029 observaciones, el modelo predijo correctamente el 79% de veces.

- **Matriz de confusión**

La matriz de confusión se utiliza para evaluar el rendimiento de un modelo de clasificación en base a sus predicciones. En este caso, la clasificación tiene dos clases: Positivo y Negativo.

**Tabla 8**

*Matriz de confusión ventas por internet*

		Predichos	
		Negativo (N) -	Positivo (P) +
Actuales	Negativo -	2052	1905
	Positivo +	1028	9044

Nota. Elaboración propia

- Verdaderos positivos (TP): 9044. Estos son los casos en los que el modelo predijo correctamente la clase positiva.
- Falsos positivos (FP): 1905. Estos son los casos en los que el modelo predijo incorrectamente la clase positiva cuando la clase verdadera era negativa.
- Falsos negativos (FN): 1028. Estos son los casos en los que el modelo predijo incorrectamente la clase negativa cuando la clase verdadera era positiva.
- Verdaderos negativos (TN): 2052. Estos son los casos en los que el modelo predijo correctamente la clase negativa.

En la matriz de confusión, podemos reconocer que el modelo tiene un alto número de verdaderos positivos (TP = 9044) y verdaderos negativos (TN = 2052), lo cual indica que está acertando en la clasificación de ambas clases. Sin embargo, también tiene un número significativo de falsos positivos (FP = 1905) y falsos negativos (FN = 1028), lo que indica que comete errores al predecir algunas muestras.

## 2. Variable dependiente: Compra de bienes o servicios por internet (tic3\_2)

**Tabla 9**

*Coeficientes compras por internet*

<b>Variable independiente</b>	<b>Coeficiente</b>	<b>Valor p</b>
<b>Banda Ancha Fija - Conexión DSL (tic2_2_1a)</b>	-0.1676	0.006
<b>Uso de internet - Utilizar servicios en la nube (tic2_6c)</b>	0.2175	0.000
<b>Uso de internet - Obtener información bienes (tic2_6d)</b>	0.4013	0.000
<b>Disponibilidad de sitio o página web (tic2_8)</b>	0.221	0.000
<b>Página Web - Acceso a catálogos de productos (tic2_8_2b)</b>	-0.1570	0.002
<b>Página web – Recepción de pedidos (tic2_8_2c)</b>	-0.2585	0.000
<b>Software abierto - Aplicaciones ofimáticas (tic4_2c)</b>	-0.2673	0.000
<b>Apoyo de TIC - Control y seguimiento de pedidos (tic4_4b)</b>	0.4587	0.000
<b>Medios de comunicación - Correo electrónico (tic5_1a)</b>	0.3562	0.002
<b>Capacitación personal no especialista TIC (tic6_2b)</b>	0.2993	0.000

Nota. Elaboración propia

Estas variables representan diferentes características empresariales relacionadas con las compras por internet; por lo cual, escogemos las más representativas en función de sus coeficientes y su p value, por lo cual, es importante identificar cuáles afectan de manera positiva y negativa a la variable dependiente.

Las variables tic2\_6c (Uso de internet - Utilizar servicios en la nube), tic2\_6d (Uso de internet - Obtener información acerca de bienes y servicios), tic2\_8 (Disponibilidad de sitio o página web en la empresa), tic4\_4b (Apoyo en uso de TIC - Control y seguimiento de pedidos), tic5\_1a (Medios de comunicación - Correo electrónico) y tic6\_2b (Capacitación al personal que no es especialista en TIC) tienen un impacto positivo en la variable tic3\_2 (compra de bienes o servicios por internet).

La utilización de servicios en la nube (tic2\_6c) permite acceder a aplicaciones y recursos de manera ágil y flexible, facilitando la realización de compras en línea. Obtener información acerca de bienes y servicios (tic2\_6d) brinda a los consumidores la posibilidad de investigar y comparar opciones antes de realizar una compra, lo cual fomenta la confianza y la toma de decisiones informadas. La disponibilidad de un sitio o página web en la empresa (tic2\_8) proporciona un canal directo y accesible para realizar compras en línea. El control y seguimiento de pedidos (tic4\_4b) facilita la gestión de las transacciones y brinda confiabilidad en el proceso de compra. El uso del correo electrónico (tic5\_1a) como medio de comunicación permite establecer una comunicación eficiente y directa con los clientes. Finalmente, la capacitación al personal que no es especialista en TIC (tic6\_2b) mejora la competencia y el conocimiento en el uso de tecnologías, lo cual se refleja en una mejor atención al cliente y las transacciones en línea.

Por el contrario, las variables tic2\_2\_1a (Banda Ancha Fija - Conexión DSL), tic2\_8\_2b (Servicios de página web - Acceso a catálogos de productos), tic2\_8\_2c (Servicios de página web - Recepción de pedidos o reservas) y tic4\_2c (Apoyo en uso de TIC - Gestión de cadena de suministro, logística, control de inventarios) tienen un impacto negativo en la variable tic3\_2 (compra de bienes o servicios por internet por parte de la empresa).

La conexión DSL (tic2\_2\_1a) puede limitar la velocidad y la capacidad de transmisión de datos, lo que dificulta el acceso y la navegación fluida en línea para realizar compras. El acceso a catálogos de productos (tic2\_8\_2b) puede ser limitado o poco intuitivo, lo que dificulta la visualización y la búsqueda eficiente de productos, lo que puede disminuir la disposición de la empresa para realizar compras en línea. La recepción de pedidos o reservas a través de servicios de página web (tic2\_8\_2c) puede presentar deficiencias en términos de precisión, fiabilidad o integración con los sistemas internos de la empresa, lo que puede generar problemas en la gestión de los pedidos y la logística. La gestión de cadena de suministro, logística y control de inventarios (tic4\_2c) puede ser ineficiente o poco optimizada, lo que puede afectar negativamente la capacidad de la empresa para gestionar adecuadamente los procesos relacionados con las compras en línea.

- **Accuracy**

Para la evaluación de nuestro modelo nuestro accuracy es de: 0.75; esto nos indica que, de las 14.029 observaciones, el modelo predijo correctamente el 75% de veces.

- **Matriz de confusión**

La matriz de confusión se utiliza para evaluar el rendimiento de un modelo de clasificación en base a sus predicciones. En este caso, la clasificación tiene dos clases: Positivo y Negativo.

**Tabla 10**

*Matriz de confusión compras por internet*

		Predichos	
		Negativo (N) -	Positivo (P) +
Actuales	Negativo -	2824	2471
	Positivo +	1029	7705

Nota. Elaboración propia

- Verdaderos positivos (TP): Hay 7705 casos en los que el modelo ha predicho correctamente la clase positiva.
- Falsos positivos (FP): Hay 2471 casos en los que el modelo ha predicho incorrectamente la clase positiva cuando la verdadera clase era negativa. Estos casos se denominan errores de Tipo I o falsos positivos.
- Falsos negativos (FN): Hay 1029 casos en los que el modelo ha predicho incorrectamente la clase negativa cuando la verdadera clase era positiva. Estos casos se denominan errores de Tipo II o falsos negativos.
- Verdaderos negativos (TN): Hay 2824 casos en los que el modelo ha predicho correctamente la clase negativa.

Para este caso, el modelo ha tenido un buen rendimiento en la predicción de la clase negativa (verdaderos negativos) con 2824 aciertos, mientras que ha tenido más dificultades en la predicción de la clase positiva, con 1029 casos incorrectos (falsos negativos).

### 8.1.5 ANÁLISIS MEDIANTE RANDOM FOREST FEATURE IMPORTANCE

Consideramos un análisis prescriptivo en el cual identificamos “Feature Importance” que se refiere a la importancia de las características para nuestro modelo de regresión, en el modelo de bosque aleatorio entrenado mediante “(rf)”. La importancia de una variable se refiere a cuánto contribuye esa variable al rendimiento predictivo del modelo.

De igual manera, al igual que en los anteriores modelos, consideramos a nuestra variable objetivo  $y = \text{tic1\_2}$  (Inversión en TIC), donde obtenemos el siguiente resultado:

**Figura 25**

*Características más importante por Feature Importance*



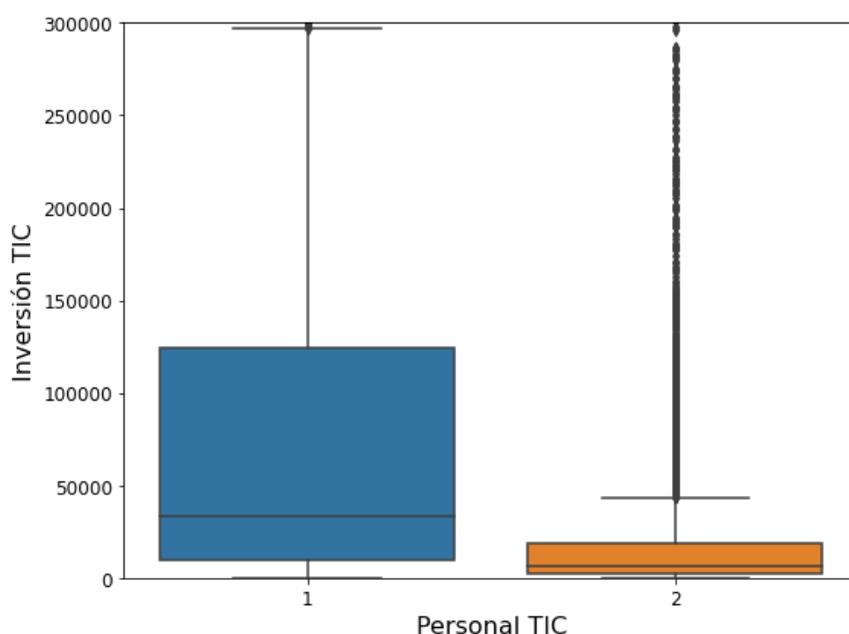
Nota. Elaboración propia

En este sentido, procedemos a realizar un análisis de las variables categóricas más representativas, en función de poder ver la relación existente entre estas y nuestra variable objetivo (Y) que es inversión en TIC (tic1\_2). Como podemos analizar en el gráfico de barras, vemos que las variables categóricas más destacadas son tic6\_1 (Especialistas TIC), tic6\_2a (Capacitación al personal especialista en TIC), tic6\_2c (No capacitación en temática TIC) y tic4\_2d (Software abierto – Servicios web/internet). Para este análisis se eligen las variables categóricas más representativas, considerando su nivel de importancia en el estudio, ya que las otras, se relacionan a categorías de equipos tecnológicos, que no necesariamente son de carácter innovador, ya que muchos de estos tienen una tendencia negativa dentro del análisis exploratorio.

Para el caso de Personal TIC, utilizamos un gráfico de caja el cual nos permite visualizar la variabilidad y distribución en función de su relación con la variable objetivo (Inversión en TIC). En tal sentido, podemos reconocer que existe una distribución de la Inversión en TIC para cada nivel de la variable Personal Especialista TIC, donde vemos que la relación más alta se da en el nivel 1, que nos indica que las empresas sí cuentan con personal especialista, siendo su mediana menor a 50.000, tomando como medida nuestro eje Y (tic1\_2). En el caso del nivel 2, que representa que “no” exista personal especialista, sus valores son considerablemente menores, al igual que su mediana, en alrededor 10.000.

**Figura 26**

*Relación entre Inversión TIC y Personal TIC*



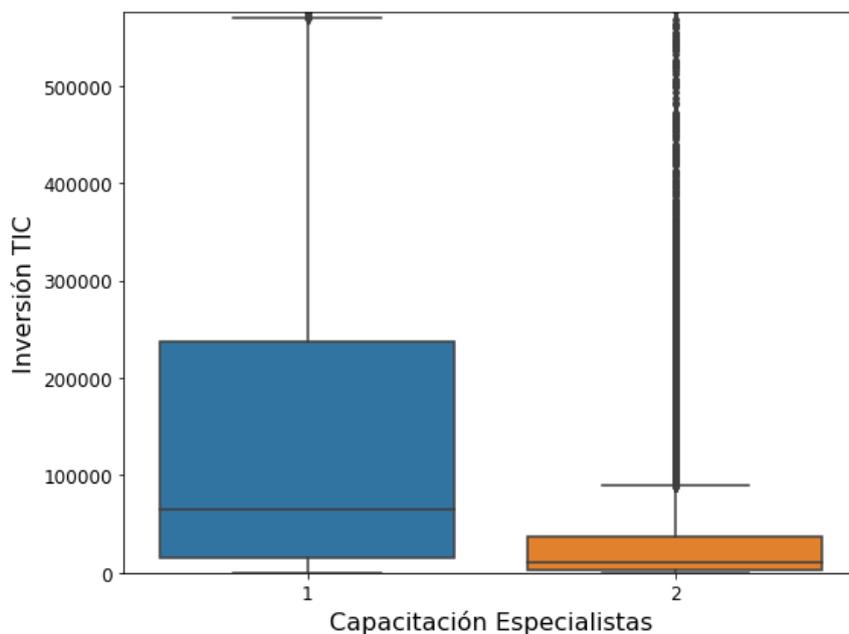
Nota. Elaboración propia

La segunda variable más representativa es la capacitación brindada al personal especialista en TIC, lo cual, nos permite de igual forma, mediante un diagrama de caja, incidir que el nivel 1 “sí”, existe una relación significativa con nuestra variable objetivo, donde si bien, su mediana se encuentra en 60.000 vemos que, en su pico más alto, supera incluso el medio millón, pues las empresas pueden ser capaces de invertir en capacitación para especialistas. Para nuestro objeto

de estudio, nos está indicando que esta inversión, puede ser representativa, al gestionarla en la formación de expertos en tecnologías.

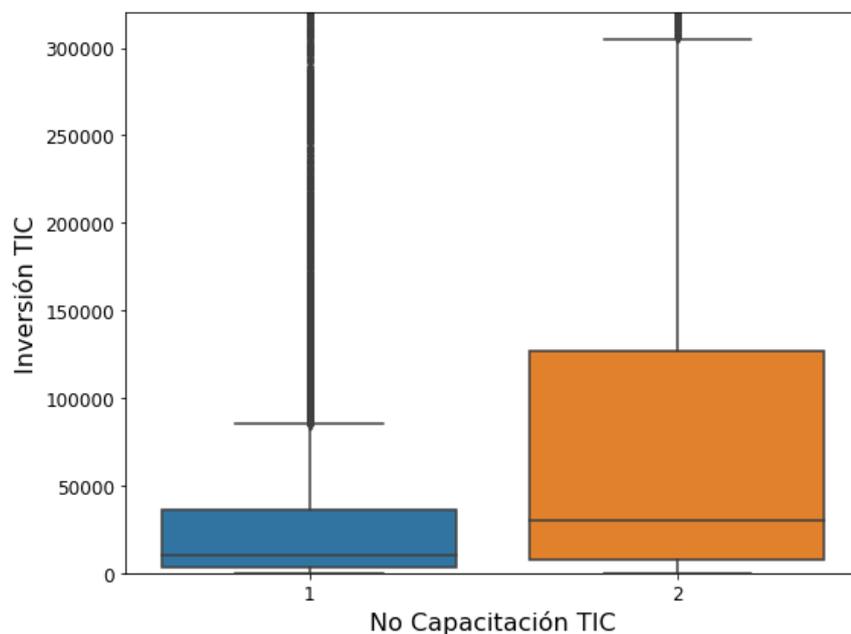
**Figura 27**

*Relación entre Inversión TIC y Capacitación Especialistas TIC*



Nota. Elaboración propia

Podemos reconocer como en el caso de nuestra variable `tic6_2c`, sucede un análisis contrario al de los 2 anteriores, ya que podemos reconocer como la caja con mayor representatividad es la del 2do nivel, en la que podemos analizar que no “no se ha realizado capacitación en la temática TIC”, de manera que, en su sentido contrario quiere decir que en efecto sí han sido realizadas capacitación en estas temáticas, lo cual para nuestro objeto de estudio guarda una relevancia significativa, por el apoyo al personal tanto especialista, como no especialista.

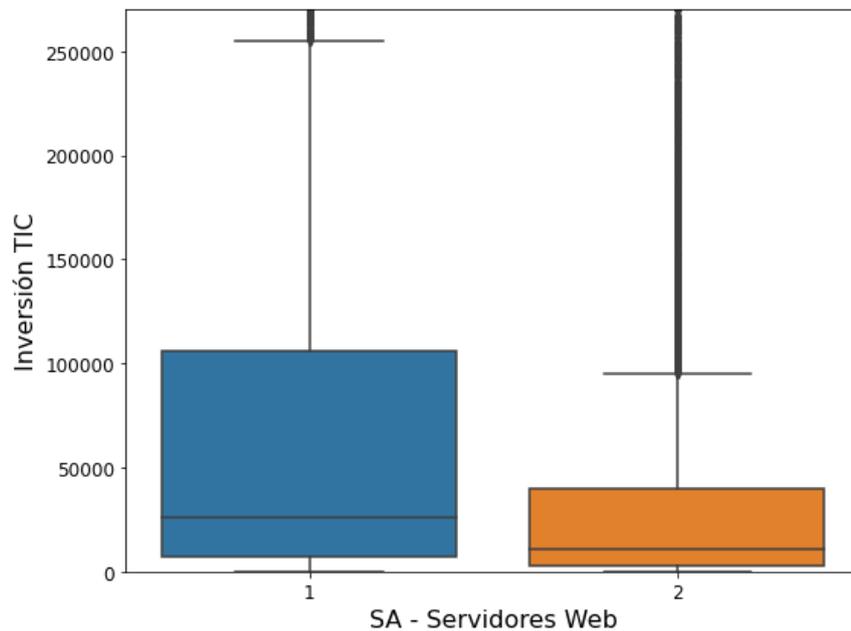
**Figura 28***Relación Inversión TIC y No Capacitación TIC*

Nota. Elaboración propia

En este caso, cambiamos un poco la dinámica ya que nos enfocamos directamente en las herramientas TIC, que, para efectos de nuestro estudio, señalamos que una variable destacada es el uso de herramientas como el caso de software abierto para servidores web o de internet. En nuestro caso, vemos que si analizamos la relación con nuestra variable objetivo (inversión en TIC), su mediana no es un valor considerablemente alto, ya que este tipo de sistemas son poco usados, alcanzando alrededor de 25.000 en inversión; sin embargo, existe un máximo de hasta 250.000, donde hay empresas que apuestan por el uso de este tipo de sistemas abiertos, deciden invertir.

**Figura 29**

*Relación entre Inversión en TIC y SA Servidores Web*



Nota. Elaboración propia

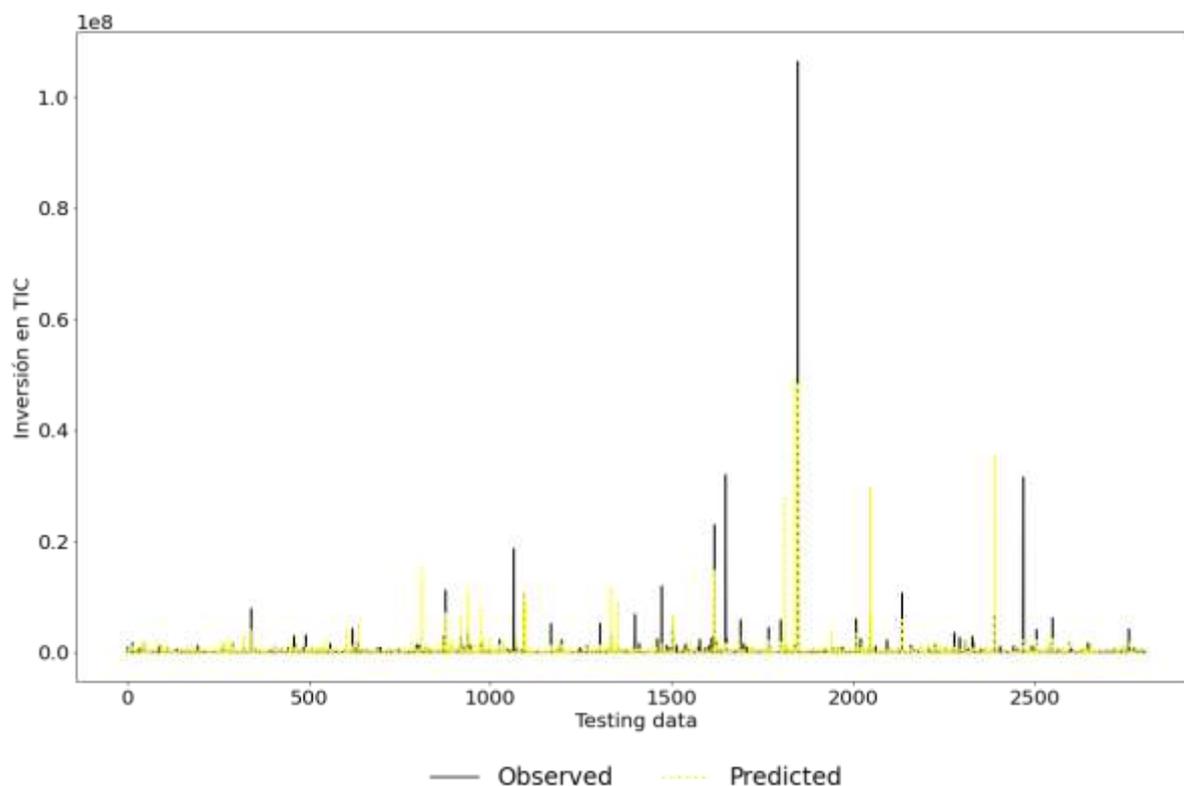
### 8.1.6 ANÁLISIS MEDIANTE RANDOM FOREST REGRESSOR

Para trabajar con este modelo de aprendizaje supervisado, nuestra variable dependiente es  $y=tic1\_2$  (Inversión TIC) una variable de tipo numérica, que nos permite analizar de manera general los demás elementos relacionados a la transformación digital empresarial, al ser considerada dentro del primer segmento de variables.

En este sentido, nos interesa poder evaluar los niveles de ajuste, en función de los datos de nuestra base *BDTICMODEL*, para de esta forma poder reconocer mediante un *plotting*, los valores observados y los que el modelo predice.

**Figura 30**

*Análisis valores observados vs. valores predichos*



Nota. Elaboración propia

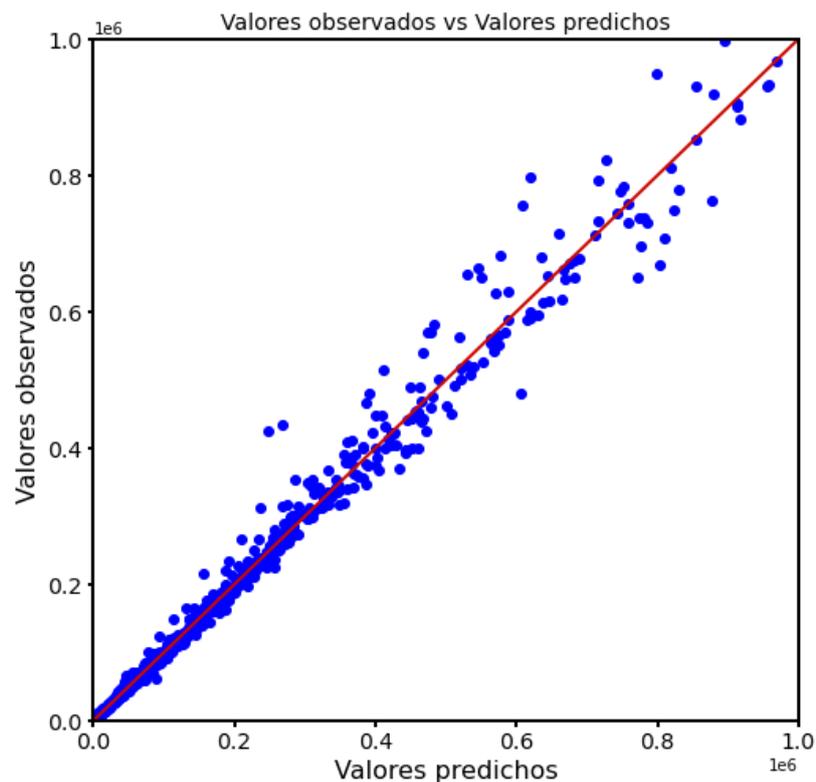
Este gráfico podemos comparar visualmente cómo se ajustan los valores predichos a los valores observados. Si los valores predichos están cerca de la línea negra sólida, significa que el modelo ha hecho predicciones precisas y está capturando adecuadamente las fluctuaciones y tendencias. Si los valores predichos se desvían significativamente de la línea negra sólida, indica que el modelo puede tener dificultades para ajustarse a los datos observados; sin embargo, en este caso, considerando que la precisión para los datos de *test\_data* es alta con una precisión de 0.89, gráficamente podríamos indicar que en la mayoría de los casos ambas líneas buscan un nivel de similitud, por lo cual, podemos incidir de que el modelo es eficiente.

De igual manera, es importante poder identificar el modelo a través de un scatterplot, a través de un gráfico de dispersión, donde utilizaremos puntos azules para representar los valores observados y los valores predichos para

(`ytest_pred`) y (`y_pred`); debemos recordar que nuestras variables independientes para este modelo son todas las variables de la base de datos, excepto la nuestra `y=tic1_2` (Inversión TIC).

### Figura 31

*Gráfico de dispersión Observados vs Predichos*



Nota. Elaboración propia

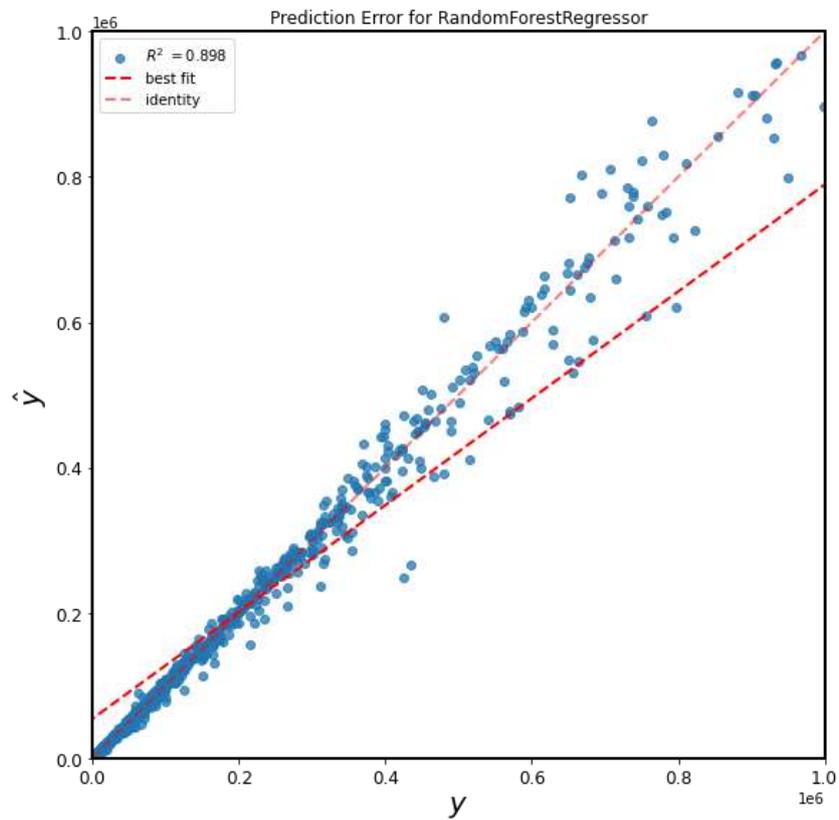
De manera complementaria, utilizamos una nueva librería que es: `"from yellowbrick.regressor import PredictionError"`: Esto importa la clase `PredictionError` de la biblioteca `Yellowbrick`, que se utiliza para crear gráficos de error de predicción en modelos de regresión.

A partir de `"PredictionError(tic1_2_rf)"`, podemos crear una instancia de predicción de error, donde nos referimos al modelo de regresión en el que se visualiza el error. Por lo cual, se evalúa el rendimiento óptimo del modelo utilizando el conjunto de prueba de características (`X_test`) y el conjunto de prueba de valores objetivo (`y_test`). El método `"score"` del visualizador indica el

coeficiente de determinación ( $r^2$ ) del modelo, que es una medida de qué tan bien se ajustan las predicciones a los valores reales.

### Figura 32

#### Ajuste para Error de Predicción



Nota. Elaboración propia

En este caso observamos que esta predicción de error se mantiene cerca de la línea de ajuste, lo cual indica que su ajuste es adecuado, al cruzarse ambas líneas.

## 9. PROPUESTA DE SOLUCIÓN

### 9.1 IMPLICACIONES PARA LA ORGANIZACIÓN

Un análisis profundo de todas las variables relacionadas a las tecnologías de la información y comunicación TIC, dentro del contexto empresarial ecuatoriano es primordial para entender el nivel de influencia positiva que obtendrán las empresas que opten por el camino de la innovación y transformación digital. Para este fin se partió desde la analítica descriptiva, predictiva y prescriptiva, las mismas que son detalladas a continuación:

- **TIC en las empresas durante el período 2016-2021.**

A lo largo de los años, se ha podido detectar mediante el análisis descriptivo el comportamiento de los diferentes tamaños de empresas en lo que respecta a la implementación de la tecnología. Si bien es cierto, la encuesta segmenta las empresas en 3 tipos de tamaño, a lo largo del estudio se observa que la denominada mediana y pequeña empresa en lo que respecta a TIC tienen un nivel de aproximación similar en la mayoría de los campos como: uso, gestión, medios de comunicación y capacitación, etc. Un aspecto sin duda relevante en el estudio es el nivel de inversión que han tenido los 3 tipos empresa, ya que mientras las denominadas grandes empresas pudieron aumentar la inversión en tecnología en más del 100%, la mediana aumento apenas un 12 % y por el contrario la pequeña empresa redujo el monto en un 37% desde 2016 al 2021.

- **Identificación de tendencias y optimización de recursos.**

Al analizar los datos y aplicar el algoritmo K-means, el cual consiste en desarrollar un proceso no supervisado para encontrar grupos significativos en los datos, la organización puede identificar las variables o indicadores que permitan desplegar una correcta estrategia de transformación digital.

Asimismo, al contar con una variedad de datos a través de grupos definidos utilizando la encuesta realizada por el INEC, hemos logrado identificar con

precisión cada elemento junto con su correspondiente opción para superar este desafío.

Bajo estos conceptos, se concluye que el factor determinante dentro de las empresas es el nivel de inversión y presupuesto con el que cuenta cada una de ellas, ya que, si bien es cierto el despliegue de infraestructura y de servicios tecnológicos en el último tiempo han permitido adquirir diferentes tipos de tecnología, estas aun han sido poco o subutilizadas en su rendimiento para el cual han sido creadas.

- **Reconocer aquellas características que afectan la transformación digital**

El análisis de datos, junto con la técnica de regresión logística, permiten identificar las características que afectan positiva y negativamente la transformación digital de las empresas. Es así como, mediante el análisis de la base de datos, se determinó que está en su mayoría está conformada por variables de tipo categóricas. Por lo tanto, mediante esta técnica se va a estudiar las ventas y compras de bienes y servicios por internet mediante las variables mencionadas. Al aplicar la regresión logística, se pueden modelar las relaciones entre estas variables y determinar qué características tienen un impacto significativo en el Ecommerce. Esto proporciona a las empresas una comprensión más clara de los factores que impulsan o limitan su capacidad de comercio electrónico, lo que les permite tomar decisiones estratégicas basadas en evidencia y diseñar estrategias más efectivas para superar los desafíos identificados.

- **Arboles de decisión para analizar datos y predecir resultados con alta precisión**

Random Forest, basado en el aprendizaje supervisado, se convierte en una herramienta valiosa para entender de mejor forma la problemática propuesta. En el caso particular se eligió la variable objetivo de entrenamiento a la Inversión en Tic (tic1\_2) que generan los 3 tipos de empresa a lo largo de 6 años. A través de

este análisis de datos, se evidenció que el modelo es capaz de predecir el monto de inversión tomando en cuenta todas las variables tanto categóricas como numéricas, con una precisión del 89 % lo que muestra el nivel de efectividad del modelo. Por lo tanto, Random Forest es capaz de aprender patrones complejos y relaciones entre las variables, lo que permite a las empresas tomar decisiones informadas en su proceso de transformación digital, logrando mejores resultados en su evolución tecnológica.

## **9.2 ESTRATEGIA ORGANIZACIONAL**

### **1. Fomentar la inversión en tecnologías digitales y software empresariales**

Una estrategia clave para abordar los desafíos en la adopción de tecnologías digitales en las empresas de Ecuador es fomentar la inversión en software empresariales y aplicaciones móviles que permitan la innovación y la transformación digital. Es importante identificar y adquirir herramientas tecnológicas adecuadas para cada empresa, considerando sus necesidades y objetivos específicos. Algunos ejemplos de software empresariales que pueden impulsar la transformación digital incluyen:

- **Sistemas de gestión empresarial (ERP):** Estos sistemas integran diferentes áreas y procesos de la empresa, como finanzas, recursos humanos, inventario y producción, en una única plataforma. Al implementar un ERP, las empresas pueden mejorar la eficiencia operativa, agilizar la toma de decisiones y optimizar la gestión de recursos. Un ejemplo de software gratuito es la plataforma llamada BITRIX que brinda servicios de CRM de manera gratuita.
- **Herramientas de análisis de datos:** El análisis de datos se ha vuelto fundamental en la toma de decisiones empresariales. La inversión en herramientas de análisis de datos, como software de visualización y minería de datos, puede ayudar a las empresas a obtener información valiosa, identificar patrones, realizar pronósticos y optimizar sus operaciones.

## **2. Telecomunicaciones eficientes**

Para una verdadera transformación digital, el garantizar la disponibilidad y expansión de la red de fibra óptica en el área empresarial es primordial, ya que sin ella la conectividad y la tardía implementación de las soluciones que brindan las empresas ya no tendrían su efectividad inicial. Esto incide directamente en el uso que se le da al internet en los diferentes tipos de empresa, si se lo analiza desde la óptica de la grande empresa, en este grupo se registra un mayor aprovechamiento de este recurso que está enfocado principalmente en incentivar la capacitación, los servicios en la nube y la telefonía IP, con menos despliegue para la entrega de productos online. A su vez y en menor proporción, la mediana y pequeña empresa también aprovecha los recursos mencionados sin tanto énfasis en el tema de capacitación. En este sentido es importante concientizar a los gerentes o administradores de la importancia del conocimiento sobre la tecnología, acompañado de prácticas laborales entorno a su nicho de negocio.

## **3. Páginas de aterrizaje como alternativa**

Para el tema de la página web, el mayor aprovechamiento se ha dado en el área de marketing, sin embargo, aún no se ha logrado explotar todas sus ventajas que el cliente espera al momento de la interacción. En este sentido en la actualidad, existen paginas diseñadas específicamente para una campaña de marketing o promoción puntual (landing page), estas son paginas simples que pueden funcionar mejor en organizaciones donde el presupuesto es una limitante para la difusión de productos.

## **4. Implementar una plataforma de comercio electrónico**

Para potenciar las ventas en línea y mejorar la experiencia del cliente, es esencial implementar una plataforma de comercio electrónico robusta y efectiva. Esta plataforma permitirá a las empresas ecuatorianas expandir su alcance y aprovechar las oportunidades que ofrece el comercio electrónico. Al seleccionar una plataforma adecuada, es importante considerar aspectos como la facilidad

de uso, la seguridad de las transacciones, la personalización y la capacidad de integración con otros sistemas empresariales. También es esencial ofrecer diversas opciones de pago y envío, para adaptarse a las preferencias y necesidades de los clientes. De forma complementaria, es necesario reconocer la importancia de la gestión eficiente de devoluciones y reclamaciones, así como la personalización de la comunicación con los clientes.

#### **5. Hardware más sofisticados.**

En las empresas grandes buscan equipos más sofisticados como computadores y laptop de última generación. En este aspecto, una alternativa relativamente económica es el uso de infraestructura montada en la nube, esto podría generar ahorros y permitir que las empresas de menor tamaño puedan manejar estos equipos que sin duda serán de gran apoyo para sus actividades económicas.

#### **6. Redes sociales como ejes de la comunicación**

La comunicación es un factor determinante para la interacción entre el cliente y las empresas, para las pequeñas y medianas empresas los medios tradicionales y que conlleven una baja inversión como es las redes sociales o la telefonía móvil son una alternativa eficaz dentro de sus negocios, por otro lado las grandes empresas que también han aprovechado de los medios tradicionales, en este caso aún se mantiene la interacción mediante los medios impresos, lo cual es una forma de mostrar su hegemonía en el mundo de los negocios. Sin duda una solución económica y rentable en la actualidad son las plataformas “gratuitas” como Facebook y WhatsApp, que gracias a su popularidad puede atraer más cliente con sus estrategias de marketing conocidas.

#### **7. Fortalecer el apoyo en el uso de tecnologías de la información y la comunicación (TIC) para la gestión empresarial**

Una estrategia clave para mejorar la toma de decisiones gerenciales y superar los desafíos de la transformación digital en las empresas ecuatorianas es fortalecer el apoyo en el uso de TIC para la gestión de las diferentes áreas de la empresa. Es fundamental reconocer la importancia de la gestión de relaciones

con los clientes y el seguimiento a la calidad de servicio como un factor clave para el éxito empresarial. En este sentido, se debe implementar y optimizar el uso de herramientas digitales que permitan una comunicación eficiente y personalizada con los usuarios, tal es el caso de herramientas innovadoras como aplicativos móviles tipo ERP, CRM, plataformas de aprendizaje y software de seguridad.

#### **8. Capacitación en TIC (Tecnologías de la Información y Comunicación)**

En los últimos años, la capacitación se ha vuelto un punto de inflexión en las empresas que buscan la innovación, En este punto la adquisición de habilidades debe ser una estrategia infaltable en todas las empresas, ya que esto incide directamente en la optimización para la toma de decisiones lo que permite a las empresas ser más competitivas al optimizar sus procesos internos, ofrecer mejores productos y servicios digitales y brindar una mejor experiencia al cliente.

### **9.3 ESTRATEGIA ORGANIZACIONAL, INNOVACIÓN Y COMPETITIVIDAD EMPRESARIAL**

La innovación desempeña un papel clave en el desarrollo de la madurez tecnológica empresarial. La adopción de tecnologías digitales y la implementación de soluciones innovadoras permiten a las empresas mejorar sus procesos internos, optimizar la gestión de recursos, aumentar la eficiencia operativa y ofrecer productos y servicios de mayor calidad. La innovación también ayuda a las empresas a adaptarse a los cambios del mercado y aprovechar nuevas oportunidades, lo que contribuye a su competitividad (Doria & López, 2023). La competitividad empresarial, por su parte, está estrechamente relacionada con la madurez tecnológica. Las empresas que logran una mayor madurez tecnológica a través de la adopción estratégica de tecnologías digitales y la innovación tienen una ventaja competitiva en el mercado.

Para nuestro objeto de estudio, se ha dirigido una propuesta de solución, que nos permite abordar estrategias de carácter organizacional, para impulsar componentes de innovación que, a su vez, generen mayor competitividad en el sector empresarial ecuatoriano.

- **Innovación:** Todas estas estrategias están orientadas hacia la innovación, ya que implican la adopción y utilización de tecnologías digitales en las empresas. La inversión en software empresariales, el desarrollo de plataformas de comercio electrónico y el uso de herramientas de análisis de datos son formas de innovar en los procesos y operaciones empresariales. Además, fortalecer el apoyo en el uso de TIC para la gestión empresarial implica la adopción de herramientas tecnológicas que fomentan la innovación en la gestión de relaciones con los clientes, el control y seguimiento de pedidos, y la gestión del conocimiento.
- **Competitividad empresarial:** Estas estrategias contribuyen a mejorar la competitividad empresarial al aprovechar las oportunidades que ofrece la transformación digital. La inversión en tecnologías digitales y software empresariales, así como la implementación de plataformas de comercio

electrónico, amplían el alcance de las empresas, les permiten llegar a nuevos mercados y competir en un entorno globalizado. Al mismo tiempo, mejorar la presencia digital y la experiencia del cliente en línea aumenta la satisfacción del cliente, lo que puede generar lealtad y ventaja competitiva. Fortalecer el apoyo en el uso de TIC para la gestión empresarial también contribuye a la competitividad al mejorar la eficiencia operativa, reducir costos y aumentar la productividad.

## 10. CONCLUSIONES Y RECOMENDACIONES

### 10.1 CONCLUSIONES

- Los clústeres son el reflejo de los 3 tamaños de empresa, en efecto las grandes empresas al invertir mayores cantidades en software apuestan por la transformación digital y la innovación, lo que les permite continuar desempeñando un papel cada vez más importante en la estrategia empresarial y en la búsqueda de la ventaja competitiva. Por el lado opuesto la mediana y pequeña empresa que tienen recursos limitados, se enfocan principalmente en satisfacer necesidades básicas como es el acceso al internet (telecomunicaciones), infraestructura de bajo costo que garantice el soporte y el servicio hacia sus clientes.
- Al analizar los modelos de regresión propuestos (Random Forest, Regresión Logística y Regresión Lineal), pudimos identificar que el Random Forest nos generó un mayor nivel de fiabilidad, considerando que tiene un nivel de precisión del 89% al explicar la variabilidad de los datos. En este sentido, su nivel de precisión nos permite generar predicciones y clasificación más confiables, así como determinar las variables con mayor incidencia y su repercusión en la inversión en TIC.
- Se observa una clara disparidad en los porcentajes de inversión en TIC entre diferentes tipos de empresas. Mientras que la gran empresa logró un significativo aumento del 103% en su inversión, la mediana empresa experimentó un incremento más moderado del 11.94%. Por otro lado, es preocupante que la pequeña empresa haya disminuido su inversión en un 37%, lo cual puede tener implicaciones en su capacidad para adaptarse y competir en el entorno digital.
- Se ha evidenciado la importancia de características clave para mejorar la innovación y la transformación digital empresarial en el contexto actual. La inversión en servicios de licencias y desarrollo de software se presenta

como una necesidad fundamental para impulsar el progreso tecnológico en las organizaciones. Asimismo, la conexión y el uso de internet desempeñan un papel crucial al permitir el acceso a servicios en la nube. Además, contar con una plataforma estable para el comercio electrónico.

- La formación en TIC ha sido dirigida principalmente hacia empleados sin conocimientos especializados, sobre todo en empresas de menor tamaño. Aunque esto podría sugerir que hay una mayor proporción de especialistas en este grupo, al evaluar la contratación de especialistas en comparación con las grandes empresas, se constata una reducción del 50%.

## 10.2 RECOMENDACIONES

- Ante la disparidad en los porcentajes de inversión en TIC entre diferentes tipos de empresas, se sugiere brindar apoyo y asesoramiento a las pequeñas y medianas empresas para que puedan incrementar su inversión en tecnologías digitales. Esto puede incluir la difusión de programas de financiamiento, la promoción de alianzas estratégicas con proveedores de tecnología y la capacitación en gestión de proyectos de transformación digital. Además, es importante concienciar sobre los beneficios y oportunidades que la adopción de tecnologías digitales puede brindar a estas empresas en términos de eficiencia operativa y competitividad.
- Dado que se ha destacado la importancia de la inversión en servicios de licencias y desarrollo de software, la conexión a internet y una plataforma estable para el comercio electrónico, se recomienda a las empresas en Ecuador que realicen una evaluación exhaustiva de sus necesidades tecnológicas y realicen inversiones estratégicas en estas áreas. Es esencial buscar proveedores confiables, realizar pruebas piloto y evaluar el retorno de la inversión para garantizar que las soluciones tecnológicas seleccionadas sean adecuadas y contribuyan a la innovación y la transformación digital de la empresa.
- Contar con un sitio web y utilizar servicios de página web para ofrecer catálogos de productos, recibir pedidos o reservas, y brindar información relevante es crucial para establecer una presencia digital sólida. Esto permitirá a la empresa llegar a nuevos clientes, expandir su mercado y facilitar la interacción con los consumidores, a partir de la implementación eficiente del comercio electrónico.
- Es crucial fomentar una mayor formación en TIC tanto para el personal no especializado como para los especialistas en empresas de menor tamaño. Además, se debe estimular la contratación de expertos en

tecnología en las organizaciones medianas y pequeñas para equilibrar la proporción de personal capacitado en TIC. Esto fortalecerá la capacidad tecnológica en todas las empresas, maximizando así el potencial de la transformación digital y asegurando su competitividad en el mercado actual. Asimismo, se deben implementar programas de capacitación continua y actualizar regularmente las habilidades técnicas de los empleados para mantenerse al día con los avances en el campo de las TIC.

- Se sugiere llevar a cabo una evaluación exhaustiva de los factores que podrían obstaculizar la adopción de tecnologías para la transformación digital de las empresas en Ecuador. Esto incluye identificar barreras como la falta de infraestructura adecuada, la brecha digital, la resistencia al cambio y la falta de conocimientos técnicos. Una vez identificados estos obstáculos, se deben desarrollar estrategias y políticas específicas para abordar cada uno de ellos.
- Para futuros análisis se recomienda utilizar el algoritmo K-prototypes. El cual permite una agrupación precisa considerando tanto valores numéricos como variables categóricas. Esto ayuda a identificar patrones y segmentar participantes en grupos homogéneos. Es una herramienta valiosa para comprender los datos de encuestas de manera más completa y representativa.

## 11. REFERENCIAS

- Agencia de Regulación y Control de las Telecomunicaciones (ARCOTEL). (2020). *Informe Estadístico Anual de las Telecomunicaciones*. Recuperado de <https://www.arcotel.gob.ec/wp-content/uploads/2021/03/Informe-Anual-2020.pdf>
- Banco Interamericano de Desarrollo (BID). (2019). *Las empresas latinoamericanas invierten poco en tecnología*. Recuperado de <https://www.iadb.org/es/noticias/las-empresas-latinoamericanas-invierten-poco-en-tecnologia>
- Botto-Tobar, M., Zambrano Vizuete, M., Torres-Carrión, P., Montes León, S., Pizarro Vásquez, G., & Durakovic, B. (Eds.). (2020). *Applied Technologies. Communications in Computer and Information Science*. doi:10.1007/978-3-030-42517-3
- Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). *Classification and Regression Trees*. New York: Chapman and Hall/CRC.
- Breiman, L. (2017). *Classification and regression trees*. Routledge.
- Cano Giner, J. (2007). *Business Intelligence: competir con información*. Fundación Banesto.
- Chitarroni, H. (2022). La regresión logística. Instituto de Investigación en Ciencias Sociales. *Universidad del Salvador*. Recuperado desde <https://racimo.usal.edu.ar/83/1/Chitarroni17.pdf>
- Dini, M., Gligo, N. & Patiño, A. (2021). Transformación digital de las mipymes. Elementos para el diseño de política. *Documentos de Proyectos (LC/TS.2021/99)*. Santiago, Comisión Económica para América Latina y el Caribe (CEPAL).
- Dominik & Modrák, M., V. & Zsifkovits, H. (2020). *Industry 4.0 for SMEs Challenges, Opportunities and Requirements: Challenges, Opportunities and Requirements*. 10.1007/978-3-030-25425-4.

- Doria Pérez, V., & López Hoyos, N. A. (2023). *Una década de estrategias competitivas empresariales en Colombia*. Recuperado desde <https://repository.ucc.edu.co/items/3e050e5e-a4ec-41a6-b5fe-d3bde8618587>
- Escorza, P.; y Valls, J. (2001), *Tecnología e Innovación en la Empresa: Dirección y Gestión*. México, D.F: Alfaomega.
- Fernández Diez, M.C. y Puig Gabarró, P. (2020). Los desafíos del comercio electrónico para las PYME. *Banco Interamericano de Desarrollo*.
- Forbes. (2020). MercadoLibre utiliza inteligencia artificial para personalizar la experiencia de compra. *Revista Forbes*. Recuperado de <https://www.forbes.com.mx/negocios-mercadolibre-utiliza-inteligencia-artificial-para-personalizar-la-experiencia-de-compra/>
- Frost, C. & Sullivan. (2020). *Servicios de telecomunicaciones para la transformación digital en Latinoamérica*. Recuperado desde <https://ww2.frost.com/wp-content/uploads/2020/11/LATAM-Digital-Transformation-Services-Market-2020.pdf>
- Gartner. (2020). *Gartner Says Worldwide IT Spending to Grow 3.2% in 2019*. Recuperado desde <https://www.gartner.com/en/newsroom/press-releases/2019-07-10-g>
- Gerstner, T., & Kugler, A. (2018). *Introduction to mathematical statistics and its applications*. Pearson.
- Hastie, T., Tibshirani, R. y Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer.
- Instituto Nacional de Estadísticas y Censos (2022). *Diseño muestral ENESEM 2020*. Quito.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An introduction to statistical learning: with applications in R*. Springer.

- Mafla, W. (2022). Obtención de curvas de caga para los diferentes tipos de usuarios residencial, comercial e industrial a través del algoritmo. [Tesis de grado]. *Escuela Politécnica Nacional*.
- Maldonado, G., Mojica, J. E., & Molina, V. M. (2013). La relación entre la Innovación, las Tics y la calidad: Una Perspectiva de La Pyme Iberoamericana. *Global Conference On Business & Finance Proceedings*, 1135-1140.
- McKinsey. (2018). *Digital Globalization: The new era of global flows*. Recuperado de <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/digital-globalization-the-new-era-of-global-flows>
- Mergel, I., Edelmann, N., & Haug, N. (2019). Defining digital transformation: Results from expert interviews. *Government Information Quarterly*, vol. 36 (4), 1-16. DOI: <https://doi.org/10.1016/j.giq.2019.06.002>
- Ministerio de Economía, Fomento y Turismo de Chile. (2020). Informe General de Resultados: Encuesta de Acceso y Uso de Tecnología de Información y Comunicación (TIC) en Empresas. *División Política Comercial e Industrial*. Unidad de Estudios.
- Ministerio de Telecomunicaciones y de la Sociedad de la Información (s.f.). Comercio electrónico, una oportunidad para el desarrollo de negocios a través de la web. *Gobierno del Ecuador*. Recuperado desde <https://www.telecomunicaciones.gob.ec/comercio-electronico-una-oportunidad-para-el-desarrollo-de-negocios-a-traves-de-la-web/>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2015). *Introduction to linear regression analysis* (5th ed.). Wiley.
- Oliva, V. (2019). *¿Cómo influye la tecnología en las empresas?*. Admisión UTEM.
- Organisation for Economic Co-operation and Development (OCDE). (2019). *The Future of Business: OECD Business and Finance Outlook 2019*. Recuperado de <https://www.oecd-ilibrary.org/docserver/31e70832->

[en.pdf?expires=1620347683&id=id&accname=guest&checksum=48F14C9B9E42B7A87E342C31D7C46B6E](https://www.researchgate.net/publication/354111111/en.pdf?expires=1620347683&id=id&accname=guest&checksum=48F14C9B9E42B7A87E342C31D7C46B6E)

Ortega, J. A. (2000). Gestión de la tecnología, innovación y competitividad empresarial. *Revista Ingeniería y competitividad*.

Prateek. (2022). K means clustering algorithm. *KeyToDataScience*.  
<https://keytodatascience.com/k-means-clustering-algorithm/>

PwC. (2017). *Industry 4.0: Building the digital enterprise*. Recuperado de  
<https://www.pwc.com/gx/en/industries/industries-4.0/landing-page/industry-4.0-building-your-digital-enterprise-april-2017.pdf>

S.N. (2019). Cervecería Nacional apuesta por la automatización para ser más eficiente. *El Universo*. Recuperado de  
<https://www.eluniverso.com/noticias/2019/10/07/nota/7539606/cerveceria-nacional-apuesta-automatizacion-ser-mas-eficiente>

Vásquez, B. & Orozco, J. (2018). Transformación digital. El poder de los medios digitales como herramienta estratégica para mejorar la competitividad en las PYMES. *Red Internacional de Investigadores en Competitividad: XVI Congreso*. Recuperado desde  
<https://www.riico.net/index.php/riico/article/view/2108>

Walterman, R. (2021). *Digitalización y transformación digital: Definición y principales diferencias*. Recuperado desde  
<https://walterman.es/diferencias-transformacion-digital-y-digitalizacion/>

## ANEXOS

## ANEXO 1. – Base de datos modificada para estudio actual.

Código	Descripción	Numérica
des_tamano	Tamaño de la empresa	Categórica
tic1_2	Valor invertido TIC	Numérica
tic1_3a	Inversión - Bienes TIC	Numérica
tic1_3b	Inversión - Software	Numérica
tic1_3c	Inversión - Consultorías en TIC	Numérica
tic1_3d	Inversión - Otros	Numérica
tic2_1	¿Disponía su empresa de conexión a internet?	Categórica
tic2_2_1a	Banda Ancha Fija – Conexión DSL	Categórica
tic2_2_1b	Banda Ancha Fija – Cable Módem	Categórica
tic2_2_1c	Banda Ancha Fija – Fibra óptica	Categórica
tic2_3	¿Disponía la empresa de conexión por Banda Ancha Móvil para acceder a internet?	Categórica
tic2_6a	Uso de internet – Enviar y recibir correos electrónicos	Categórica
tic2_6b	Uso de internet – Telefonía a través de Internet/VoIP	Categórica
tic2_6c	Uso de internet – Utilizar servicios en la nube	Categórica
tic2_6d	Uso de internet – Obtener información acerca de bienes y servicios	Categórica
tic2_6f	Uso de internet – Efectuar transacciones bancarias	Categórica
tic2_6g	Uso de internet – Acceder a otros servicios financieros	Categórica
tic2_6h	Uso de internet – Brindar servicio al cliente	Categórica
tic2_6i	Uso de internet – Entregar productos digitales en línea	Categórica
tic2_6k	Uso de internet – Capacitar al personal	Categórica
tic2_8	¿Disponía de sitio o página web su empresa ?	Categórica
tic2_8_2b	Servicios de página web – Acceso a catálogos de productos	Categórica
tic2_8_2c	Servicios de página web – Recepción de pedidos o reservas	Categórica
tic3_1	Durante 2021, ¿su empresa vendió bienes o servicios por internet?	Categórica
tic3_1_2b	Obstáculos ventas internet – Logística	Categórica
tic3_1_2c	Obstáculos ventas internet – Pago	Categórica
tic3_1_2d	Obstáculos ventas internet – Seguridad informática y protección de datos	Categórica
tic3_1_2e	Obstáculos ventas internet – Normativa legal del país	Categórica
tic3_1_2f	Obstáculos ventas internet – Costo de venta por internet mayor a beneficio	Categórica
tic3_2	Durante 2021, ¿su empresa compró bienes o servicios por internet?	Categórica
tic4_1a	Número dispositivos - Computadoras de escritorio	Numérica
tic4_1b	Número dispositivos - Laptop o notebook	Numérica
tic4_1d	Número dispositivos - Smartphone	Numérica
tic4_1e	Número dispositivos - Tablet	Numérica
tic4_2a	Software abierto - Sistemas operativos (tipos de software de código abierto)	Categórica
tic4_2b	Software abierto - Navegadores de internet	Categórica
tic4_2c	Software abierto - Aplicaciones ofimáticas	Categórica
tic4_2d	Software abierto - Servidores web/internet	Categórica

tic4_2e	Software abierto - Otros como aplicaciones tipo ERP o CRM, plataformas de aprendizaje, software de seguridad	Categoría
tic4_4a	Apoyo en uso de TIC - Gestión de relaciones con los clientes	Categoría
tic4_4b	Apoyo en uso de TIC - Control y seguimiento de pedidos	Categoría
tic4_4c	Apoyo en uso de TIC - Gestión de cadena de suministro, logística, control de inventarios	Categoría
tic4_4d	Apoyo en uso de TIC - Gestión de finanzas y presupuestos	Categoría
tic4_4e	Apoyo en uso de TIC - Gestión de recursos humanos	Categoría
tic4_4f	Apoyo en uso de TIC - Servicio y soporte a ventas	Categoría
tic4_4g	Apoyo en uso de TIC - Apoyo al desarrollo productivo	Categoría
tic4_4h	Apoyo en uso de TIC - Gestión del conocimiento	Categoría
tic5_1a	Medios de comunicación - Correo electrónico	Categoría
tic5_1b	Medios de comunicación - Teléfonos	Categoría
tic5_1d	Medios de comunicación - Call center	Categoría
tic5_1f	Medios de comunicación - Redes sociales	Categoría
tic5_1h	Medios de comunicación - Medios impresos	Categoría
tic6_2a	Capacitación al personal especialista en TIC	Categoría
tic6_2b	Capacitación al personal que no es especialista en TIC	Categoría
tic6_2c	No se ha realizado capacitación en temática TIC	Categoría
Total_Personal	Total personal	Numérica
Total_Espec	Total Especialistas	Numérica