

Entregable final - Reto 1

DataJamp CRC 2021

Equipo: LAB101 - Universidad Nacional de Colombia

Integrantes: Ana María Barbosa Cárdenas
Carlos Alberto Galindo Rojas
Jhonatan Alejandro Gordillo Silva

1. Propuesta:

Creación de índices de acceso a servicios de telecomunicaciones (internet fijo, telefonía fija y televisión) en cada municipio para presentar la evolución en el tiempo. Así mismo, clasificar el comportamiento del índice en cada municipio para aproximarse a la homogeneidad o heterogeneidad entre el acceso a las TICs para los distintos municipios del país. Finalmente, se presenta un indicador de corte transversal para identificar el porcentaje de cobertura del servicio sobre el total de la población en el último período del que se tiene información para cada tipo de servicio.

2. Justificación:

La llegada de la pandemia por el COVID-19 y el aislamiento obligatorio, reflejó la falta de cobertura en tecnologías de la información y la comunicación en algunas regiones del país. Sin embargo, dada la falta de índices técnicos y precisos que hagan seguimiento a este dato, se hace difícil tomar decisiones de política pública que permitan avanzar en materia de conectividad. Por ello, es indispensable construir indicadores de cobertura a las TICs, que permitan mostrar la evolución de las regiones y poder clasificarlas para enfocar las decisiones de política pública en función a mejorar el acceso a las telecomunicaciones en los lugares más apartados.

3. Desarrollo:

Se usa como insumo para la creación de índices las siguientes bases de datos:

a. Cantidad de accesos a internet fijo:

Describe la cantidad de accesos a internet existentes en cada municipio. Se dispone de información con periodicidad trimestral desde el año 2010 hasta el 2020.

b. Líneas de telefonía local:

Informa sobre la cantidad de líneas conectadas y en funcionamiento. Se posee información con periodicidad trimestral desde el año 2003 hasta el 2020.

c. Suscriptores de TV por suscripción:

Contiene la información de suscriptores de televisión por suscripción. Se presenta información con periodicidad trimestral desde el año 2010 hasta el 2020.

d. Población DANE:

Contiene la información poblacional de cada uno de los municipios, con códigos DANE del último censo Nacional 2018.

Es importante resaltar que toda la programación y tratamiento de los datos se realizó con ayuda del software Rstudio.

4. Procedimiento:

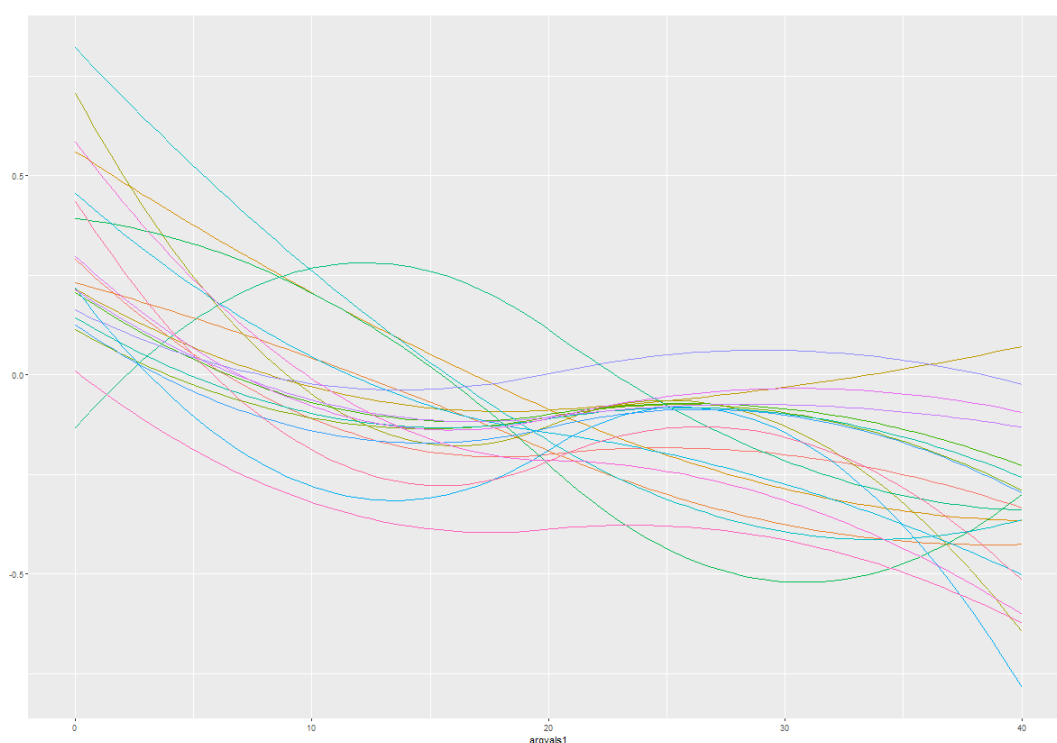
a. Creación de índices:

- i. Índices de cobertura en el tiempo: se crearon índices de cobertura para el servicio de telefonía fija, internet fijo y televisión para cada municipio del país en el que se tenía información disponible. El índice representa la variación anual para mostrar la evolución de la cobertura en cada municipio en el tiempo disponible de la base de datos.
- ii. Índices de corte transversal: teniendo en cuenta el estado actual (último trimestre del 2020) de los accesos a internet fijo, televisión y telefonía fija, se tomó la población de cada municipio y se construyó el porcentaje de accesos en función a su población.
Luego de crear el índice, a través de la distribución del mismo señalamos los percentiles claves para clasificar los comportamientos del índice en: “bajo”, “normal” y “alto”. Finalmente, contamos en cada departamento cuántos municipios están clasificados en cada uno de los comportamientos y esos porcentajes se vieron reflejados en el producto final que son los mapas de análisis.
- iii. Datos funcionales: dado que se tienen índices para cada municipio a lo largo del tiempo, se decide trabajar estos datos como datos funcionales, que consiste en analizar las curvas generados por los diferentes valores que toma el índice a lo largo del tiempo y de esta forma resumir el comportamiento de cada municipio de una manera práctica y clasificar los municipios según el grado de homogeneidad.

b. Clasificación:

Para la clasificación se consideraron dos escenarios, el primero consiste en tomar a todos los municipios de Colombia y agruparlos independientemente del departamento al que pertenecieran, esta fue una primera clasificación a nivel nacional. El otro escenario consiste en tratar cada departamento como una población y dentro de esta población agrupar los municipios, esto fue una clasificación a nivel departamental, con el fin de comparar el comportamiento de los diferentes municipios a nivel nacional y el comportamiento que presentan dentro de sus departamentos. Así mismo, permite conocer la homogeneidad que pueda existir dentro de los departamentos y la heterogeneidad que pueda existir entre los departamentos.

A Continuación se presentan las curvas del comportamiento del índice de telefonía fija en el departamento de Córdoba para cada uno de los municipios:

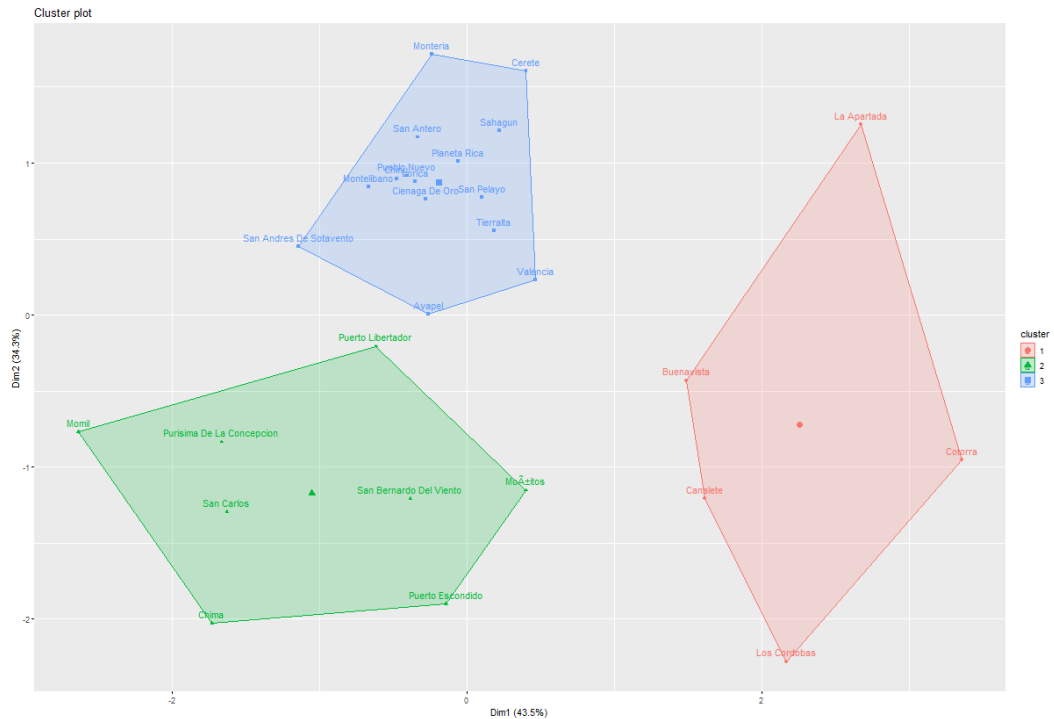


Fuente: Elaboración propia

Es importante resaltar, que se dispone de todos los índices a lo largo del tiempo para cada municipio de los diferentes departamentos para cada tipo de servicio (telefonía fija, internet fijo y televisión), pero por efectos de presentación se eligió el departamento de Córdoba para su ilustración.

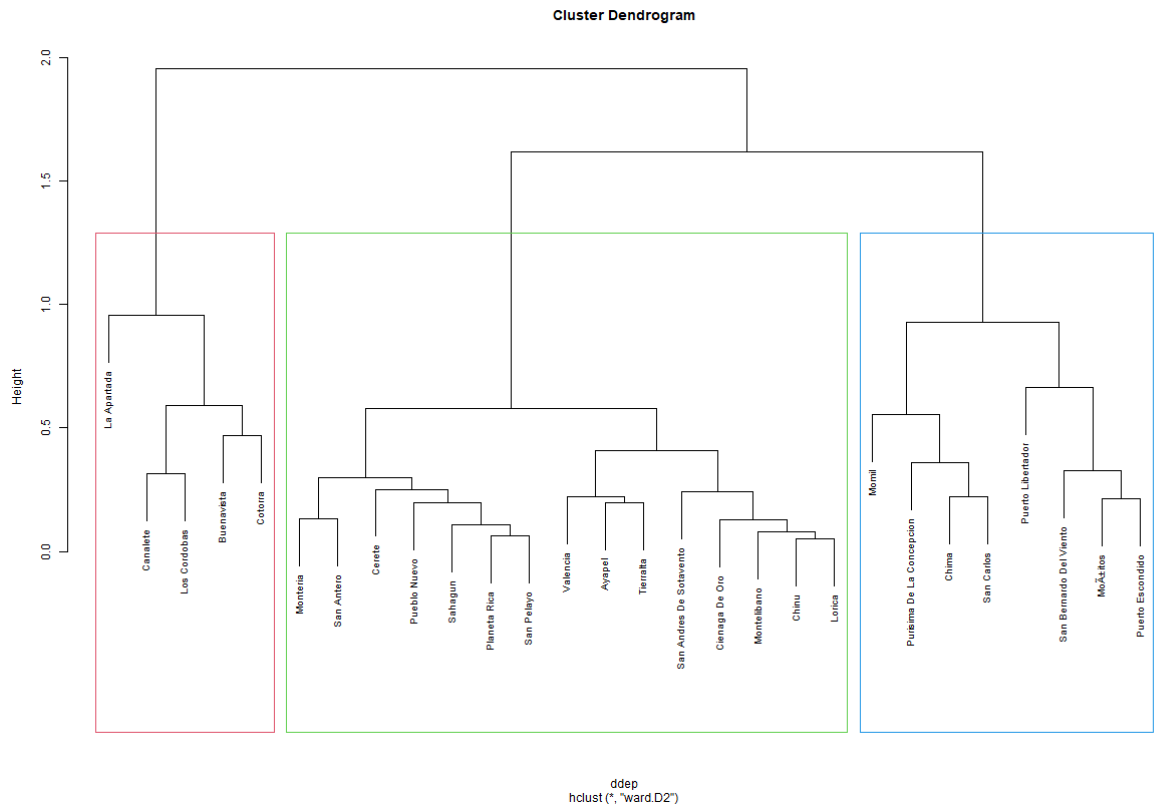
Para poder agrupar estas curvas según su comportamiento, se hizo uso de los datos funcionales, los cuales nos permiten tratar cada curva como un objeto multivariado y posteriormente poder hacer análisis sobre estos datos, una vez se tienen las curvas resumidas en datos multivariados, hacemos uso de el

metodo de ACP (Análisis de Componentes Principales) el cual nos permite visualizar una agrupación de los objetos multivariados. A continuación se muestra un clúster que se desarrolló para el departamento de Córdoba (usando la metodología K-means) agrupando las curvas que describen el comportamiento del índice de telefonía fija a lo largo del tiempo para cada municipio:



Fuente: Elaboración propia

Otra manera de visualizar los grupos, por departamento es por medio de dendrogramas. Estos nos permiten conocer los municipios que presentan diferencias marcadas y los posibles municipios más homogéneos que se encuentran dentro de cada departamento. A continuación se muestra para el departamento de Córdoba el respectivo dendrograma (usando la metodología K-means) que agrupa los municipios por comportamientos similares según el índice de telefonía fija a lo largo de tiempo:



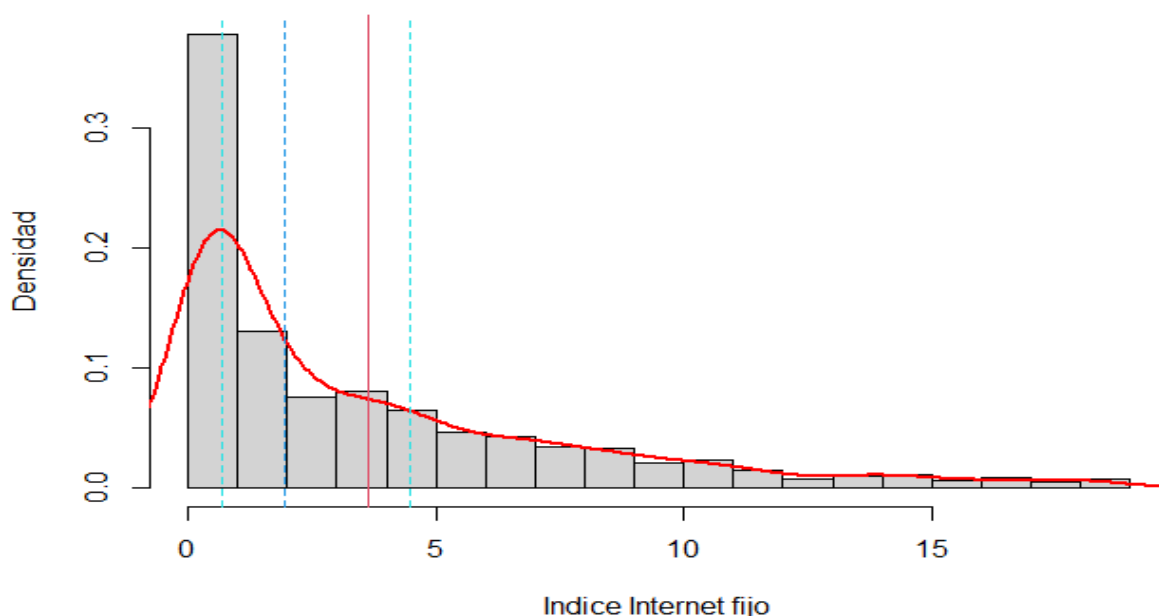
Fuente: Elaboración propia

5.Resultados:

5.1 Índice de Internet fijo:

Para obtener el índice municipal de la cantidad de accesos en función a la población ir al anexo 5.1. El primer paso que se hizo para analizar el comportamiento del índice fue recortar al percentil 95, para obtener una idea más acertada del comportamiento general.

Histograma y densidad



Fuente: Elaboración propia

Lo que observamos en la anterior gráfica es que existe una acumulación muy pronunciada en porcentajes muy bajos, es decir, que la mayoría de municipios tienen accesos muy bajos en comparación a la población total, por otro lado encontramos las líneas aguamarina que señalan los percentiles 30 y 70 que hace referencia al valor de el 0.68% y 4.47% respectivamente, es decir, como observamos la aculumación de la densidad está alrededor de el percentil 30, la línea roja hace referencia a la media 3.62% y la línea azul hace referencia a la mediana 1.95%.

Luego de tener en cuenta los percentiles señalados anteriormente se procedió a clasificar a los municipios, si tenían valores menores al percentil 30 entonces se consideraba como “Bajo”, si estaba entre el percentil 30 y 70 entonces se consideraba como “Medio” y si estaba por encima del percentil 70 entonces se consideraba como “Alto”. Una vez se tiene la clasificación de cada uno de los municipios entonces se procede a hacer un proceso de conteo por cada uno de los departamentos:

Con la anterior información se crea una tabla de frecuencias, para estandarizar la información en escala porcentual.

Nota: Si se desea conocer la totalidad de la clasificación consultar el anexo 5.2

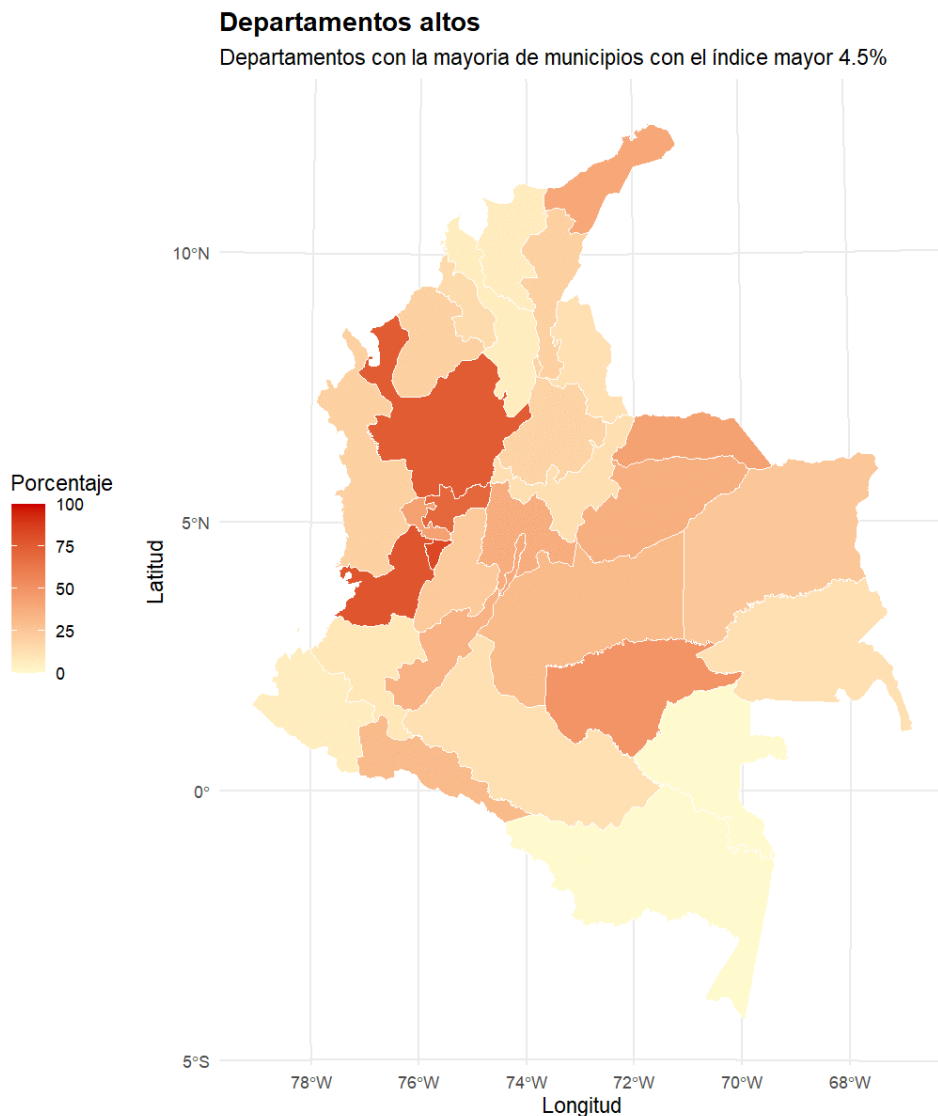
Mapa de Resultados

A continuación se construyó una especie de mapa de calor que muestran las concentraciones de municipios en cada departamento dependiendo de la distribución de los mismos en el índice.

El mapa tiene 4 momentos (para la consulta de los mapas por individual ir al Anexo de mapas).

- El primer momento muestra la concentración de los Departamentos que tienen un alto porcentaje de municipios con accesos a Internet mayores al percentil 70, es decir mayores al 4.5% del índice.
- El segundo momento muestra la concentración de los Departamentos que tienen un porcentaje de municipios medio, entre el percentil 30 y 70 lo cual corresponde al 0.7% y 4.5% del índice.
- El tercer momento muestra la concentración de los Departamentos que tienen un alto porcentaje de municipios con accesos a Internet mayores al percentil 30, es decir menores al 0.7% del índice.
- El cuarto y último momento muestra la concentración de los departamentos que no tienen información.

Entre el color se acerca a tonos más profundos indica mayor concentración porcentual.



Fuente: Elaboración propia

Lo que se observa en el mapa anterior es lo siguiente:

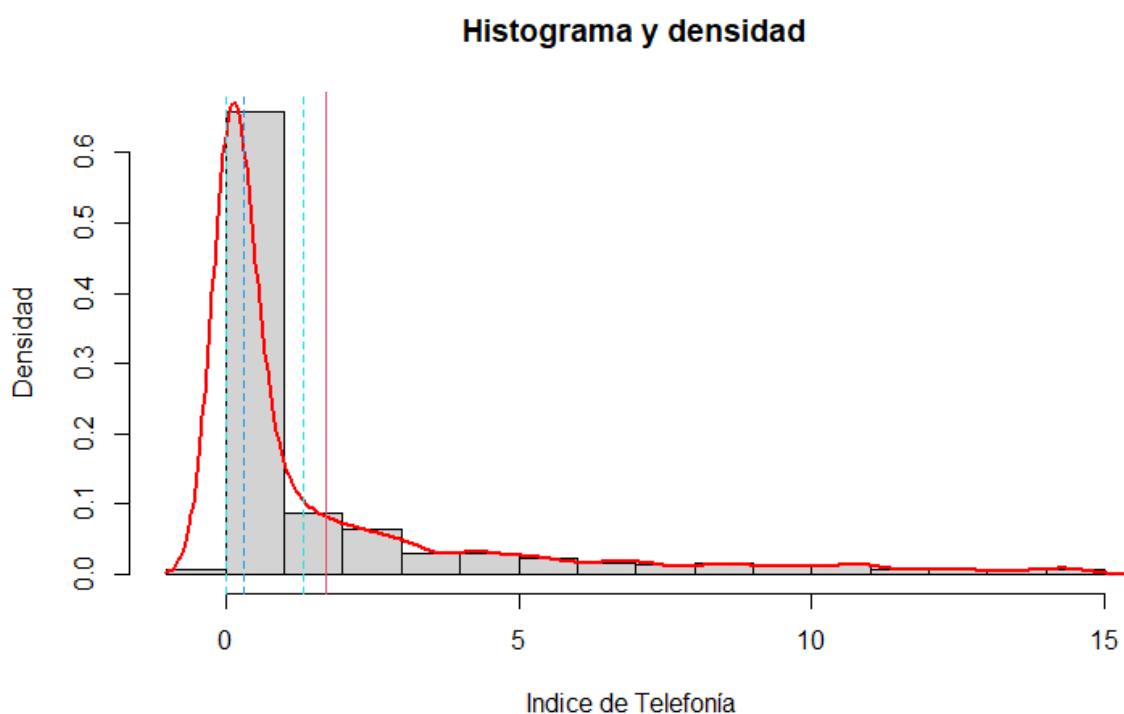
1. Hay departamentos de los cuales se posee muy poca información con código que corresponden a San Andrés, Amazonas, Guainía, Guaviare, Vaupés.

2. Los Departamentos donde la mayoría de los Municipios tiene conexiones mayores al percentil 70, correspondientes a Antioquia, Caldas, Quindío, Valle.
3. Por otro lado, aquellos que tienen un nivel más bajo que los demás, correspondiente a Cauca, Magdalena.

5.2 Índice de Telefonía fija

Para obtener el índice municipal de la cantidad de accesos en función a la población ir al anexo 5.1.

El primer paso que se hizo para analizar el comportamiento del índice fue recortar al percentil 95, para obtener una idea más acertada del comportamiento general.



Fuente: Elaboración propia

Lo que observamos en la anterior gráfica es que existe una acumulación muy pronunciada en porcentajes muy bajos, es decir, que la mayoría de municipios tiene accesos a telefonía muy bajos en comparación a la población total, por otro lado encontramos las líneas aguamarina que señalan los percentiles 10 y 70 que hace referencia al valor de el 0.01% y 1.8% respectivamente, la línea roja hace referencia a la media 1.71% y la línea azul hace referencia a la mediana 0.32%.

Luego de tener en cuenta los percentiles señalados anteriormente se procedió a clasificar a los municipios, si tenían valores menores al percentil 10 entonces se consideraba como

“Bajo”, si estaba entre el percentil 10 y 70 entonces se consideraba como “Medio” y si estaba por encima del percentil 70 entonces se consideraba como “Alto”.

Nota: Si se desea conocer la totalidad de la clasificación consultar el anexo 5.2

Una vez se tiene la clasificación de cada uno de los municipios entonces se procede a hacer un proceso de conteo por cada uno de los departamentos y con la anterior información se crea una tabla de frecuencias, para estandarizar la información con escala porcentual .

Mapa de Resultados

A continuación se construyó una especie de mapa de calor que muestran las concentraciones de accesos a Telefonía fija de los municipios en cada departamentos.

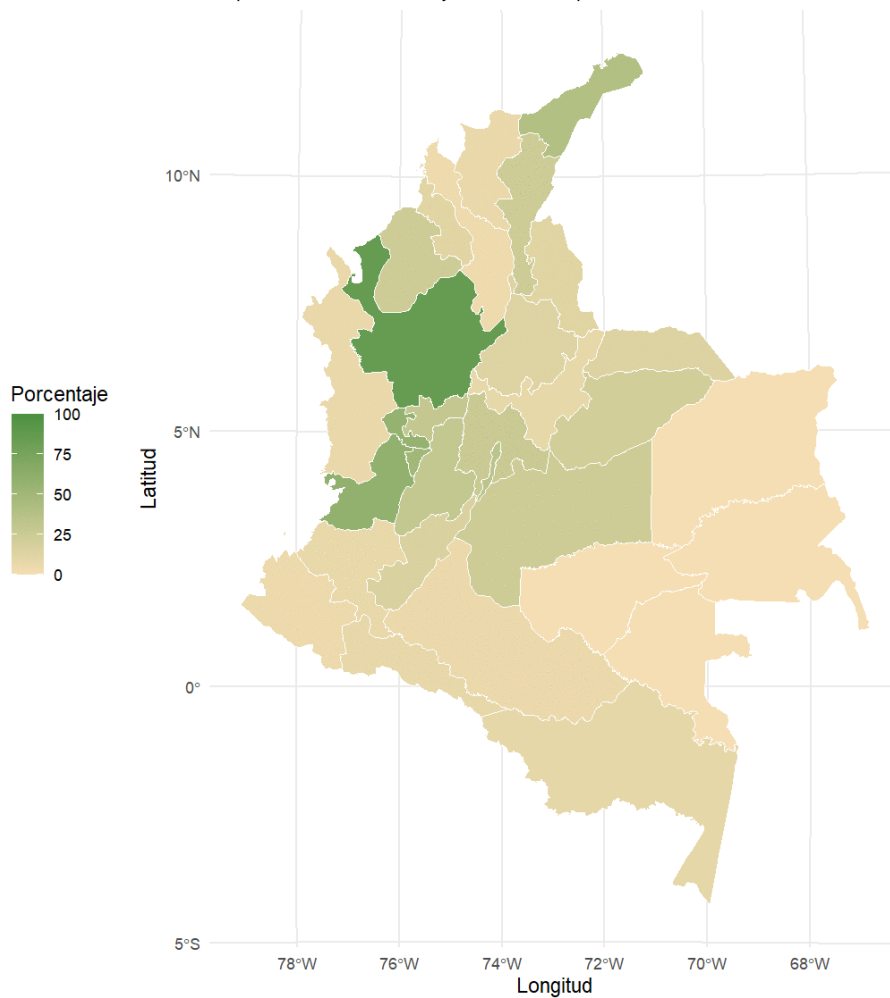
El mapa tiene 4 momentos (para la consulta de los mapas por individual ir al Anexo de mapas).

- El primer momento muestra la concentración de los Departamentos que tienen un alto porcentaje de municipios con accesos al servicio de telefonía fija mayores al percentil 70, es decir mayores al 1.7% de la población.
- El segundo momento muestra la concentración de los Departamentos que tienen un porcentaje de municipios medio, entre el percentil 10 y 70 lo cual corresponde al 0.01% y 1.7% de la población con acceso al servicio de telefonía fija.
- El tercer momento muestra la concentración de los Departamentos que tienen un alto porcentaje de municipios con accesos al servicio de telefonía fija menores al percentil 10, es decir menores al 0.01% de la población.
- El cuarto y último momento muestra la concentración de los departamentos que no tienen información.

Entre el color se acerca a tonos más profundos indica mayor concentración porcentual.

Departamentos accesos altos a Telefonía fija

Departamentos con la mayoría de municipios con índices menores al 1.7%

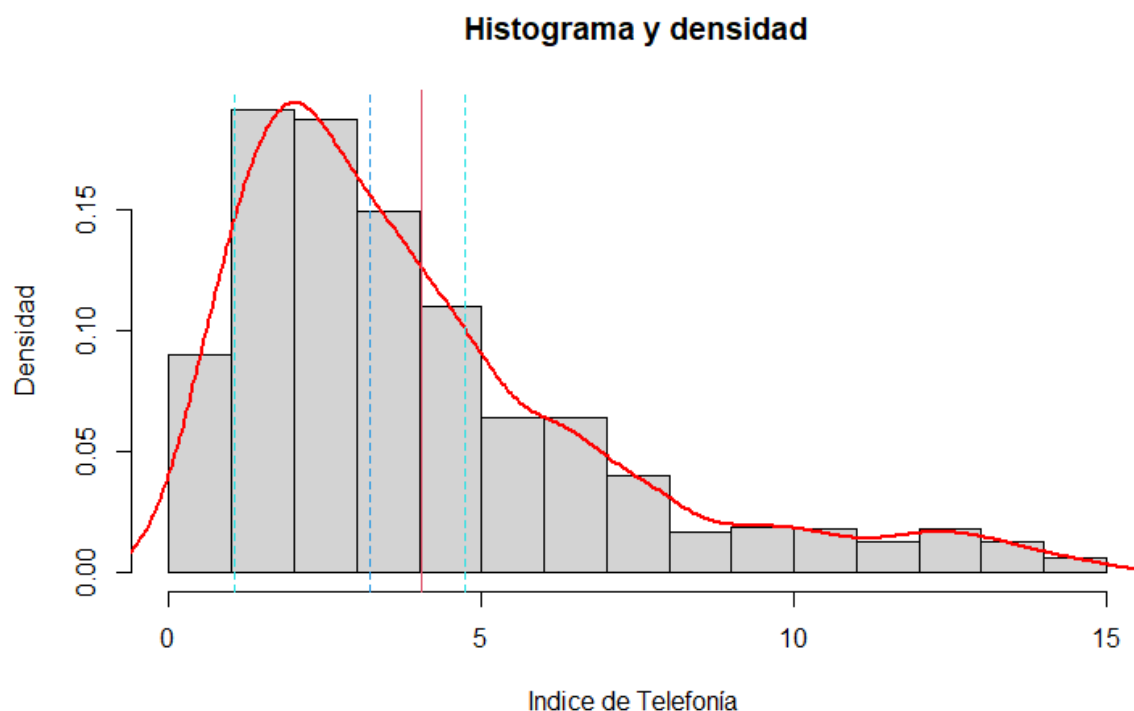


Lo que podemos observar a una gran cantidad de departamentos en la parte sur del País que tiene ausencia de información, además de que muy pocos municipios tienen comportamientos medios.

5.3 Índice de Televisión

Para obtener el índice municipal de la cantidad de accesos en función a la población ir al anexo 5.1.

El primer paso que se hizo para analizar el comportamiento del índice fue recortar al percentil 95, para obtener una idea más acertada del comportamiento general.



Fuente: Elaboración propia

Lo que observamos en la anterior gráfica es que existe una acumulación muy pronunciada en porcentajes muy bajos, es decir, que la mayoría de municipios tiene accesos muy bajos en comparación a la población total, por otro lado encontramos las líneas aguamarina que señalan los percentiles 10 y 70 que hace referencia al valor de el 1.05% y 4.72% respectivamente, lo que se observa es que la concentración está en picos más alejados de cero diferentes a los otros dos servicios,, la línea roja hace referencia a la media 4.03% y la línea azul hace referencia a la mediana 3.21%.

Luego de tener en cuenta los percentiles señalados anterioremente se procedio a clasificar a los municipios, si tenían valores menores al percentil 10 entonces se consideraba como “Bajo”, si estaba entre el percentil 10 y 70 entonces se consideraba como “Medio” y si estaba por encima del percentil 70 entonces se consideraba como alto.

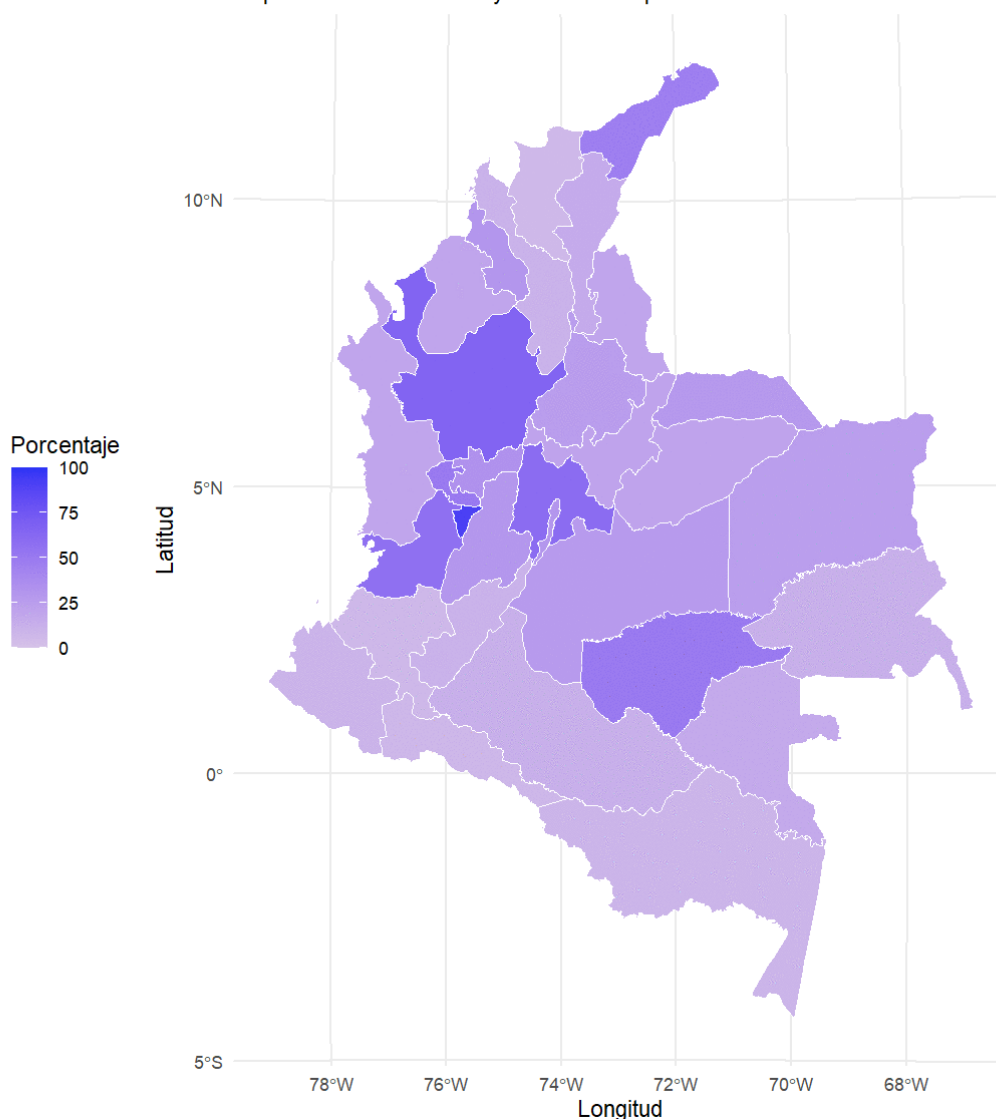
Nota: Si se desea conocer la totalidad de la clasificación consultar el anexo 5.2

Una vez se tiene la clasificación de cada uno de los municipios entonces se procede a hacer un proceso de conteo por cada uno de los departamentos:

Con la anterior información se crea una tabla de frecuencias, para estandarizar la información con escala porcentual y se crea un Mapa de Análisis que muestra el comportamiento de los municipios en términos de los índices municipales, con el fin de captar la variabilidad de cada uno de los departamentos.

Departamentos accesos altos a Televisión

Departamentos con la mayoría de municipios con índice menores al 4.7%



Fuente: Elaboración propia

Se observa ciertos comportamientos en el mapa, en principio accesos altos están concentrados en la zona cafetera, por otro lado la mayoría de los municipios están en concentraciones medias, muy pocos departamentos tienen acceso bajos, y el patrón de ausencia de datos que se repite en los anteriores servicios está vigente en la telefonía.

6. Conclusiones:

- a. Los índices contruidos en esta investigación, permiten comparar en el tiempo la variación anual en la cobertura de las TICs, con un nivel de desagregación regional.
- b. No existe información registrada a la fecha disponible de manera pública que permita construir indicadores de cobertura para la telefonía móvil con desagregación regional.

- c. A nivel general la distribución de los índices de los servicios de internet fijo y de telefonía fija tienen acumulaciones cercanas a cero, mientras que el pico del servicio de televisión está un poco más alejado a valores bajos.
- d. La telefonía se caracteriza por no estar distribuida en valores medios, mientras que el internet y la televisión se distribuye en valores extremos, la televisión en valores altos y el internet en valores bajos.
- e. La clasificación de los municipios según su evolución en el tiempo del índice de cobertura de las TICs, refleja el nivel de heterogeneidad en el comportamiento del índice no solo en los municipios de diferentes departamentos, sino en los municipios del mismo departamento.
- f. En los resultados del indicador de cobertura TIC nos permite identificar las diferencias entre los departamentos y a su vez identificar los municipios dentro de los departamentos que presentan un comportamiento diferente, un ejemplo de esto es la diferencia entre el comportamiento del departamento de Cundinamarca en relación con el comportamiento de Vaupés, donde se puede ver que las curvas oscilan entre diferentes valores, en Cundinamarca valores más grandes que Vaupes, esto se puede relacionar con el PIB de cada departamento, siendo Cundinamarca el departamento con mayor PIB y Vaupés el departamento con menor PIB, dando una correlación entre estas dos variables económicas.

7. Bibliografía:

- a. Comisión de Regulación de Comunicaciones. (2021). Postdata Más allá de los datos. Obtenido de <https://www.postdata.gov.co/search/type/dataset>
- b. Champely S, Doledec S. 1997. How to separate long-term trends from periodic variation in water quality monitoring.
- c. Faraway JJ. 1997. Regression analysis for a functional response.