

Análisis Descriptivo del Estado y la Velocidad de los Accesos Fijos a Internet por Proveedor y Segmento en Colombia (MinTIC, 2022–2025)

Universidad de Caldas - UTraining- MinTIC

Sara Medina Ruiz CC. 1004668000

Jefferson David Arteaga CC. 1089378880

Bootcamp Analítica de Datos Nivel Explorador

e-mail: medinasara825@gmail.com, jeffer.max1000@gmail.com

Febrero 22 de 2026

Pereira, Risaralda, Colombia

Resumen- El presente análisis se desarrolló a partir de un conjunto de datos del sector de Tecnologías de la Información y las Comunicaciones (TIC), enfocado en el estudio del acceso a internet fijo, las velocidades de conexión y las tecnologías asociadas. De manera previa, se realizó un proceso de limpieza y validación de la información, que incluyó la normalización del formato de las variables numéricas, el tratamiento de valores faltantes y la identificación de datos atípicos.

El análisis se orienta a identificar las zonas con mayor vulnerabilidad o menor nivel de acceso a internet, así como a evaluar la calidad del servicio a través de las velocidades de subida y bajada. Adicionalmente, se analizan las variaciones temporales de los accesos según los distintos segmentos del mercado, y se estudia la calidad del servicio ofrecido por los proveedores, junto con su participación en el mercado.

Palabras Claves- Acceso a internet fijo, Brecha digital, Colombia, Conectividad, Proveedores de internet, Tecnología, TIC, Velocidad de conexión

Abstract — The present analysis was conducted using a dataset from the Information and Communications Technology (ICT) sector, focused on the study of fixed internet access, connection speeds, and associated technologies. Prior to the analysis, a data cleaning and validation process was carried out, including the normalization of numerical variable formats, the treatment of missing values, and the identification of outliers.

The analysis aims to identify areas with greater vulnerability or lower levels of internet access, as well as to evaluate service quality through upload and download speeds. Additionally, temporal variations in access across different market segments are examined, and the quality of service provided by internet providers is analyzed along with their market share.

Keywords — Fixed internet access, Digital divide, Colombia, Connectivity, Internet service providers, Technology, ICT, Connection speed.

I. INTRODUCCION

La analítica de datos se ha consolidado como una herramienta fundamental para la toma de decisiones y el análisis de fenómenos complejos en diversos sectores, tanto en el ámbito académico como en el institucional. En particular, el análisis exploratorio de datos permite comprender el comportamiento de grandes volúmenes de información y extraer patrones relevantes que apoyan la formulación de diagnósticos y políticas públicas.

En este contexto, se desarrolla un proyecto de analítica de datos aplicado al sector de Tecnologías de la Información y las Comunicaciones (TIC), orientado al estudio del acceso a internet fijo, las velocidades de conexión y las tecnologías asociadas en Colombia. El análisis se llevó a cabo utilizando el lenguaje de programación Python, junto con librerías ampliamente utilizadas en ciencia de datos como Pandas, Numpy, Matplotlib y Seaborn, empleando Google Colab como entorno de trabajo para facilitar la ejecución y reproducibilidad del análisis.

El conjunto de datos original contaba con aproximadamente 2,1 millones de registros. A partir de un proceso de limpieza, validación y agrupación de la información, en el que se trataron valores faltantes, registros administrativos y datos atípicos, el conjunto de datos (dataset) final quedó conformado por cerca de 1,2 millones de registros, equivalentes a 52 millones de accesos fijos a internet, garantizando así una base de datos consistente y adecuada para el análisis.

Sobre este conjunto de datos depurado, se realizó un análisis exploratorio enfocado en la distribución de los accesos fijos a internet por proveedor, región y segmento, así como en la evaluación de la calidad del servicio a través de las velocidades de subida y bajada. El objetivo principal es identificar desigualdades territoriales, diferencias entre

proveedores y posibles brechas en el acceso y desempeño del servicio de internet fijo en el país.

A partir del estudio de los accesos fijos y las velocidades reportadas, se identifican diferencias relevantes entre segmentos y regiones, así como patrones de concentración de mercado y cambios en las preferencias tecnológicas. Este trabajo busca comprender cómo se distribuye la calidad del servicio, qué tecnologías lideran el mercado y cómo se comporta la oferta en términos territoriales y competitivos.}

II. CONTENIDO

Objetivo General- Aplicar técnicas de análisis exploratorio de datos en Python con el fin de analizar el estado general y la velocidad de los accesos fijos a internet en Colombia, diferenciando por departamento, proveedor y segmento comercial, y evaluando su evolución a lo largo del tiempo.

Objetivos Específicos-

1. Seleccionar el conjunto de datos del sector TIC [1] relacionada con los accesos fijos a internet, a partir de fuentes oficiales y confiables.
2. Realizar el proceso de limpieza y preparación de los datos, incluyendo el tratamiento de valores faltantes, atípicos e inconsistencias, para garantizar su calidad y fiabilidad.
3. Aplicar técnicas de análisis exploratorio de datos para identificar patrones, tendencias y diferencias en la calidad y el acceso a internet, considerando variables como departamento, proveedor y segmento comercial.
4. Resolver las preguntas de investigación a partir del entendimiento de la problemática abordada
5. Presentar los resultados del análisis de manera clara y efectiva, adaptando el contenido y el lenguaje al público objetivo, con el fin de cumplir los requisitos de evaluación y obtener la certificación del bootcamp.

Planteamiento del Problema:

El acceso a internet se ha consolidado como un servicio esencial para el desarrollo social, económico y educativo de los territorios, al constituirse en un medio fundamental para el acceso a la información, la generación de conocimiento, la productividad empresarial y la participación ciudadana. En el contexto colombiano, el **Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC)** produce y difunde estadísticas oficiales del sector TIC a partir del aprovechamiento de registros administrativos, con el propósito de garantizar información confiable, oportuna y comparable para la formulación y seguimiento de políticas públicas.

De acuerdo con [1], esta operación estadística tiene como objetivo general producir y difundir información sobre el mercado de los servicios de telefonía fija, telefonía móvil, internet fijo e internet móvil, en términos de accesos, ingresos, tráfico y cobertura, con desagregación nacional y territorial.

La información proviene principalmente del registro administrativo denominado Sistema de Información Único del Sector de Telecomunicaciones (SIUST), y se publica con periodicidad trimestral, garantizando cobertura para los 32 departamentos y 1121 municipios del país.

En este marco, el conjunto de datos seleccionado para el presente estudio corresponde a los microdatos de accesos fijos a internet publicados por el MinTIC en el portal de datos abiertos. Este dataset incluye variables relacionadas con número de accesos, proveedores, tecnologías de acceso (como fibra óptica, cable, DSL, entre otras), velocidades de descarga y carga, segmentos de mercado y desagregación geográfica a nivel departamental y municipal. La riqueza temática y la granularidad territorial de esta información permiten realizar análisis descriptivos y exploratorios que aportan evidencia empírica sobre la evolución del servicio de internet fijo en Colombia.

No obstante, a pesar de la disponibilidad de estos datos abiertos y su alineación con estándares internacionales propuestos por la Unión Internacional de Telecomunicaciones (UIT), persiste la necesidad de profundizar en el análisis del comportamiento territorial y tecnológico de los accesos fijos a internet. En particular, resulta relevante identificar:

- ¿Cómo ha evolucionado el número de accesos fijos a internet en el tiempo y cuáles son las tendencias de crecimiento o desaceleración?
- ¿Existen brechas territoriales significativas en términos de penetración del servicio entre departamentos?
- ¿Qué tecnologías predominan en la provisión del servicio y cómo se relacionan con las velocidades promedio ofrecidas?
- ¿Se evidencian diferencias en la oferta de velocidades según segmento de mercado o ubicación geográfica?
- ¿Cómo se distribuye el mercado entre proveedores y qué niveles de concentración pueden observarse?

Estas preguntas no pueden resolverse únicamente con afirmaciones descriptivas simples, sino que requieren la aplicación de técnicas de análisis exploratorio de datos, tratamiento de valores atípicos, análisis de tendencias temporales y comparación territorial mediante indicadores estandarizados. En consecuencia, el presente proyecto se orienta a realizar un análisis integral del comportamiento de los accesos fijos a internet en Colombia, utilizando herramientas de analítica de datos que permitan transformar la información estadística en conocimiento útil para la comprensión de la dinámica del sector TIC y sus implicaciones en el desarrollo territorial.

Metodología:

La presente investigación adopta un enfoque cuantitativo basado en analítica de datos aplicada a información estadística oficial del sector TIC en Colombia. El objetivo metodológico es transformar registros administrativos en evidencia analítica

que permita comprender la evolución, distribución territorial y composición tecnológica de los accesos fijos a internet.

La estrategia metodológica se estructura en etapas secuenciales que permiten descomponer el problema general en subproblemas.

La investigación utiliza como fuente principal los microdatos de accesos fijos a internet publicados por el *Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC)*, en el marco de la operación estadística *Estadísticas del Sector de las Tecnologías de la Información y las Comunicaciones*.

Se adopta un enfoque de Análisis Exploratorio de Datos (EDA) como metodología principal, dado que:

- El estudio busca identificar patrones, tendencias y brechas.
- No se persigue inicialmente inferencia causal ni modelamiento predictivo.
- El objetivo es comprender la estructura y comportamiento del fenómeno antes de aplicar modelos más complejos.

Etapas del proceso metodológico

Etapa 1. Entendimiento de los datos

Se realiza revisión de:

- Ficha metodológica oficial.
- Diccionario de datos.
- Estructura de los microdatos.
- Naturaleza de las variables (numéricas, categóricas, temporales, geográficas).

Esta etapa permitió comprender la lógica de construcción de indicadores y evitar errores de interpretación derivados del carácter administrativo de la información.

Etapa 2. Limpieza y preparación de datos

Se desarrollan las siguientes tareas:

- Identificación y tratamiento de valores nulos.
- Detección y análisis de valores atípicos en velocidades.
- Estandarización de nombres de tecnologías.
- Conversión de tipos de datos.
- Agregaciones por territorio y periodo.

Este proceso es fundamental en registros administrativos debido a posibles inconsistencias de reporte entre proveedores.

Etapa 3. Análisis descriptivo

El análisis se desarrolla en cuatro dimensiones:

- a) **Análisis temporal;** Tendencias de crecimiento.
- b) **Análisis territorial;** Comparación de accesos por departamento.
- c) **Análisis tecnológico;** Participación por tipo de tecnología, asociación entre tecnología y velocidad promedio.
- d) **Análisis de estructura de mercado**

- Número de proveedores.
- Participación relativa por proveedor.
- Exploración de concentración del mercado

Etapa 4. Visualización y comunicación

Se emplean técnicas de visualización de datos para facilitar la interpretación:

- Gráficos de líneas para tendencias.
- Gráficos de barras para participación tecnológica.
- Diagramas comparativos territoriales.
- Representaciones en millones para mejorar legibilidad.

La visualización se justifica como herramienta analítica y no únicamente estética, ya que permite identificar patrones no evidentes en tablas numéricas.

Herramientas utilizadas

El análisis se realizó mediante:

- Python como lenguaje de programación principal.
- Librerías:
 - pandas (manipulación de datos) [2]
 - numpy (operaciones numéricas) [3]
 - matplotlib (visualización) [4]
 - geopandas (cuando se requirió análisis territorial) [5]

Alcance y limitaciones metodológicas

- La investigación es de carácter descriptivo, no causal.
- Se trabaja con datos de oferta (proveedores), no con encuestas de demanda.
- La calidad de los resultados depende de la consistencia del reporte administrativo.

Estas limitaciones se reconocen para evitar sobreinterpretaciones.

Resultados obtenidos.

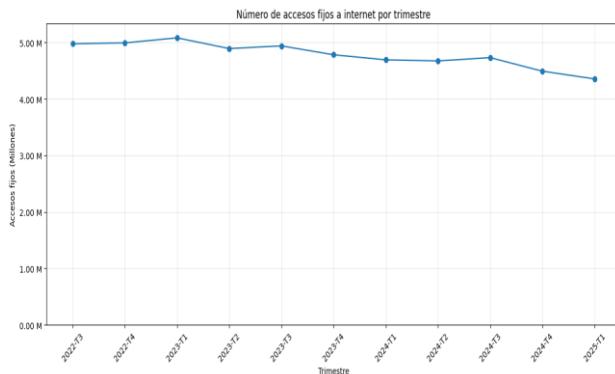
El desarrollo del proyecto se llevó a cabo siguiendo una metodología estructurada que incluyó las etapas de comprensión del dataset, limpieza y depuración de datos, construcción de variables derivadas, análisis exploratorio y generación de visualizaciones para la interpretación de resultados.

En una primera fase, se realizó la revisión y depuración de la base de datos, identificando valores atípicos en variables como velocidad de bajada y número de accesos fijos. Se aplicaron criterios técnicos diferenciados por tecnología para corregir o eliminar registros inconsistentes, garantizando así mayor confiabilidad en los análisis posteriores. Esta etapa permitió reducir sesgos en métricas agregadas y asegurar coherencia en los cálculos.

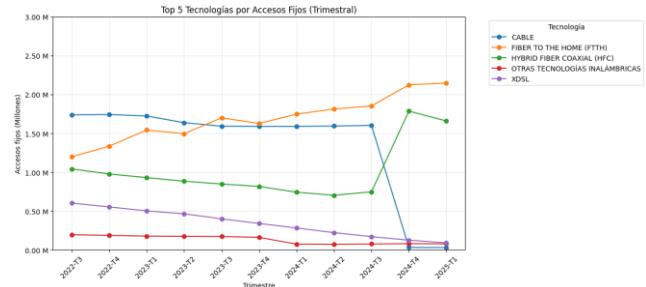
Posteriormente, se construyeron variables derivadas, como la velocidad promedio ponderada por número de accesos, con el fin de obtener indicadores más representativos del comportamiento real del mercado. La ponderación permitió evitar distorsiones generadas por proveedores con pocos accesos, pero velocidades elevadas, reflejando de manera más precisa la experiencia promedio de los usuarios.

En la fase de análisis exploratorio, se realizaron agregaciones por segmento, tecnología, proveedor y departamento. Los resultados evidenciaron diferencias significativas en velocidades entre segmentos, así como una concentración del mercado en tres principales proveedores. Asimismo, se identificaron brechas territoriales importantes, particularmente en departamentos con menor número de accesos, menor diversidad tecnológica y menor competencia entre proveedores.

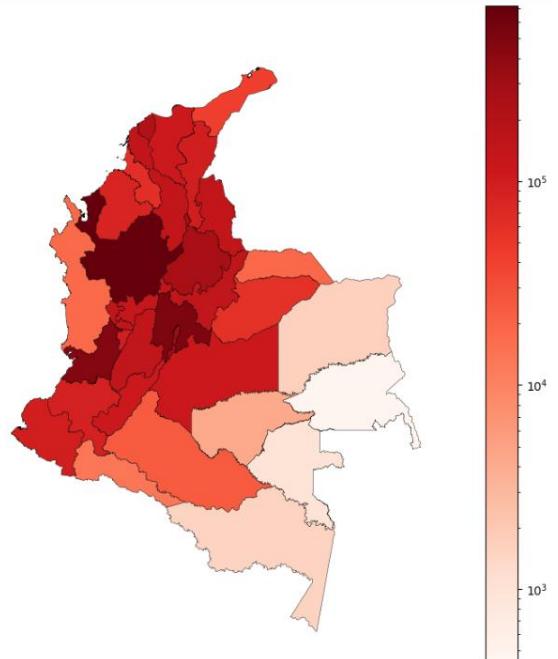
Las gráficas de evolución temporal mostraron una tendencia general decreciente en los accesos fijos a internet, mientras que los análisis por tecnología evidenciaron una mayor preferencia por soluciones como FTTH y HFC. Estos resultados sugieren un proceso de transición tecnológica y posibles cambios en las preferencias de los usuarios hacia otras modalidades de conexión.



Aunque en un primer análisis la reducción en el número de accesos fijos a internet podría interpretarse como un deterioro en la conectividad del país, esta conclusión sería apresurada si no se consideran dinámicas estructurales del mercado de telecomunicaciones. En particular, el fenómeno puede explicarse desde la perspectiva de la sustitución tecnológica, donde los hogares migran progresivamente de conexiones fijas tradicionales hacia soluciones móviles, impulsadas por la expansión de redes 4G y 5G, mayor cobertura, portabilidad y esquemas de precios más flexibles. En este sentido, la disminución de accesos fijos no necesariamente refleja una caída en el acceso total a internet, sino una transformación en los patrones de consumo y en las preferencias de los usuarios. Por tanto, el análisis integral de la conectividad en Colombia requiere incorporar variables asociadas al acceso móvil para comprender adecuadamente la evolución del mercado y evitar interpretaciones sesgadas basadas únicamente en conexiones fijas.



El comportamiento de la tecnología CABLE presenta una discontinuidad estructural a partir de 2024-T4, evidenciada por una caída abrupta en el número de accesos. Dado que esta reducción no sigue una tendencia progresiva, sino que ocurre de manera repentina, es probable que esté asociada a un cambio metodológico o a una reclasificación tecnológica en la fuente de datos, más que a una disminución real en la demanda. Esta hipótesis se refuerza al observar el crecimiento simultáneo de tecnologías como HFC, lo que sugiere un posible traslado de registros entre categorías.



Mapa de Calor de Número de accesos por departamento, en escala logarítmica del último trimestre reportado (2025-T1)

En conjunto, las tablas y visualizaciones generadas permitieron sintetizar los hallazgos clave del estudio, facilitando la comparación entre segmentos, regiones y tecnologías. Los diagramas del proceso metodológico y los gráficos incluidos en el informe respaldan de manera visual las conclusiones alcanzadas y evidencian la coherencia entre la metodología aplicada y los resultados obtenidos.

III. CONCLUSIONES

El desarrollo del presente proyecto permitió fortalecer competencias en analítica de datos aplicada a un contexto real, integrando procesos de limpieza, transformación, análisis exploratorio y visualización de información. Desde el punto de vista del proceso de enseñanza-aprendizaje, se consolidaron habilidades en el manejo de datos estructurados, interpretación de métricas como promedios ponderados y análisis comparativo entre segmentos, tecnologías y regiones. Si bien las bases teóricas permitieron abordar el proyecto de manera adecuada, se evidenció que un mayor dominio en técnicas de modelamiento estadístico y análisis inferencial habría enriquecido la profundidad de los resultados. El nivel de dificultad fue progresivo, especialmente en la etapa de depuración de datos y construcción de indicadores agregados, lo que representó un aprendizaje significativo.

En cuanto al uso de herramientas, el empleo de Python y librerías como Pandas y Matplotlib facilitó el procesamiento de grandes volúmenes de información y la generación de visualizaciones claras para el análisis. Como fortaleza, se destaca la capacidad de realizar transformaciones complejas y cálculos ponderados que permitieron obtener métricas más representativas del comportamiento real del mercado. Como debilidad, se identificó que algunos procesos iniciales de limpieza y validación de datos podrían haberse estructurado de manera más sistemática para evitar reprocesamientos posteriores. No obstante, el uso de estas herramientas permitió desarrollar un análisis integral y reproducible.

Respecto a los objetivos específicos del proyecto, se logró caracterizar la distribución de los accesos fijos a internet por segmento, identificando diferencias significativas en las velocidades promedio y evidenciando que los segmentos con mayor número de accesos no necesariamente presentan las mayores velocidades. Asimismo, se analizaron las brechas territoriales, encontrando que departamentos como Vaupés y Guainía presentan menores niveles de acceso, menor diversidad tecnológica y menor número de proveedores, lo que sugiere desigualdades en infraestructura y cobertura. En términos tecnológicos y de mercado, se confirmó una mayor preferencia por tecnologías como FTTH y HFC, así como una alta concentración en tres proveedores principales: Claro, Tigo y Movistar.

De manera general, el proyecto permitió comprender las dinámicas del mercado de internet fijo desde una perspectiva cuantitativa, evidenciando que la aparente tendencia decreciente en los accesos fijos no necesariamente implica una reducción en el acceso total a internet, sino que puede estar asociada a cambios en las preferencias hacia otras modalidades de conexión, no analizadas en el presente proyecto. En este sentido, la principal aplicación de los resultados radica en aportar evidencia para la toma de decisiones en materia de políticas de conectividad, análisis de mercado y planificación de infraestructura digital.

IV. RECOMENDACIONES

Futuros proyectos sobre las TIC en Colombia deberían ampliar el alcance del estudio incorporando enfoques más profundos que permitan no solo describir el comportamiento del servicio de internet, sino también explicar sus causas y proyectar su evolución. Para ello, sería pertinente integrar variables sociodemográficas y económicas, desarrollar análisis territoriales más detallados que permitan identificar brechas digitales estructurales, y evaluar la relación entre tipo de tecnología, calidad del servicio y competencia entre proveedores. Asimismo, se recomienda avanzar hacia estudios que analicen el impacto de la conectividad en el desarrollo regional, la educación y la productividad, de manera que la información publicada por el Ministerio de Tecnologías de la Información y las Comunicaciones pueda transformarse en insumo estratégico para la toma de decisiones y la formulación de políticas públicas orientadas al cierre de desigualdades digitales.

REFERENCIAS

- [1] M. TIC. “Datos TIC”. ONTIC. Accedido el 22 de febrero de 2026. [En línea]. Disponible: https://ontic.mintic.gov.co/857/w3-propertyvalue_916082.html#data=%7B%22orfilter%22%3D%22753959%22%2C%22page%22%3D%227D
- [2] Python. “User Guide — pandas 3.0.1 documentation”. pandas - Python Data Analysis Library. Accedido el 22 de febrero de 2026. [En línea]. Disponible: https://pandas.pydata.org/docs/user_guide/index.html
- [3] Python. “NumPy documentation — NumPy v2.4 Manual”. NumPy. Accedido el 22 de febrero de 2026. [En línea]. Disponible: <https://numpy.org/doc/stable/>
- [4] Python. “Matplotlib documentation — Matplotlib 3.10.8 documentation”. Matplotlib — Visualization with Python. Accedido el 22 de febrero de 2026. [En línea]. Disponible: <https://matplotlib.org/stable/index.html>
- [5] Python. “Documentation — GeoPandas 1.1.2+0.g81214bf.dirty documentation”. GeoPandas 1.1.2 — GeoPandas 1.1.2+0.g81214bf.dirty documentation. Accedido el 22 de febrero de 2026. [En línea]. Disponible: <https://geopandas.org/en/stable/docs.html>

ANEXOS

- Anexo 1. Limpieza TIC - Proyecto Análisis de Datos.ipynb
- Anexo 2. Análisis descriptivo TIC - Proyecto Análisis de Datos.ipynb
- Anexo 3. Presentación