



TRABAJO DE GRADO
Opción Seminario-Diplomado.

**ANÁLISIS DE DATOS DE PENETRACIÓN DE INTERNET FIJA Y MÓVIL EN
COLOMBIA**

Corporación Universitaria Remington.
Nombre de la facultad: Facultad de Ingeniería
Nombre del programa académico: Ingeniería de Sistemas

Samuel Giovanni Colmenares Patiño
Nombre del Tutor del trabajo de grado: John Fredy Mira Mejía.
Opción de Trabajo de grado Seminario-Diplomado.
2024.

Dedicatoria

A mi madre que siempre me observa desde arriba, a mi padre que siempre me dio ánimo cuando todo parecía fallar y a mi tía que siempre creyó en mí....

Tabla de Contenidos

Resumen.....	4
Marco conceptual y contextual	5
Marco contextual	5
Marco conceptual.....	6
Datos Abiertos	6
Power Query	6
Power BI	6
Proceso de datos	6
Desarrollo e implementación del aprendizaje.....	8
Definición de Objetivos	8
Fuentes de información.....	8
Depuración y Organización de datos	10
Depuración Tabla móvil	10
Depuración Tabla internet fijo	11
Análisis e interpretación de resultados	11
Interpretación general	11
Interpretación ininternet móvil	12
Interpretación internet fijo	12
Interpretación general	12
Referencias.....	15

Resumen

En este trabajo de grado analizaremos el alcance que tiene internet en Colombia en los últimos años, tanto para conexiones fijas como móvil, es discriminando por departamentos y comparándolo contra la población de cada año, entre otros. Con esta información pretendo determinar que tan cerca está la tecnología, no solamente de ciudades capitales conocidas, sino de aquellos lugares apartados que, debido a nuestra particular geografía, representan un desafío considerable llevar la información a los colombianos.

Para obtener los datos usados, recurrí a el sitio “Datos abiertos”, un espacio de dominio público patrocinado por el Mintic, de allí usamos dos conjuntos de datos diferentes, uno para internet fijo y otro para internet móvil. Para realizar el modelado de la data, utilicé Power Query de Excel, mientras que para la generación de gráficas decidí usar Power BI de Microsoft por su versatilidad capacidad de personalización de los gráficos y opciones de filtrado.

Las conclusiones están orientadas a entender el crecimiento de las conexiones, su índice de penetración y demostrar la desigualdad de conectividad en algunas zonas del país.

Palabras clave

Índice de penetración: Este indicador nos ayuda a conocer el nivel de adopción del servicio de internet fijo y móvil en un periodo de tiempo determinado. Se calcula con la siguiente ecuación:

$$\frac{\text{Número de accesos}}{\text{Población}} \times 100$$

Proveedor: Hace referencia a la empresa prestadora del servicio de internet, sea móvil o fijo; también conocido como operador.

Accesos de internet: Nos indica el número de accesos reportados por el proveedor como aquellos que se comercializaron a sus clientes.

Población: Término que hace referencia a un grupo de habitantes que habitan en un espacio geográfico, ya sea un departamento, municipio o centro poblado

Hogar: Agrupación de uno o más habitantes que residen en la misma casa.

Marco conceptual y contextual

Marco contextual

El internet en Colombia presenta un panorama diverso y controversial, reflejando una alta brecha digital entre las ciudades principales y las zonas rurales. Aparte de la conectividad, encontramos también una gran diferencia en las velocidades de conexión.

Esto refleja las diferentes dinámicas tecnológicas, socioeconómicas y de ubicación geográfica de cada municipio.

En medio de este panorama diferencial, es importante conocer el crecimiento de adopción de este servicio público esencial para crear estrategias que permitan cerrar la brecha de alcance a las tecnologías de la información y como desde nuestro futuro rol como ingenieros, podemos aportar a dicho crecimiento. Este análisis está enfocado en la penetración del internet en Colombia y como esto impacta en la velocidad de conexión, donde hay una gran variación en la calidad del servicio, un ejemplo de ello es que según la encuesta de tecnologías de la información (ENTIC) del DANE un 68,4% de los hogares consideran buena su conexión a internet fijo; en cambio no están muy felices con las velocidades de conexión donde se encontraban con una disponibilidad de entre 2 Mbps a 10 Mbps, lo cual indica una necesidad de mejora en la infraestructura de red. Para el caso de las conexiones móviles encontramos panoramas más desoladores, donde el alcance de conexión está sólo disponible en las cabeceras municipales, pero en las zonas rurales apartadas no llega la conectividad o cuenta con un alcance de tecnología de 2G y pocas veces en 3G, alcanzando velocidades entre 9,6 Kbps y 1,8 Mbps, que para la disposición actual de información es muy baja y representa más un inconveniente que una ventaja.

Para este análisis consideré diferentes aspectos aportados por los conjuntos de datos encontrados en el portal Datos Abiertos del gobierno nacional, aquellos puntos clave fueron:

- Accesos a internet
- Penetración por departamento
- Penetración por municipio
- Tecnologías (Para el caso de móviles)
- Desafíos y oportunidades

Tuvimos en cuenta el crecimiento de accesos a internet fijo desde el año 2015 al 2022, contrastado con el crecimiento poblacional y para el caso móvil el crecimiento entre los años 2020 y 2022 en cada uno de los proveedores de estos servicios, contrastado también con el crecimiento poblacional del país.

Marco conceptual

En la creación de este análisis sobre la penetración del internet fijo y móvil en Colombia, inicié con la recolección de datos, obtenidos de la fuente más confiable que conozco para este tipo de proyectos: Datos Abiertos.

Datos Abiertos

Es la información pública del Ministerio de las Tecnologías de la Información y las Comunicaciones, esta información se deja al alcance de todas las personas en diferentes formatos que permiten su análisis y modelado. La razón por la que el Estado pone a disposición esta información obedece a las iniciativas del gobierno para promover la participación ciudadana y la transparencia. El lugar donde se encuentra disponible es www.datos.gov.co.

Existen tres tipos de datos en datos abiertos:

- **Primarios:** Es el estado inicial de los datos, no se han procesado, se denominan datos crudos de la fuente de origen.
- **Interoperables:** Es el estado procesado inicial donde se facilita su acceso.
- **Libres – sin restricciones:** Estado en el cual ya se encuentran procesados y disponibles para que cualquier persona haga uso de ellos.

Contar con esta información es una gran ventaja, porque contar con data nos ofrece un panorama real de los eventos del país y nos permite tomar decisiones acordes a los resultados.

Power Query

Es una herramienta con una interfaz intuitiva de Power BI y Excel creada por Microsoft, que sirve para importar, transformar y analizar datos desde diferentes fuentes; con esta herramienta es posible conectarse con varias fuentes de datos, sea de manera individual o al mismo tiempo, como por ejemplo bases de datos, archivos de texto, Excel, API's de internet, entre otros. También permite la transformación de datos con diferentes operaciones como lo son: Ordenar datos, filtrar, eliminar columnas, crear columnas calculadas, cambiar tipos de datos, entre otros. Luego de la transformación nos ayuda a cargar los datos en Excel o Power BI para analizarlos o generar reportes.

Power BI

Es un compendio de aplicaciones, servicios y conectores que tienen como objetivo transformar los datos recibidos de diferentes fuentes convirtiéndolos en datos coherentes adaptados para ser representados por gráficas y reportes dinámicos. Es posible acceder a Power BI de diferentes formas: Power BI Desktop, Power BI Online (SaaS) y Power BI Mobile. Los reportes creados pueden ser publicados en un servidor local (Power BI Server) o en la nube.

Proceso de datos

Realizar el proceso de los datos es fundamental en cualquier análisis, debido a que nos ofrece la materia prima para establecer hipótesis, medir y tomar decisiones. Normalmente conta de siete pasos:

1. **Definición de objetivos:** Es el paso cero del proceso, aquí debemos definir los objetivos del análisis y sus límites, para orientar la recolección por el camino deseado. En mi caso tomé la decisión de buscar datos sobre el internet en Colombia y el alcance que este tiene.
2. **Identificar fuentes de información:** En este paso se debe investigar y encontrar fuentes de datos fiables ya sean datos originales (datos obtenidos de la fuente) o datos secundarios (datos compilados). En mi caso use datos secundarios obtenidos de datos abiertos.
3. **Estrategia de recopilación:** Aquí se determinan la forma en que se obtendrán los datos, por ejemplo: entrevistas, encuestas, observación, etc. En mi caso no aplica este paso.
4. **Depuración y organización de datos:** En este paso se organizan y depuran los datos acordes a los KPI que se propusieron en la estrategia inicial. En mi caso usé *Power Query* de *Excel* y *Power BI* para modelar la data que pensaba representar mediante gráficos.
5. **Análisis de datos:** Mediante diferentes técnicas estadísticas se interpreta la información, ya sea de manera manual o con el uso de herramientas de análisis y con esto se obtienen conclusiones significativas.
6. **Interpretar resultados:** Entender y comprender los resultados en relación con los KPI's propuestos en el diseño.
7. **Aplicar resultados:** A partir de la interpretación, tomamos los hallazgos para próximas investigaciones o se utilizan para tomar decisiones.

Desarrollo e implementación del aprendizaje

El internet pasó de ser lujo y entretenimiento a ser la fuente de información principal del momento, convirtiéndose en una necesidad y en un servicio básico, ya que no tenerlo vulnera el derecho fundamental a “Estar informado”. Esta es una de las motivaciones principales para escoger el internet como temática principal y enlazarlo con la penetración que tiene y ha tenido durante los últimos años.

Como expresé en párrafos anteriores, traté de seguir siete pasos fundamentales en el procesamiento de datos, integrando lo aprendido en los cursos virtuales, los cuales aportaron un acercamiento a diferentes tecnologías que había escuchado en varias ocasiones pero que no me había atrevido a estudiar ni practicar; con este aprendizaje en mente, a continuación, los detalles de mi análisis:

Definición de Objetivos

Como mencioné anteriormente, mi intención es analizar la penetración del internet tanto fijo como móvil; por tanto, adapto como KPI principal el *índice de penetración*, esto quiere decir que mis fuentes de datos deben contar con este valor o tener los insumos necesarios para calcularla. Otro KPI's para tener en cuenta son: cantidad de accesos o conexiones o accesos, así como la población, esta última, debería contar con por lo menos los departamentos, unidad poblacional de Colombia que tiene una gran brecha tecnológica y económica.

Otro de los objetivos importantes es entender el crecimiento de las conexiones que ha tenido el internet en Colombia basado en los datos mencionados anteriormente.

Fuentes de información

Como se menciona en puntos anteriores, mi fuente de datos fue Datos Abiertos, porque considero que tiene una muestra importante, con lo cual podré generar resultados más acertados; ya que, a mayor tamaño de la muestra, más cerca de la realidad estaré. Adicional, datos abiertos cuenta con accesos directos a datos poblacionales reales obtenidos de censos estatales y toda esta información se encuentra pre-modelada y en un formato fácil de trabajar, en este caso opté por formatos CSV por su adaptabilidad y por ser reconocida por diferentes lenguajes programáticos y herramientas de análisis y modelado.

Luego de un recorrido por el sitio de Datos Abiertos, seleccioné tres grupos de datos que se adaptaban a las necesidades descritas en mis objetivos y contaban con la data planeada, estos fueron:

1. [Internet Fijo Penetración Municipio | Datos Abiertos Colombia](#) : Este conjunto de datos cuenta con 35.600 registros y cuenta con información de:

- a. Año
- b. Trimestre
- c. Departamento
- d. Municipio
- e. Cantidad Accesos fijos
- f. Población DANE
- g. Índice

Como se puede observar, cuenta con la información necesaria para completar los KPI's propuestos en el diseño y cuenta con gran cantidad de datos, por tanto es la indicada.

2. [Cobertura móvil por tecnología, departamento y municipio por proveedor | Datos Abiertos Colombia](#) : Este conjunto de datos cuenta con 407.000 registros y columnas de datos que son:
 - a. Año
 - b. Trimestre
 - c. Proveedor
 - d. Departamento
 - e. Municipio
 - f. Cabecera municipal
 - g. Centro Poblado
 - h. Cobertura 2G
 - i. Cobertura 3G
 - j. Cobertura HSPA+, HSPA+DC
 - k. Cobertura 4G
 - l. Cobertura LTE
 - m. Cobertura 5G

Este conjunto de datos tiene una gran cantidad de registros y variables que pueden ofrecer tableros enriquecidos y con un gran acercamiento a la realidad.

3. [Internet Móvil suscriptores por proveedor | Datos Abiertos Colombia](#) : Este conjunto de datos orientado al internet móvil un poco más pequeño cuenta con 90 registros y cuenta con información de:
 - a. Año
 - b. Proveedor
 - c. Trimestre
 - d. Suscriptores
 - e. Población DANE
 - f. Índice de penetración

Este conjunto cuenta con menos columnas que los anteriores, pero cuenta con información precisa de la penetración de internet móvil.

Con los conjuntos definidos, decidí tomar solo dos conjuntos de datos: el 1 y el 3; descarté el número 2 debido a que aunque tenía una gran cantidad de datos de poblaciones y tecnologías móviles, no contaba con el índice de penetración ni con la población de las poblaciones específicas, impidiendo obtener el KPI principal que es el índice de penetración; adicional, las tecnologías móviles sólo informaban si el centro poblado contaba o no con ellas, dificultando el modelado y brindando información poco valiosa para el análisis propuesto.

Depuración y Organización de datos

Como se mencionó en el marco conceptual, la data fue modelada con Power Query de Excel, comparto los enlaces de las tablas procesadas porque por su tamaño quedarían fuera de proporción incluirlas en este documento, las URL son:

- [1. Tabla filtrada internet fijo](#)
- [2. Tabla filtrada movil tecnologia](#)
- [3. Tabla filtrada movil definitiva](#)

Para tener un primer acercamiento a cada una de las tablas procesadas de la siguiente manera:

Depuración Tabla móvil

Para generar esta tabla usamos el conjunto de datos número 3, allí aplicamos la siguiente lista de cambios:

- Renombramiento de proveedores por su nombre comercial
- Renombramiento de encabezado de columnas
- Eliminación de campos vacíos o en cero
- Filtrado de columna trimestre: Dado a que solo quería valores por año, dejé los valores únicamente del último trimestre, ya que son acumulativos.

Podemos observar como quedó la tabla ya modelada en la siguiente imagen:

Tabla 1 - Tabla de móvil modelada

✕ ✓ fx = Table.RenameColumns("#Filtered Rows",{"POBLACION DANE", "POBLACION"}, {"ÍNDICE DE PENETRACIÓN", "ÍNDICE"})

	A ⁰ PROVEEDOR	1 ² 3 AÑO	1 ² 3 TRIMESTRE	1 ² 3 SUSCRITORES	1 ² 3 POBLACION	1.2 ÍNDICE
1	AVANTEL	2021	4	33021	51117378	0,06
2	AVANTEL	2020	4	410262	50407647	0,81
3	CLARO	2022	4	9229304	51682692	17,86
4	CLARO	2020	4	7359734	50407647	14,6
5	CLARO	2021	4	8309469	51117378	16,26
6	ETB	2022	4	115911	51682692	0,22
7	ETB	2020	4	147328	50407647	0,29
8	ETB	2021	4	89828	51117378	0,18
9	BAQUETAD	2021	4	4011270	51117378	0,07

Depuración Tabla internet fijo

Para generar esta tabla usamos el conjunto de datos número 1, allí aplicamos la siguiente lista de cambios:

- Renombramiento de encabezado de columnas
- Renombramiento de Departamento “San Andrés”
- Eliminación columna código departamento
- Eliminación columna código municipio
- Filtrado de columna trimestre: Dado a que solo quería valores por año, dejé los valores únicamente del último trimestre, ya que son acumulativos.

Podemos observar como quedó la tabla ya modelada en la siguiente imagen:

Tabla 2 - Tabla filtrada internet fijo

	AÑO	TRIMESTRE	DEPARTAMENTO	MUNICIPIO	No. ACCESOS FIJOS A INTERNET	POBLACIÓN DANE	INDICE
1	2022		4 CHOCÓ	CÉRTEGUI	179	5967	3
2	2021		4 BOLÍVAR	CALAMAR	564	24034	2,35
3	2022		4 BOYACÁ	SUSACÓN	8	2798	0,29
4	2022		4 CAUCA	PUERTO TEIADA	3925	43624	9
5	2021		4 BOLÍVAR	PINILLOS	47	25814	0,18
6	2021		4 BOYACÁ	TUTA	584	8641	6,76
7	2022		4 VALLE DEL CAUCA	GUACARÍ	3038	35239	8,62
8	2022		4 BOLÍVAR	SANTA ROSA	345	22566	1,53
9	2021		4 CAUCA	LA VEGA	598	25455	2,35
10	2021		4 ANTIOQUIA	YARUMAL	3859	43308	8,91
11	2022		4 CAUCA	GUAPI	293	28878	1,01
12	2022		4 CUNDINAMARCA	GACHETÁ	460	9279	4,96
13	2022		4 ANTIOQUIA	BETULIA	2484	16320	15,22
14	2022		4 ANTIOQUIA	SAN CARLOS	1929	16687	11,56

Análisis e interpretación de resultados

Para realizar el análisis y la interpretación se realizaron varios gráficos y tablas representados en dos dashboards, uno para móvil y otro para fijo, construidos en Power BI. Dadas las características de este documento es complicado mostrar todas las interacciones con los dashboards de donde se obtuvieron los insights, aunque decidí publicarlo en línea para que cualquier persona que lea este análisis tenga la oportunidad de interactuar con ellos, el link es:

[Dashboards móvil y fijo - público](#)

Interpretación general

A nivel general, podemos interpretar lo siguiente:

- Lo primero que podemos observar es que la penetración del internet en Colombia ha aumentado a buen ritmo en los últimos años, aunque el crecimiento es más notable en departamentos como **Antioquia, Valle de Cauca y Bogotá D.C.** (Para datos se considera departamento la ciudad capital del país)
- El año con mayor índice de penetración tanto en fijo y móvil es el **2022**.
- El crecimiento de las conexiones aumentó a buen ritmo pero de manera **lineal**.

Interpretación ininternet móvil

A nivel de internet móvil, podemos interpretar lo siguiente:

- Claro es el proveedor con mayor número de conexiones, cubre entre el **48,84%** y el **54,28%** del total de usuarios a nivel nacional, el resto está repartido entre los demás operadores.
- La mayor pérdida de usuarios la tuvo Avantel, quien entre **2020** y **2021** perdió **400.000** usuarios.
- El proveedor que tuvo el crecimiento más exponencial fue **WOM** quien entre **2021** y **2022** creció en un **172,6%**.
- El mayor índice de penetración lo obtuvo **Claro** en el **2022** con el **17,86**.
- El mayor crecimiento de usuarios móviles se obtuvo entre **2021** y **2022** con un crecimiento de más de **3 millones** de **usuarios nuevos**.

Interpretación internet fijo

A nivel de internet fijo, podemos interpretar lo siguiente:

- El año con mayor crecimiento de accesos fue el **2020**, con un incremento de **862.973 nuevas conexiones**.
- Los municipios con mayor cantidad de accesos son:
 - Bogotá D.C. 2,27 Millones
 - Medellín 769.787
 - Cali 548.236
- Los departamentos con menos cantidad de conexiones son:
 - Amazonas 1.005
 - Guanía 818
 - Vaupés 105
- Los **5** departamentos con menor cantidad de conexiones, fueron los únicos en decrecer conexiones entre **2020** y **2022**.

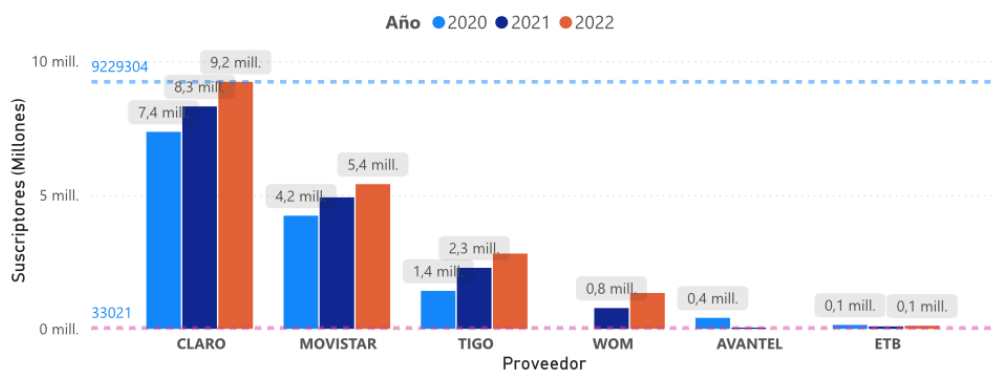
Interpretación general

A continuación los dashboards creados:

Figure 1- Dashboard Internet móvil

Penetración Internet en Colombia - Móvil

Suscriptores por Proveedor al año

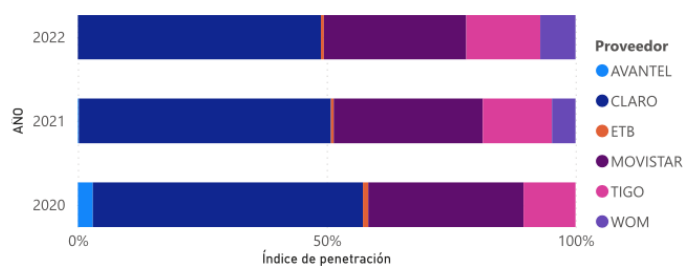


5,97

Índice de penetración

Año	Población DANE
2020	50407647
2021	51117378
2022	51682692

Porcentaje de Penetración por Proveedor



Alcance de Internet en la población

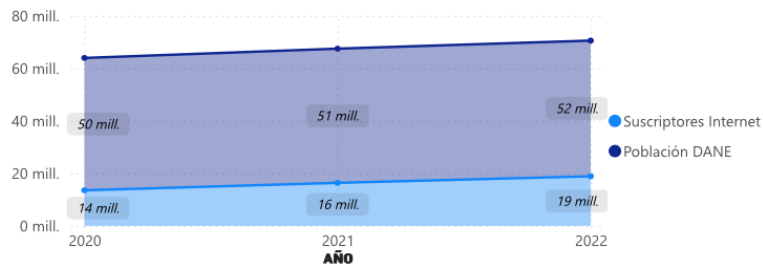


Figure 2- Dashboard Internet Fijo

Penetración Internet en Colombia -Fijo



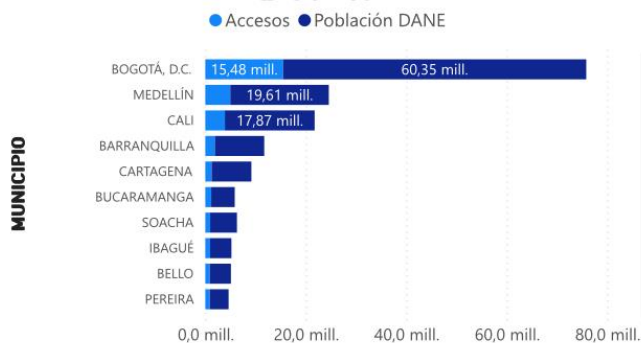
4,29

Índice de penetración

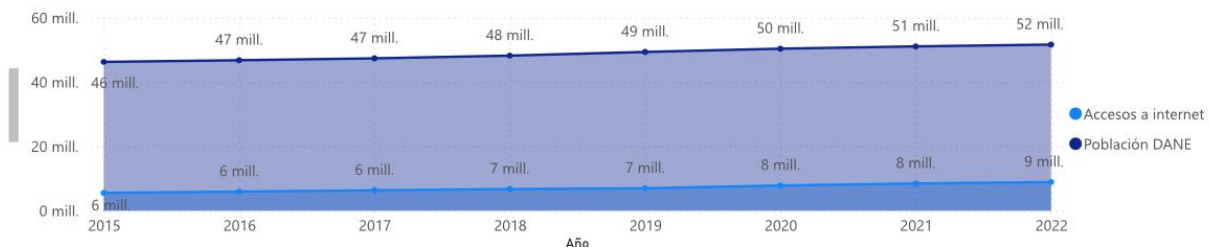
Accesos por Departamento al año

DEPARTAMENTO	2015	2016	2017	2018	2019	2020	2021	2022
BOGOTÁ D.C.	1610511	1728459	1818094	1876271	1916910	2088680	2171274	2269823
ANTIOQUIA	957544	1033038	1098914	1164322	1232910	1364120	1512907	1645915
VALLE DEL CAUCA	609043	657524	707051	749987	782539	864090	892963	918562
CUNDINAMARCA	290582	318243	342365	391959	417242	517323	573571	612263
ATLÁNTICO	285611	299204	328389	346822	354675	392527	426313	471543
SANTANDER	293905	311546	329803	324351	328405	366945	405974	429054
BOLÍVAR	158599	165240	177184	186967	185788	197100	234257	265887
RISARALDA	138897	148983	149943	168620	178877	211170	222026	220933
TOLIMA	137491	143181	147084	154721	152485	180695	200147	214413
NORTE DE SANTANDER	123841	121940	129802	143285	154898	181696	198306	185060
Total	5551745	5940639	6353018	6736079	6985051	7848024	8470457	8910502

Top Accesos por Municipio VS



Cubrimiento acceso a internet



Referencias

- Alexandra, M. G. D., Johana, O. A. L., Yudis, V. S., & Anderson, M. S. J. (2022). *Análisis sobre la penetración de internet en Colombia basados en el informe de primer semestre de 2022 del MINTIC*. Repositorio Institucional Compensar. <https://repositoriocrai.ucompensar.edu.co/entities/publication/16f8ce4a-e217-47d6-91ee-74d479b3b223>
- Bolaños, D. G. (2024, 12 enero). ¿Cuál es la velocidad máxima de subida y descarga en 2G, 3G y 4G? *MovilZona*. <https://www.movilzona.es/2016/02/17/cual-es-la-velocidad-maxima-de-subida-y-descarga-en-2g-3g-y-4g/>
- DataScope. (2023, 14 septiembre). *¿Cómo realizar una recolección de datos efectiva?* DataScope. <https://datascope.io/es/blog/releccion-de-datos/>
- Datos abiertos*. (s. f.-a). <https://gobiernodigital.mintic.gov.co/portal/Iniciativas/Datos-abiertos/>
- Datos abiertos*. (s. f.-b). <https://mintic.gov.co/portal/inicio/Glosario/D/5300:Datos-Abiertos>
- Internet fijo penetración municipio*. (s. f.). <https://www.opendatanetwork.com/dataset/www.datos.gov.co/fut2-keu8>
- Ministerio de las Tecnologías de la Información y las comunicaciones [MINTIC]. (s. f.). *Boletín trimestral de las TIC: Cifras cuarto trimestre de 2021* (1.^a ed.). https://colombiatic.mintic.gov.co/679/articles-209445_archivo_pdf.pdf
- Ptyx507x. (2024, 24 enero). *¿Qué es Power Query? - Power Query*. Microsoft Learn. <https://learn.microsoft.com/es-es/power-query/power-query-what-is-power-query>
- Westreicher, G. (2022, 24 noviembre). *Población*. Economipedia. <https://economipedia.com/definiciones/poblacion.html>

