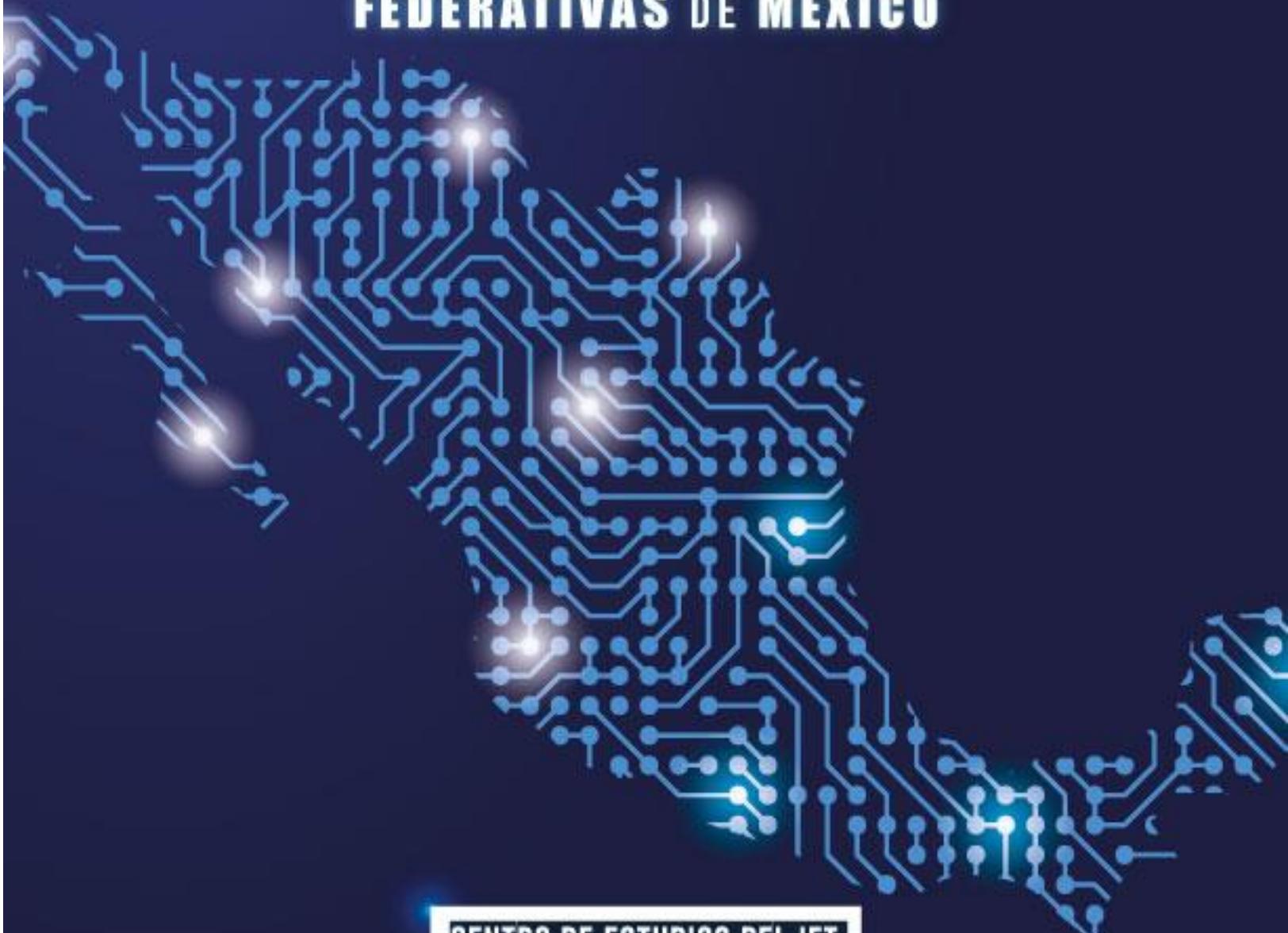


REPORTE ANALÍTICO DE INDICADORES

**INDICADORES DE
TELECOMUNICACIONES
E ÍNDICE DE BIENESTAR
DIGITAL PARA ENTIDADES
FEDERATIVAS DE MÉXICO**



CENTRO DE ESTUDIOS DEL IFT

TERCER CUATRIMESTRE 2021.

MTRO. JOSÉ ALBERTO CANDELARIA BARRERA

Reporte Analítico de Indicadores

Indicadores de telecomunicaciones e índice
de bienestar digital para entidades
federativas de México.

Mtro. José Alberto Candelaria Barrera

Introducción

Las telecomunicaciones tienen hoy en día una gran relevancia en la vida de las personas y en las actividades productivas de los países. De manera creciente su uso ha permitido ampliar y facilitar la comunicación entre los individuos, les ha dado acceso a servicios de educación, salud y financieros, así como al conocimiento y a la recreación. En lo económico, las telecomunicaciones son necesarias para promover la productividad, el crecimiento y la recuperación de la actividad productiva.

El progreso tecnológico de las telecomunicaciones es un fenómeno multifactorial que requiere de la consideración de distintos indicadores para ser capturado, por lo que su medición integral es compleja. La existencia de diferentes variables que señalan el avance de los distintos aspectos que conforman el desarrollo tecnológico dificulta establecer conclusiones y definir una estrategia de conectividad acorde con las necesidades del país. Asimismo, el progreso en materia de telecomunicaciones implica identificar la adopción efectiva de los servicios por parte de la población, ya que sólo a través de la apropiación de los mismos es que la digitalización se convierte en un factor multiplicador del bienestar y el crecimiento ¹.

El objetivo del trabajo es proveer un diagnóstico completo del avance estatal en el desarrollo digital, así como integrar información útil para identificar los rezagos en la materia, a partir de 19 indicadores básicos, clasificados en cuatro dimensiones: infraestructura de telecomunicaciones; adopción de tecnologías de la información y comunicaciones (TIC's), inclusión digital para la actividad económica de las personas, y el uso del internet.

Utilizando distintas técnicas estadísticas multivariadas se integran clústers o agrupamientos de las entidades federativas con el objeto de identificar aquellos estados de la República Mexicana que presentan niveles similares de avance en conectividad, en la adopción y uso de las tecnologías y en servicios de telecomunicaciones.

Adicionalmente, se construye un índice de desarrollo digital para las entidades federativas, el cual permite establecer un ordenamiento respecto al avance relativo de cada una. Esto es, el ejercicio identifica qué entidades reportan un desarrollo similar en cuanto a la dotación integral de los servicios de telecomunicaciones, así como en su adopción, en función de los 19 indicadores básicos.

¹ Katz, R. & Koutroumpis, P. Measuring digitization: A growth and welfare multiplier. *Technovation* Volume 33, Issues 10-11, October–November 2013, Pages 314-319.

A las dimensiones definidas se les asigna una ponderación para reflejar la relevancia que tienen en función de los ejes de política pública de telecomunicaciones y desarrollo digital, establecidos por el gobierno federal, a través de la Secretaría de Comunicaciones y Transportes², así como, a partir del objetivo 3 del Programa Sectorial derivado del Plan Nacional de Desarrollo 2019-2024. En estos instrumentos de planeación se reconoce la relevancia de la promoción de infraestructuras de telecomunicaciones, del incremento de la cobertura y la disponibilidad de los servicios de telecomunicaciones y radiodifusión, así como del desarrollo de capacidades y habilidades digitales.

En efecto, el desarrollo de capacidades y habilidades digitales es importante porque se vincula con la obtención de empleos mejor remunerados (Pirzada y Khan³), o con el acceso a más y mejor información (Van Deursen & Van Dijk⁴). En México para 2020 se alcanzó la cifra de 84 millones de usuarios de Internet, considerando exclusivamente a la población mayor de seis años, de los cuales el 51.3% son mujeres. Aun así, todavía existen millones de personas que no tienen acceso a Internet en México⁵. Es así, que en función de los ejes de política pública y de los retos a los cuales se enfrenta actualmente el país, se opta por asignar una mayor ponderación a la dimensión de infraestructura de telecomunicaciones, la cual es de 40%; en tanto que a las tres dimensiones restantes se les asigna un ponderador de 20%, reflejando su similar relevancia.

Los resultados de este análisis multifactorial sobre el avance del sector de las telecomunicaciones a nivel estatal permiten establecer seis grupos, cada uno con niveles similares de desarrollo digital. A partir del ordenamiento obtenido de los agrupamientos y del índice de desarrollo digital se concluye que Baja California, Baja California Sur, CDMX, Colima y Querétaro integran el grupo que muestra un mayor avance en el conjunto de los distintos servicios de telecomunicaciones con respecto al resto de entidades. En el otro extremo se encuentran Guerrero, Oaxaca y Chiapas. Así también, se detecta que existen mayores brechas entre estados en materia de la penetración de las líneas fijas y en la teledensidad de los servicios móviles. El reporte corrobora que los estados con menor nivel de vida son también los que tienen menos equipamiento y adopción de los servicios de telecomunicaciones.

El trabajo busca proveer al lector con un insumo útil para identificar brechas entre las treinta y dos entidades federativas, y generar información analítica útil para el diseño e implementación de políticas públicas que acorten esas diferencias.

² Página 9, https://www.dof.gob.mx/nota_detalle.php?codigo=5596042&fecha=02/07/2020

³ Pirzada, K. & Khan F. Measuring Relationship between Digital Skills and Employability. European Journal of Business and Management www.iiste.org ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online). Vol.5, No.24, 2013.

⁴ Van Deursen, A. & Van Dijk J. Improving digital skills for the use of online public information and services. Government Information Quarterly. Volume 26, Issue 2, April 2009, Pages 333-340.

⁵ <https://www.inegi.org.mx/programas/dutih/2020/#Tabulados>

Información estadística y consideraciones metodológicas

Para el ejercicio que se presenta en este reporte, se utilizan diecinueve indicadores del sector de las telecomunicaciones diferenciados en cuatro categorías o dimensiones: infraestructura, la inclusión digital, usos del internet. Estas categorías sirven para agrupar las variables, y, a la vez, para adjudicarles un peso ponderado el cual es de utilidad al momento de construir el índice de bienestar digital. A continuación, se explica cada una de dichas categorías; asimismo, se muestra el peso ponderado que se les asigna individualmente, así como los indicadores que las integran.

1. Infraestructura de TIC´s.

Esta categoría⁶ incluye indicadores que miden la conectividad, que es la base para la prestación de todos los servicios digitales. Las variables agrupadas en esta dimensión muestran la extensión con la cual la población tiene acceso a infraestructura de telecomunicaciones, ya sea que se trate de redes fijas, móviles o de TV restringida. En este sentido se utilizan dos mediciones principales: la penetración y la teledensidad de las TIC´s. A esta categoría se le asigna una ponderación de 40% e incluye siete indicadores. Su fuente de información es el Banco de Información en Telecomunicaciones (BIT) del Instituto Federal de Telecomunicaciones⁷, salvo para el porcentaje de población cubierta por tecnología de red 4G para servicio móvil cuya fuente es el Sistema de Análisis de Telecomunicaciones y Conectividad (SATYC)⁸.

- i. Penetración de telefonía fija (BIT)
- ii. Penetración de banda ancha fija (BIT)
- iii. Penetración de TV restringida (BIT)
- iv. Teledensidad de telefonía móvil (BIT)
- v. Teledensidad de internet móvil (BIT)
- vi. Penetración de banda ancha fija con acceso de fibra óptica (BIT)
- vii. Porcentaje de población cubierta por tecnología de red 4G para servicio móvil

2. Adopción de TIC´s.

La segunda dimensión⁹ con la cual se trabaja refleja tanto el equipamiento con el que cuentan las personas en sus hogares, como la implementación exitosa de dichas tecnologías la cual se ve reflejada en su utilización. A esta dimensión se le asigna una ponderación del 20%. La fuente de información de todas las variables

⁶ En *Measuring the Digital Economy A New Perspective* de la OECD publicado diciembre de 2014 se utilizan las dimensiones de infraestructura de tecnologías de la información y adopción de tecnología.

⁷ <https://bit.ift.org.mx/BitWebApp/>

⁸ <http://satyc.ift.org.mx:8010/connect/analyst/mobile/#/login>

⁹ Idem.

es la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2020¹⁰, del INEGI.

- viii. Porcentaje de usuarios que cuentan con teléfono celular
- ix. Porcentaje de usuarios que cuentan con Smartphone
- x. Porcentaje de usuarios de telefonía celular con plan tarifario (postpago)
- xi. Porcentaje de usuarios de internet que lo utilizan diariamente

3. Inclusión digital y actividad económica.

La categoría de inclusión digital y actividad económica¹¹ refleja tres instrumentos que se consideran básicos, los cuales son: la tecnología, el acceso a la red y el dominio de esas herramientas. Una explicación de lo anterior sería que no basta simplemente con que las personas tengan una computadora o un teléfono inteligente conectados a internet para considerar a dicha persona como un incluido digitalmente, se precisa saber qué hacer con estas tecnologías. Es decir, la inclusión digital es también la rutina diaria para optimizar el tiempo y sus posibilidades. En este sentido es que se consideran indicadores como el porcentaje de usuarios que realizan pagos o compras por internet. A esta categoría se le asigna una ponderación del 20%. La fuente de información de todas las variables es la Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2020¹², del INEGI.

- xii. Porcentaje de usuarios que realizan compras por internet
- xiii. Porcentaje de usuarios que realizan pagos por internet
- xiv. Porcentaje de usuarios que usan computadora para actividades laborales
- xv. Porcentaje de usuarios que realizan interacciones con el gobierno

4. Usos de internet.

La última categoría es la de usos del internet¹³; ésta consiste en el uso adecuado, oportuno, pertinente y eficiente de la tecnología de banda ancha con el fin de facilitar la realización de las actividades y el desarrollo de los procesos. En este sentido, se consideran variables como: el porcentaje de usuarios que lo utilizan para comunicación o para obtener información. A esta categoría se le asigna una ponderación del 20%. La fuente de información de todas las variables es la

¹⁰ <https://www.inegi.org.mx/programas/dutih/2020/#Tabulados>

¹¹ En The Digital Economy and Society Index (DESI) 2020 de la Comisión Europea se construye un índice compuesto, el cual incluye como dimensión la integración de tecnología digital, que tiene como subdimensiones a la digitalización empresarial y el comercio electrónico.

¹² Idem

¹³ Katz, R. & Koutroumpis, P. *Using a digitization index to measure the economic and social impact of digital agendas.*

https://www.researchgate.net/publication/270806225_Using_a_digitization_index_to_measure_the_economic_and_social_impact_of_digital_agendas

Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2020¹⁴, del INEGI.

- xvi. Porcentaje de usuarios internet que lo utilizan para entretenimiento
- xvii. Porcentaje de usuarios de internet que lo utilizan para ver contenidos audiovisuales
- xviii. Porcentaje de usuarios de internet que lo utilizan para comunicación
- xix. Porcentaje de usuarios de internet que lo utilizan para obtener información

La muestra se integra por las treinta y dos entidades federativas de México. Los indicadores cuya fuente es el BIT se encuentran al cuarto trimestre del año 2020, los que provienen de la ENDUTIH corresponden también al año 2020. En tanto que la única variable cuya fuente es el SACYT (Porcentaje de población cubierta por tecnología de red 4G para servicio móvil) se encuentra al cuarto trimestre del año 2019.

Diagnóstico General

Los diecinueve indicadores dan cuenta del desarrollo digital de las entidades federativas de México, y son base del análisis de este estudio. El cuadro 1 siguiente muestra el promedio nacional y la dispersión o disparidad (desviación estándar) que se registra en cada uno de los indicadores. En general no se observa una fuerte dispersión de los datos en ninguno de los 19 indicadores.

Cuadro 1. Indicadores de desarrollo digital. Promedio y desviación estándar de los indicadores

N°	Dimensión	Nombre del indicador	Media	Desviación Estándar
1	1	Porcentaje de población cubierta por tecnología de red 4G para servicio móvil	87.8	2.05
2	1	Penetración telefonía fija	59.2	4.46
3	1	Penetración banda ancha fija	57.9	3.26
4	1	Penetración TV restringida	59.6	1.81
5	1	Teledensidad de telefonía móvil	97.8	1.9

¹⁴ *Idem*

6	1	Teledensidad de internet móvil	80.8	2.55
7	1	Penetración de banda ancha fija con accesos de fibra óptica	21.5	1.97
8	2	Porcentaje de usuarios que cuentan con teléfono celular	75.7	1.28
9	2	Porcentaje de usuarios que cuentan con smartphone	91.1	0.44
10	2	Porcentaje de usuarios de internet que lo utilizan diariamente	73.1	0.89
11		Porcentaje de usuarios de telefonía celular con plan tarifario de pospago	15.2	0.87
12	3	Porcentaje de usuarios de internet que compran productos o servicios	28.1	1.0
13	3	Porcentaje de usuarios de internet que interactúan con el gobierno	31.5	0.94
14	3	Porcentaje de usuarios que realizan pagos por internet	1.03	0.83
15	3	Porcentaje de usuarios que usan computadora para actividades laborales	45.2	0.81
16	4	Porcentaje de usuarios de internet que lo utilizan para entretenimiento	88.6	0.47
17	4	Porcentaje de usuarios de internet que lo utilizan para comunicarse	93.7	0.24
18	4	Porcentaje de usuarios de internet que lo utilizan para buscar información	90.9	0.3
19	4	Porcentaje de usuarios de internet que lo utilizan para ver contenidos audiovisuales	76.6	0.97

En el Anexo 1 de este reporte se presenta una descripción detallada de las cifras estatales para cada indicador.

Determinación de agrupamientos para las entidades federativas

Para integrar grupos de entidades federativas que se asemejan en cuanto a los diecinueve indicadores de telecomunicaciones seleccionados y presentados en el Cuadro 1, se utiliza la técnica de partición conocida como K-Medias¹⁵. Esta técnica utiliza la distancia entre los datos para saber si éstos son parecidos o diferentes entre sí, y es aceptada generalmente por la robustez de sus resultados. La técnica permite identificar a los estados del país que son relativamente homogéneos entre sí, con base en las categorías establecidas, de manera simultánea. Esto es, considerando los diecinueve indicadores de manera simultánea se forman grupos integrados por entidades federativas con la mayor similitud entre sí, pero diferentes respecto a otros grupos de estados definidos.

En el análisis se emplea una segunda técnica de agrupamiento, el método Ward¹⁶. El empleo de dos técnicas diferentes permite contrastar los resultados y tener mayor solidez en las conclusiones. Para una mayor descripción véase el Anexo 2 de este reporte.

A partir de la aplicación de este ejercicio estadístico, se definieron seis grupos o clústers como el número óptimo a considerar. Para llegar a esta conclusión se utiliza una gráfica conocida como Screeplot, (Gráfica A20 del Anexo Metodológico), que muestra el número óptimo de agrupamientos.

En función de los resultados que producen las dos técnicas, se consideró que el método de Ward utilizando la distancia Manhattan es el más adecuado para el análisis, ya que otorgó resultados más robustos. El Cuadro 2 presenta los agrupamientos definidos.

¹⁵ Para una descripción detallada, véase: Peña, D., 2002. Análisis de Datos Multivariados, McGraw Hill.

¹⁶ El segundo método se trata de una técnica jerárquica conocida como método Ward. La idea de ambas técnicas es que aquellas entidades que compartan características semejantes estarán juntas en un mismo grupo, y a su vez, separadas de los otros grupos definidos. Este texto lo quité porque contradice lo que señala el siguiente párrafo.

Cuadro 2. Agrupamientos

Agrupamiento	Entidades Federativas
1	Baja California, Baja California Sur, CDMX, Colima, Querétaro
2	Aguascalientes, Chihuahua, Jalisco, Estado de México, Morelos, Nuevo León, Sonora, Sinaloa, Tamaulipas,
3	Campeche, Quintana Roo, Yucatán
4	Hidalgo, Puebla, Tabasco, Tlaxcala, Veracruz
5	Coahuila, Durango, Guanajuato, Michoacán, Nayarit, San Luis Potosí, Zacatecas
6	Chiapas, Guerrero, Oaxaca

La metodología empleada permite establecer que los estados que integran cada grupo o *clúster* exhiben niveles similares en la penetración de los diferentes servicios de telecomunicaciones fijas y móviles, y en las variables de inclusión y uso digital, con base en los diecinueve indicadores considerados simultáneamente. Así, por ejemplo, la Ciudad de México comparte un nivel similar a entidades como Baja California, Baja California Sur, Colima y Querétaro, considerando el agregado de los indicadores, por lo que quedan integradas en un mismo grupo. No así Oaxaca, Guerrero o Chiapas que presentan el mayor rezago y que son similares entre sí en cuanto al reto en materia de desarrollo digital, por lo que se incluyen en un grupo aparte.

Cabe destacar que las técnicas de agrupamiento permiten integrar en grupos a las entidades federativas con niveles similares de avance o rezago, pero no dan un orden en cuanto al mayor o menor desarrollo alcanzado.

A continuación, se estiman los índices que permiten determinar cuáles son los grupos o *clústers*, y las entidades federativas, que presentan un mayor avance respecto a su desarrollo digital.

Ordenamiento de las Entidades Federativas y Grupos

Para ordenar los agrupamientos con base en los indicadores elegidos, se estiman para cada entidad federativa dos métricas alternativas: un índice de variables

estandarizadas (IVE) basado en la suma de las diecinueve variables estandarizadas¹⁷; y un índice de desarrollo digital (IDD) para cada grupo.

El IVE se define como la suma de las diecinueve variables estandarizadas; que permite estimar un índice para cada grupo. El IVE toma mayor valor en la medida que el nivel de equipamiento, adopción y uso de los servicios del agrupamiento, es más elevado. Para facilitar la lectura, el índice se estima en una escala de *valores que va de 1 a 100*.¹⁸

De manera adicional, en este reporte se procede a la elaboración del IDD. El objetivo es contar con una herramienta que ofrezca una alternativa para clasificar a las entidades. Este se estima para cada grupo (Cuadro 2); y en la siguiente sección se presenta el IDD para cada entidad federativa (Cuadro 3). Para estimar el IDD se utiliza la técnica de análisis de componentes principales¹⁹, la cual es de utilidad cuando se trabaja con bases de datos multivariadas, como lo es este caso. El análisis de los componentes principales permite observar en un solo índice la estructura de los 19 indicadores en su totalidad.

La matemática subyacente al tema del análisis de los componentes principales es compleja, por lo tanto, no se aborda en el presente documento²⁰, pero la esencia es la siguiente: se cuenta con un conjunto de datos con muchos indicadores o variables, es decir, multivariado, el cual es simplificado convirtiendo a las variables originales en un número menor de "componentes principales". Es así como los componentes principales son la estructura subyacente de los datos²¹.

En función de lo anterior, el diseño del IDD conlleva la aplicación de ponderadores a cada una de las cuatro dimensiones definidas, para después construir el índice agregando los resultados. Esta metodología es la misma que aplica The Economist Intelligence Unit en su E-readiness rankings 2009²².

Contar con dos índices diferentes para el ordenamiento de las entidades federativas da solidez al ejercicio, además de que permite comparar los resultados que se obtienen con dos diferentes técnicas de ordenamiento, una ponderada y otra no ponderada.

¹⁷ La estandarización consiste en una transformación lineal de las diecinueve variables con objeto de que tengan un promedio de cero y una desviación estándar igual a uno. Este proceso es necesario para poder sumar variables expresadas en unidades diferentes.

¹⁸ Esto se hace obteniendo las diferencias entre el valor obtenido por el grupo 1 con respecto a cada uno de los cinco grupos restantes; posteriormente, a cada valor resultante se le resta a 100 para obtener el re escalamiento final.

¹⁹ Peña, D., 2002. Análisis de Datos Multivariantes, McGraw Hill.

²⁰ *Ibidem*

²¹ Los componentes principales determinan la dirección (o direcciones) hacia la cual se encuentra la mayor varianza de los datos considerados. Por tanto, lo que se trata es de encontrar la línea recta que mejor se extienda a través de los datos al momento de ser proyectada. Este sería el primer componente principal. Para mayor explicación, véase: Peña, D., 2002. Análisis de Datos Multivariantes, McGraw Hill

²² E-readiness rankings 2009. A report from the Economist Intelligence Unit. The Economist. <http://graphics.eiu.com/pdf/e-readiness%20rankings.pdf>

El Cuadro 2 presenta los índices estimados, a partir de los cuales se ordenan los 6 grupos previamente definidos. Como puede constatarse, básicamente los dos índices ofrecen un mismo ordenamiento. Se desprende también que tanto un índice no ponderado²³ (IVE) como uno ponderado (IDD) dan el mismo un resultado similar.

Cuadro 2. Entidades federativas según agrupamiento

Orden de agrupamiento	Estados	Índice de variables estandarizadas	Índice de Desarrollo Digital	IDH
1	Baja California, Baja California Sur, CDMX, Colima, Querétaro	100	2.1	0.805
2	Aguascalientes, Chihuahua, Jalisco, Estado de México, Morelos, Nuevo León, Sonora, Sinaloa, Tamaulipas,	83	0.9	0.792
3	Campeche, Quintana Roo, Yucatán	79	0.3	0.733
4	Hidalgo, Puebla, Tabasco, Tlaxcala, Veracruz	65	-0.7	0.766
5	Coahuila, Durango, Guanajuato, Michoacán, Nayarit, San Luis Potosí, Zacatecas	63	-1.3	0.756
6	Chiapas, Guerrero, Oaxaca	52	-2.9	0.706

A partir del ordenamiento obtenido, se tiene que el *clúster* formado por Ciudad de México, Baja California, Baja California Sur, Colima y Querétaro tiene los índices más altos, con un valor igual a 100 para el IVE y 2.1 tratándose del IDD. El segundo *clúster* conformado por Aguascalientes, Chihuahua, Jalisco, Estado de México, Morelos, Nuevo León, Sonora, Sinaloa, Tamaulipas tiene un IVE igual a 83 y un IDD igual a 0.9. En ese orden sigue el grupo de Campeche, Quintana Roo y Yucatán con un IVE de 79 y un IDD igual a 0.3, y así sucesivamente. El Mapa 1 resume los agrupamientos.

²³ El índice de variables estandarizadas no usa explícitamente un ponderador ya que se define como un promedio aritmético de las variables estandarizadas. Por lo anterior, se da igual peso a cada una de las 19 variables.

Se observa que el grupo de Coahuila, Durango, Guanajuato, Michoacán, Nayarit, San Luis Potosí y Zacatecas se encuentra mejor posicionado que el grupo de Hidalgo, Puebla, Tabasco, Tlaxcala y Veracruz.

Finalmente, el grupo de Chiapas, Guerrero y Oaxaca presenta el mayor rezago con un IVE de 52 y un IDD de -2.9.

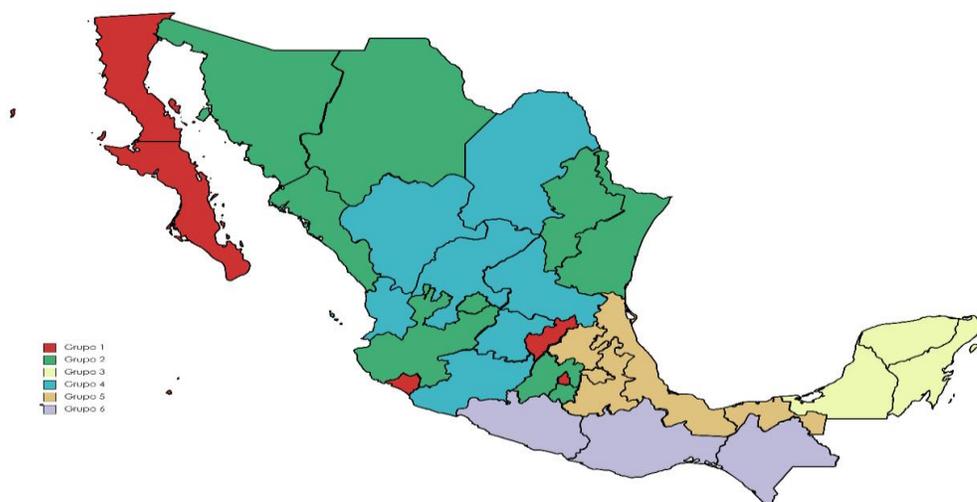
Las telecomunicaciones y la prestación del servicio de internet generan un efecto benéfico en el nivel de vida de las personas y en su actividad productiva. A su vez se reconoce que la población con mayor nivel de vida cuenta con más educación y acceso a las tecnologías, lo que favorece la adopción de las mismas. Así, el análisis se enriquece si se toma en consideración un indicador del nivel de bienestar social de la población y se compara con el ordenamiento que proporcionan el IVE y el IDD. Específicamente se emplea en este reporte el Índice de Desarrollo Humano del Programa de Naciones Unidas para el Desarrollo²⁴ (IDH, por sus siglas) para las entidades federativas México, el cual mide el nivel de desarrollo de cada estado atendiendo a variables como la esperanza de vida, la educación o el ingreso per cápita.

Para cada agrupamiento se obtiene el promedio del IDH de las entidades federativas que integran a cada clúster (ver Cuadro 2). El primer clúster conformado por Ciudad de México, Baja California, Baja California Sur, Colima y Querétaro alcanza un IDH promedio igual a 0.805. Por su parte, el segundo clúster integrado por Aguascalientes, Chihuahua, Jalisco, Estado de México, Morelos, Nuevo León, Sonora, Sinaloa y Tamaulipas tiene un IDH menor respecto al del primer grupo, con un valor de 0.792.

Se observa la concordancia esperada entre el IDD y el IDH, confirmando la relación entre el nivel de vida y el desarrollo digital. Únicamente en el caso de un agrupamiento se registra un rompimiento del orden esperado. Específicamente el clúster 4, formado por Coahuila, Durango, Guanajuato, Michoacán, Nayarit, San Luis Potosí y Zacatecas tiene un IDH promedio de 0.766, superior al 0.733 del grupo 3 de Campeche, Quintana Roo y Yucatán.

Finalmente el grupo conformado por Chiapas, Guerrero y Oaxaca tiene el IDH promedio más bajo, con una cifra igual a 0.706. Las cifras sugieren que el nivel de equipamiento de infraestructura de telecomunicaciones y la adopción de los servicios es mayor en los grupos de entidades que también presentan un mayor nivel de vida, de acuerdo con el IDH.

²⁴ https://es.wikipedia.org/wiki/Anexo:Entidades_federativas_de_M%C3%A9xico_por_IDH



Análisis de la información estatal

Para completar el panorama presentado a nivel de agrupamiento, se estimó el índice de desarrollo digital o IDD para cada una de las entidades federativas. Estas cifras se presentan en el Cuadro 3, en el que también se presentan los valores para cada una de las 4 dimensiones que, ponderadas, integran el IDD. Tal y como puede observarse la Ciudad de México reporta el IDD más elevado con un índice de 3.2, la segunda posición en el IDD es para Baja California Sur con 2.1, seguida por Baja California con 1.9 y Nuevo León con 1.8. Las posición general llega a cambiar cuando se analizan los casos individuales por dimensión; por ejemplo, para la categoría 4, *Usos del Internet*, la CDMX con un 1.10 se encuentra por debajo de Baja California Sur con 1.94 y Colima con 3.06.

En contraste Guerrero, Oaxaca y Chiapas acusan el menor desarrollo digital, con valores para su IDD de -2.1, -3.1 y -3.7, respectivamente.

Aunado a lo anterior, en la última columna del mismo Cuadro 3 se muestra el IDD para cada entidad federativa calculado para el año 2020. Los cambios no son demasiado drásticos, manteniendo Ciudad de México la primera posición con un valor de 3.2 el cual disminuyó ligeramente respecto al 3.6 del año anterior. A ésta le siguen Baja California Sur con 2.1 y Baja California con 1.9. Esta última entidad federativa vio reducido su índice con respecto al 2.7 que obtuvo en 2020. En las primeras dos columnas se pueden apreciar las posiciones numeradas tanto del presente año (2021) como del anterior (2020), así como un signo de igual (=) al lado del acrónimo de aquellas entidades federativas que no vieron un cambio en su posición, y flechas hacia arriba (↑) o hacia abajo (↓) en los casos que sí lo experimentaron. Siendo así, Baja California subió dos lugares, mientras que Nuevo

León cayó dos lugares en la clasificación. Colima subió cuatro lugares, en tanto que Morelos pasó del puesto catorce al octavo. Hubo algunas caídas importantes como la de Quintana Roo que paso del quinto al noveno lugar o Jalisco del octavo al décimo. Las entidades peor posicionadas siguen siendo las mismas y no vieron modificado su puesto, tratándose de Guerrero, Oaxaca y Chiapas.

Cuadro 3. Estimación estatal del índice de Desarrollo Digital

Posición 2021	Posición 2020	Entidad	Dimensión 1. (40%)	Dimensión 2. (20%)	Dimensión 3. (20%)	Dimensión 4. (20%)	Índice de Desarrollo Digital (IDD) 2021	Índice de Desarrollo Digital (IDD) 2020
1	1	CDMX =	4.45	3.27	2.56	1.10	3.2	3.6
2	4	BCS ↑	2.42	1.70	1.90	1.94	2.1	2.2
3	3	BC =	2.31	1.99	0.03	3.06	1.9	2.7
4	2	NL ↓	2.32	2.52	0.77	1.31	1.8	2.7
5	9	COL ↑	1.71	2.38	2.10	1.09	1.8	1.3
6	6	QRO =	1.15	2.10	2.46	1.84	1.7	1.5
7	7	SON =	2.68	0.77	0.00	-0.07	1.2	1.4
8	14	MOR ↑	1.33	0.31	0.81	1.87	1.1	0.6
9	5	QROO ↓	0.36	0.98	2.47	1.30	1.1	1.6
10	8	JAL ↓	1.76	1.08	-0.03	0.63	1.0	1.3
11	10	AGS ↓	0.53	0.84	1.41	1.24	0.9	1.1
12	15	SIN ↑	1.83	0.35	-0.12	0.12	0.8	0.5
13	13	MEX =	1.00	1.14	0.45	-0.08	0.7	0.6
14	12	CHIH ↓	0.14	0.95	-0.13	1.58	0.5	0.8
15	17	TAM ↑	0.93	0.30	-1.23	0.27	0.2	0.4
16	11	YUC ↓	-0.61	1.25	1.98	-0.98	0.2	0.9
17	16	COAH ↓	0.33	0.37	-0.33	0.20	0.2	0.5
18	24	DGO ↑	-0.62	0.03	-0.76	-0.05	-0.4	-1.3
19	28	CAM ↑	-1.72	0.36	0.08	0.80	-0.4	-1.5
20	20	GTO =	-0.70	-0.05	-0.70	-0.11	-0.5	-0.8
21	18	NAY ↓	0.21	-1.00	-0.43	-1.43	-0.5	-0.3
22	19	SLP ↓	-0.98	-0.54	-0.53	-1.04	-0.8	-0.4
23	21	MICH ↓	-1.04	-1.67	-1.22	0.06	-1.0	-1.1
24	23	VER ↓	-1.50	-2.03	-0.25	-0.52	-1.2	-1.3
25	26	PUE ↑	-1.56	-1.94	-1.41	0.50	-1.2	-1.3
26	25	TAB ↓	-1.28	-1.76	-0.05	-1.97	-1.3	-1.3
27	27	TLAX =	-0.91	-1.06	-0.63	-3.03	-1.3	-1.3
28	29	ZAC ↑	-2.14	-1.26	-1.56	-1.01	-1.6	-1.6
29	22	HGO ↓	-1.74	-1.64	-1.32	-1.81	-1.7	-1.3

30	30	GRO =	-2.64	-2.58	-1.31	-1.27	-2.1	-2.6
31	31	OAX =	-3.96	-3.12	-2.06	-2.27	-3.1	-3.0
32	32	CHIS =	-4.05	-4.03	-2.93	-3.28	-3.7	-4.1

Conclusiones

El presente Reporte permite obtener una perspectiva del desarrollo en materia de telecomunicaciones y desarrollo digital de las 32 entidades federativas de México. Mediante la realización de un análisis de clústers y la construcción de un índice de desarrollo digital se clasifican las entidades de acuerdo con el avance alcanzado a partir de diecinueve indicadores de telecomunicaciones clasificados en cuatro dimensiones básicas: la infraestructura de telecomunicaciones; la adopción de TIC's; la inclusión digital para la actividad económica, y los usos del internet. A estas dimensiones se les asigna una ponderación en función de los ejes de política pública establecidos por el gobierno federal²⁵, la cual incluye en sus programas la promoción de infraestructuras, el incremento de la cobertura y la disponibilidad de los servicios de telecomunicaciones y radiodifusión, así como el desarrollo de habilidades y capacidades digitales.

El análisis de clústers conlleva la aplicación de dos técnicas de agrupamiento: el método de K-medias y el método Ward. Se opta por seleccionar seis clústers o agrupamientos. Por otra parte, el diseño del índice de desarrollo digital se lleva a cabo mediante el análisis de componentes principales.

El reporte permite establecer que entidades federativas como Baja California, Baja California Sur, Ciudad de México, Colima y Querétaro forman un grupo con avance similar, mediante el método Ward y la distancia Manhattan (ver Cuadro 2). Este resultado se corrobora con la estimación del índice de desarrollo digital que resulta más elevado para las entidades antes señaladas (ver Cuadros 2 y 3).

Por otra parte, los estados de Guerrero, Oaxaca y Chiapas se encuentran a la zaga en el índice de desarrollo digital (ver Cuadros 2 y 3) y consistentemente, conforman el grupo más rezagado de entre los seis que se clasifican.

Por último, se reconoce que la creciente importancia de las telecomunicaciones en las actividades económicas y en el nivel de vida de las personas hace imperativo el continuar por la senda del crecimiento y la inversión en el sector. Es así que con base en los resultados alcanzados en el presente estudio se recomienda continuar el diseño de políticas públicas y regulación para impulsar un mayor crecimiento en la penetración de los servicios de las telecomunicaciones.

²⁵ https://www.dof.gob.mx/nota_detalle.php?codigo=5596042&fecha=02/07/2020

De manera prioritaria se deberá atender el equipamiento y los esfuerzos de inclusión digital en las entidades que conforman los grupos 6, 5 y 4.

ANEXO METODOLÓGICO

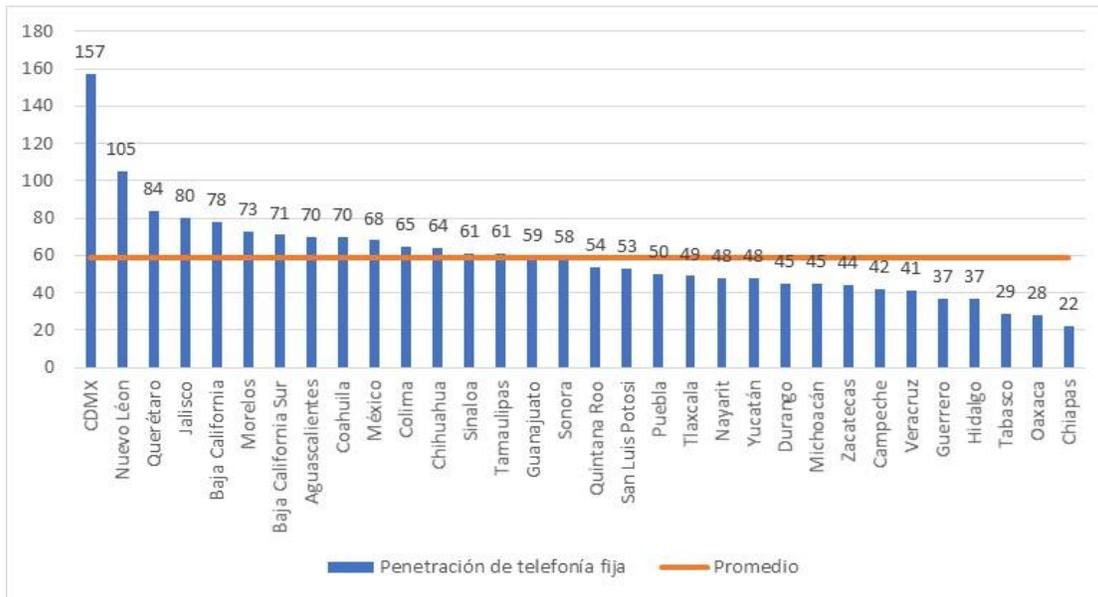
Anexo 1.

Diagnóstico estatal de las variables de telecomunicaciones por entidad federativa

El objetivo del presente trabajo es identificar las similitudes o disimilitudes que guardan las treinta y dos entidades federativas del país, con respecto al acceso y adopción por parte de su población de las tecnologías de la información y comunicaciones; así como de la inclusión digital o los usos del internet. Los diecinueve indicadores de telecomunicaciones que se utilizan en el presente reporte se analizan de forma gráfica a continuación.

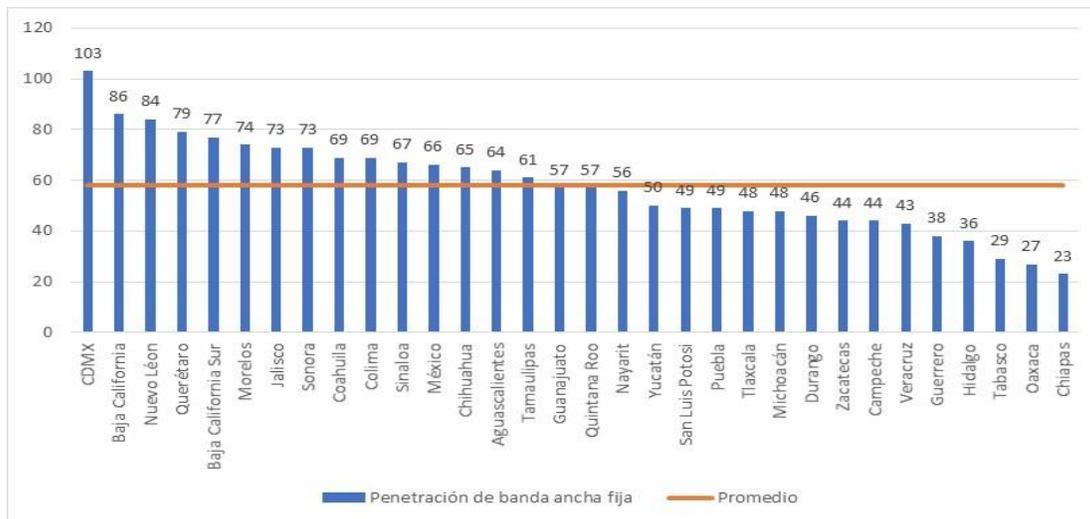
La Gráfica A1 muestra la penetración de telefonía fija por hogares para cada una de las treinta y dos entidades federativas del país. En este caso la CDMX presenta una penetración de 157 líneas de telefonía fija por cada 100 hogares, seguido de Nuevo León y Querétaro con 105 líneas y 84 líneas por cada 100 hogares, respectivamente. Por otra parte, los estados de Tabasco, Oaxaca y Chiapas acusan el mayor rezago con 29, 28 y 22 líneas de telefonía fija por cada 100 hogares. Estas cifras corroboran la amplitud de la brecha digital regional: la CDMX tiene casi ocho veces la penetración reportada en Chiapas. La media de las 32 entidades federativas está representada mediante una línea horizontal, siendo igual a 59.

Gráfica A1. Penetración de telefonía fija



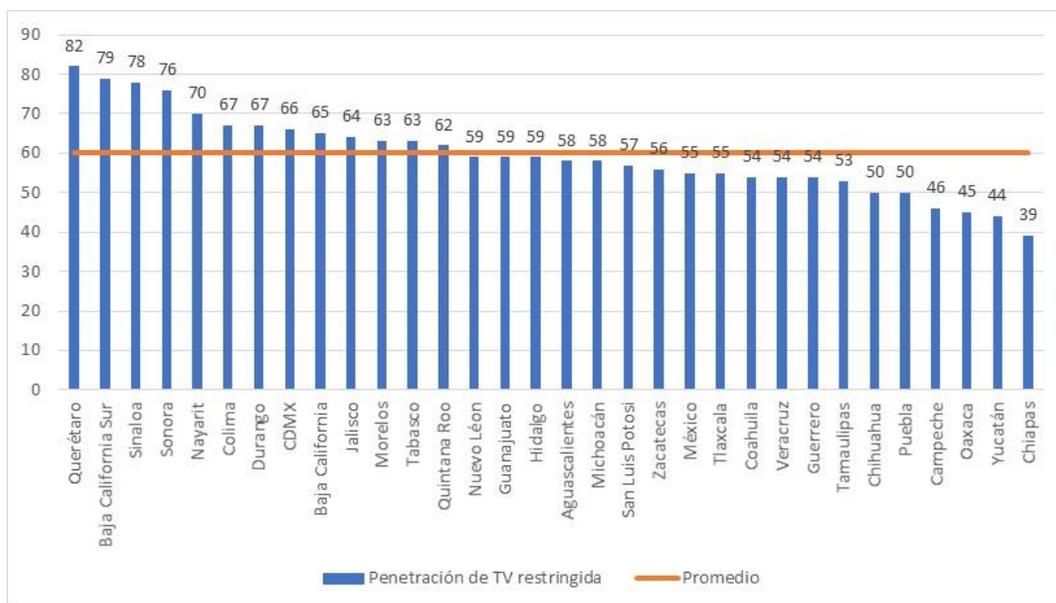
La Gráfica A2 muestra la variable de penetración de banda ancha fija por cada 100 hogares. Esta variable muestra un panorama similar a la variable anterior; es decir, la Ciudad de México se encuentra en el primer sitio con 103 accesos de banda ancha fija por cada 100 hogares, le siguen los estados de Baja California y Nuevo León con 86 y 84 accesos por cada 100 hogares, respectivamente. Las últimas tres posiciones se encuentran ocupadas por Tabasco, Oaxaca y Chiapas con 29, 27 y 23 accesos de banda ancha fija por cada 100 hogares, respectivamente. La media nacional es igual a 58 accesos de banda ancha fija por cada 100 hogares.

Gráfica A2. Penetración de banda ancha fija



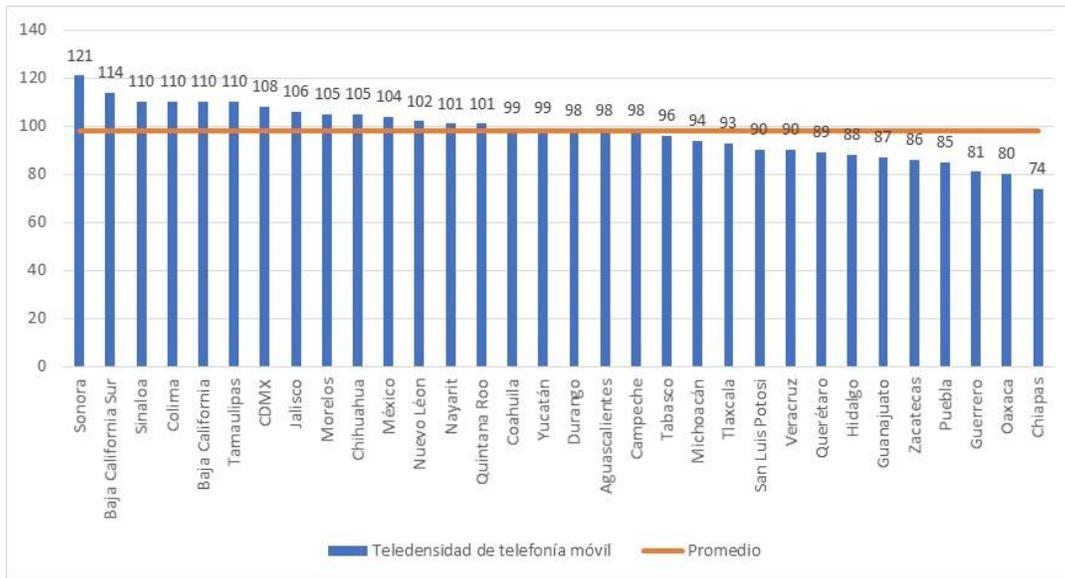
La gráfica A3 corresponde a la penetración de TV restringida por cada 100 hogares. La primera posición la ocupa el estado de Querétaro con 82 suscripciones a TV restringida por cada 100 hogares. A continuación, se encuentran Baja California Sur y Sinaloa con 79 y 78 suscripciones por cada 100 hogares, respectivamente. Los menores niveles de penetración los acusan los estados de Oaxaca, Yucatán y Chiapas con 45, 44 y 39 suscripciones por cada 100 hogares, respectivamente. Por otra parte, la media nacional corresponde a 60 suscripciones de TV restringida por cada 100 hogares.

Gráfica A3. Penetración de TV restringida



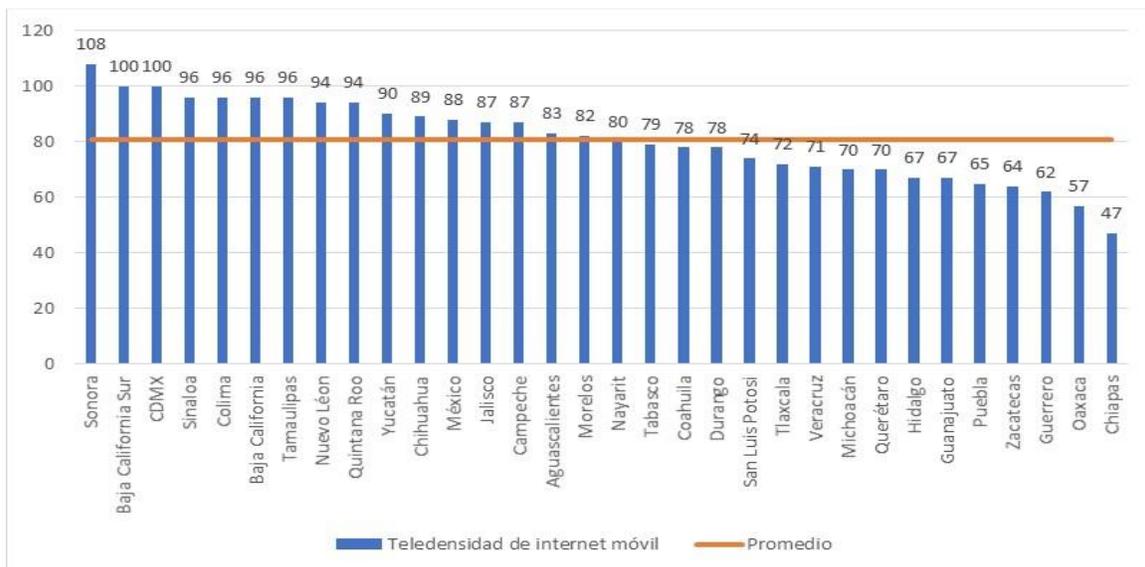
En la Gráfica A4 se observa la teledensidad de telefonía móvil por cada 100 habitantes. El primer sitio es para Sonora con 121 líneas de telefonía móvil por cada 100 habitantes; le siguen Baja California Sur y Sinaloa con 114 y 110 líneas de telefonía móvil por cada 100 habitantes. En el otro extremo se encuentra Guerrero, Oaxaca y Chiapas con 81, 80 y 74 líneas de telefonía móvil por cada 100 habitantes. La media nacional es igual a 98 líneas de telefonía de móvil por cada 100 habitantes.

Gráfica A4. Teledensidad de telefonía móvil



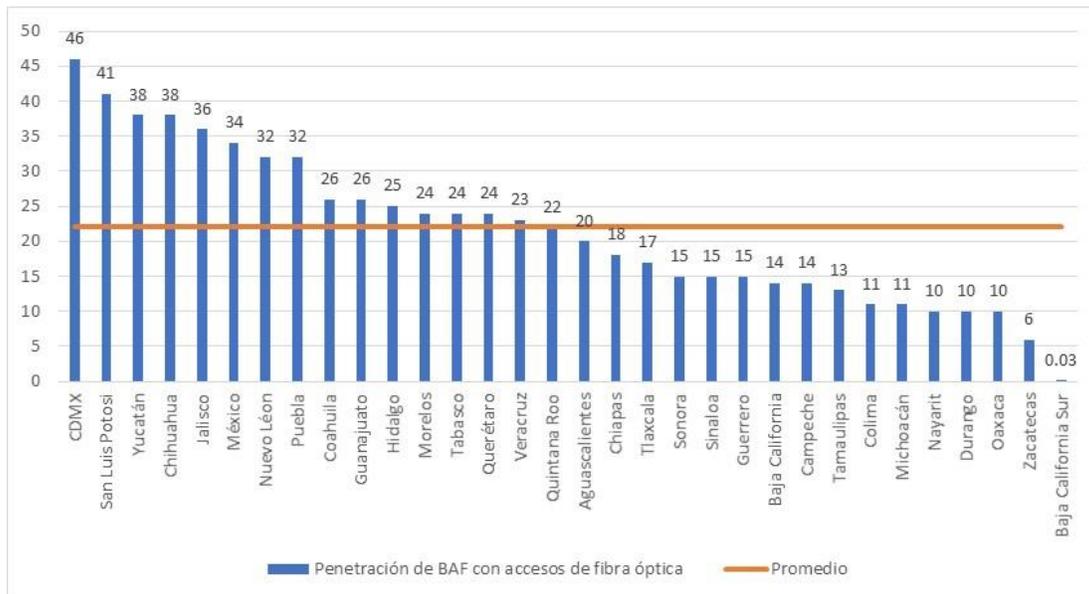
La gráfica A5 muestra la variable de teledensidad de internet móvil por cada 100 habitantes. El primer sitio lo ocupa Sonora con 108 suscripciones a internet móvil por cada 100 habitantes, seguido de Baja California Sur y CDMX con 100 suscripciones por cada 100 habitantes cada uno. Guerrero, Oaxaca y Chiapas son los estados con menores cifras de teledensidad de internet móvil con 62, 57 y 47 suscripciones por cada 100 habitantes, respectivamente. En tanto que el promedio nacional se encuentra en 81 suscripciones a internet móvil por cada 100 habitantes.

Gráfica A5. Teledensidad de internet móvil



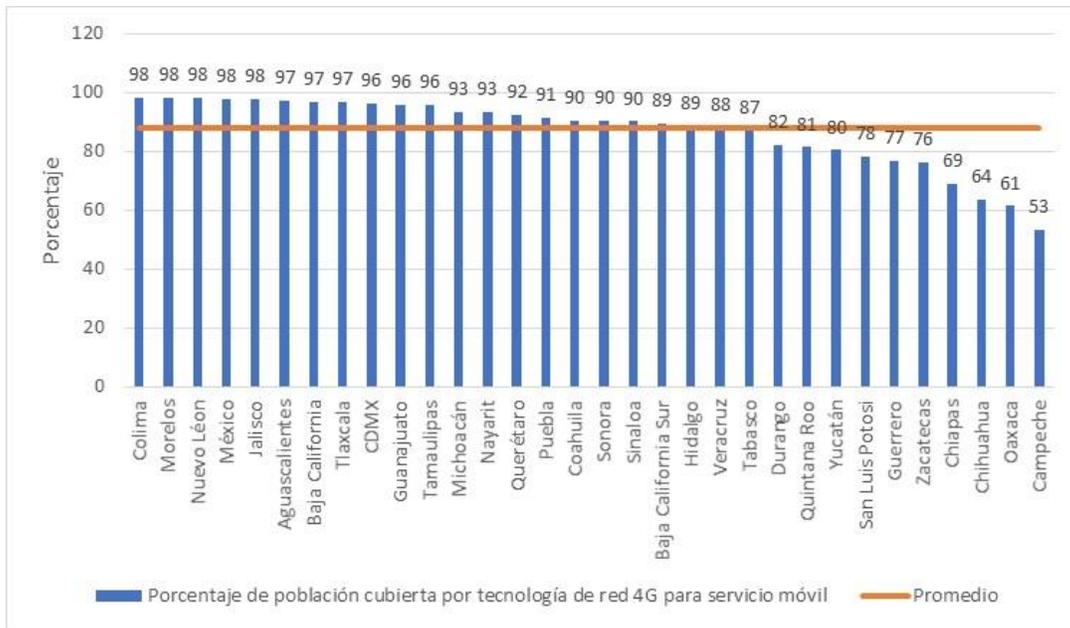
La Gráfica A6 corresponde a la variable de penetración de banda ancha fija con acceso de fibra óptica por cada 100 hogares. El primer lugar lo ocupa CDMX con 46 accesos por cada 100 hogares, seguido de San Luis Potosí con 41 accesos por cada 100 hogares y Yucatán con 38 accesos por cada 100 hogares. Oaxaca, Zacatecas y Baja California Sur presentan el mayor rezago. Por su parte el promedio nacional corresponde a 22 accesos por cada 100 hogares.

Gráfica A6. Penetración de banda ancha fija con acceso de fibra óptica



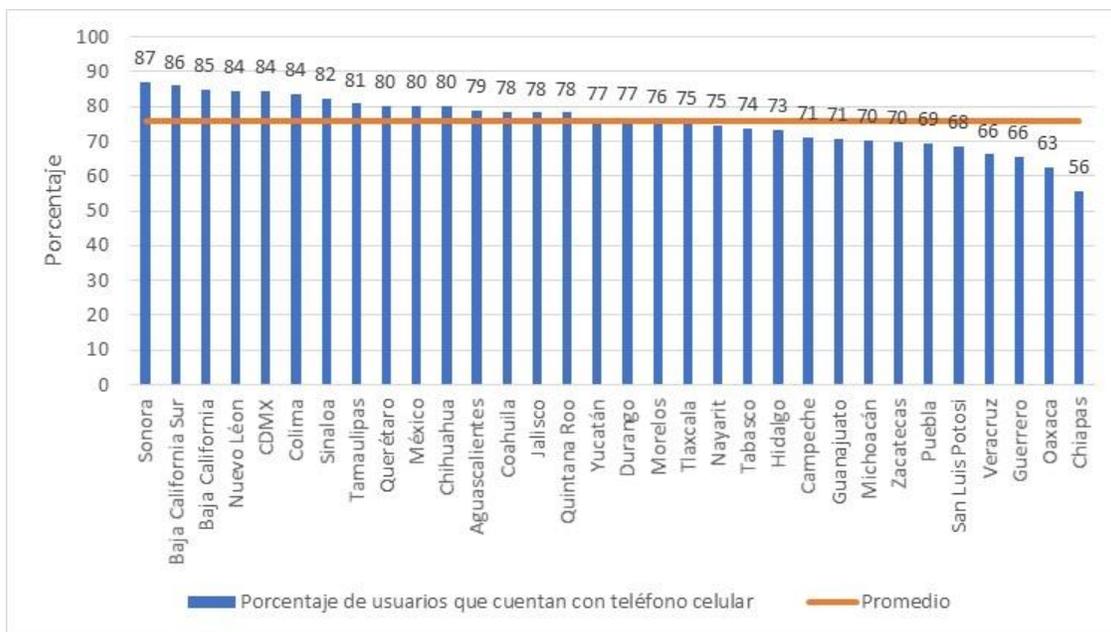
La Gráfica A7 muestra el porcentaje de población cubierta por tecnología de red 4G para servicio de telefonía móvil. El primer lugar es para Colima, Morelos y Nuevo León con 98 por ciento. La menor cobertura de 4G se registra en Oaxaca y Campeche con 61% y 53%, respectivamente. En cuanto a la media nacional ésta es igual a 88%.

Gráfica A7. Porcentaje de población cubierta por tecnología de red 4G para servicio móvil



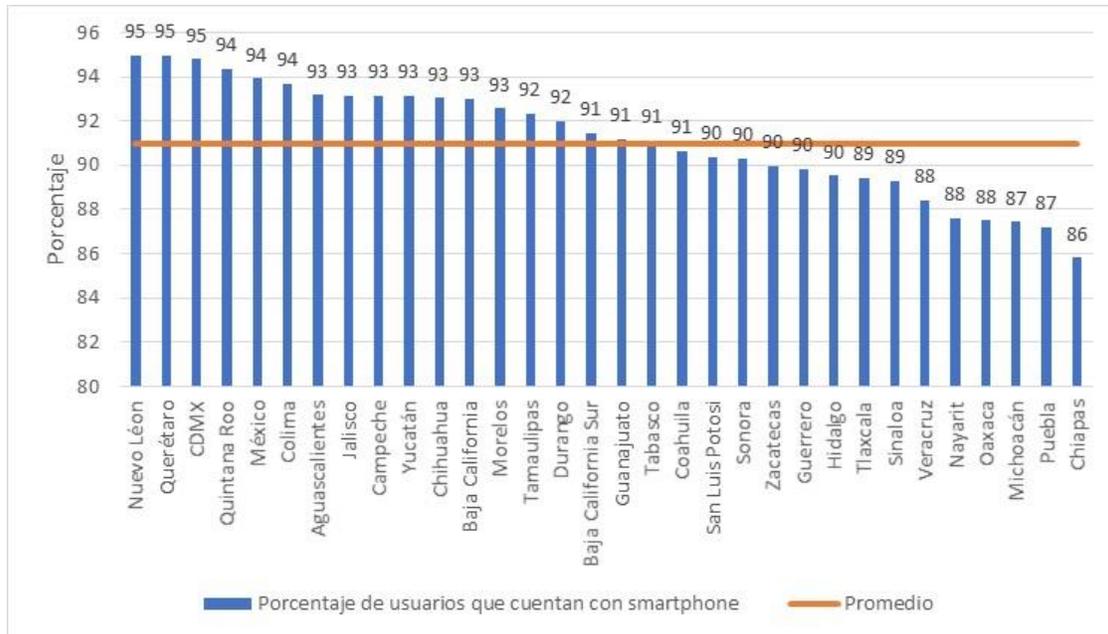
La Gráfica A8 muestra al porcentaje de usuarios que cuentan con teléfono celular siendo Sonora la entidad federativa que ocupa el primer lugar con un 88 por ciento, seguido de Baja California con un 84% y Nuevo León un 82%. En el otro extremo se encuentran Guerrero, Oaxaca y Chiapas con 66%, 63% y 56%, respectivamente. El promedio del porcentaje de usuarios a nivel nacional que cuentan con teléfono celular es del 76%.

Gráfica A8. Porcentaje de usuarios que cuentan con teléfono celular



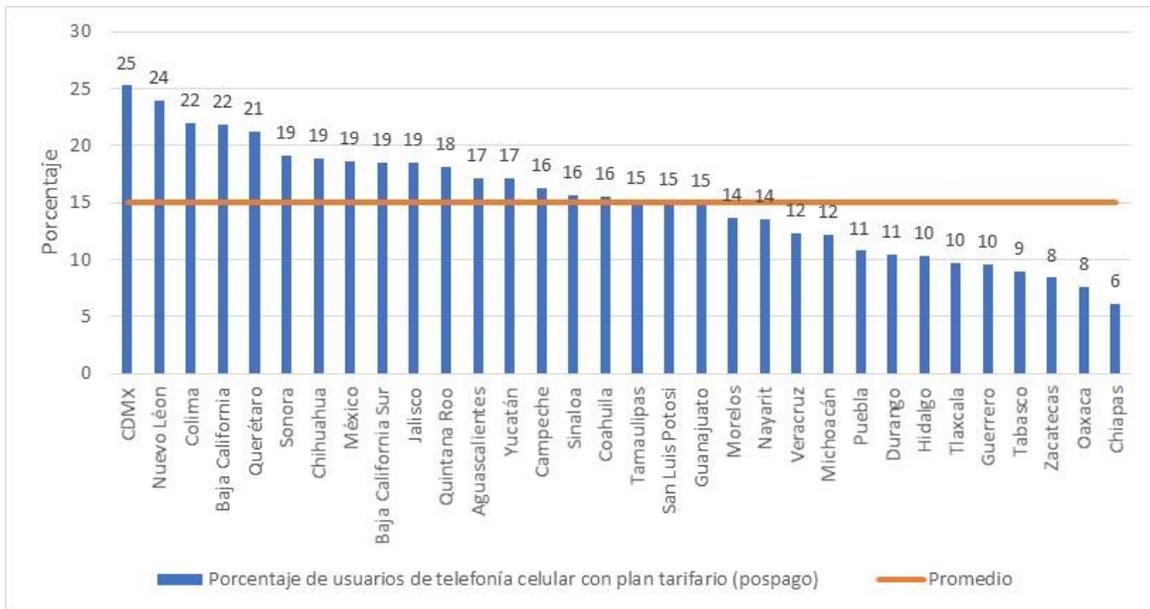
Por su parte, la gráfica A9 presenta el porcentaje de usuarios que cuentan con Smartphone o teléfono inteligente. Nuevo León, Querétaro y CDMX ocupan el primer lugar con 95 por ciento. Los puestos a la zaga los ocupan Michoacán, Puebla y Chiapas con 87%, 87% y 86%, respectivamente. A nivel nacional el promedio es de 91%.

Gráfica A9. Porcentaje de usuarios que cuentan con Smartphone



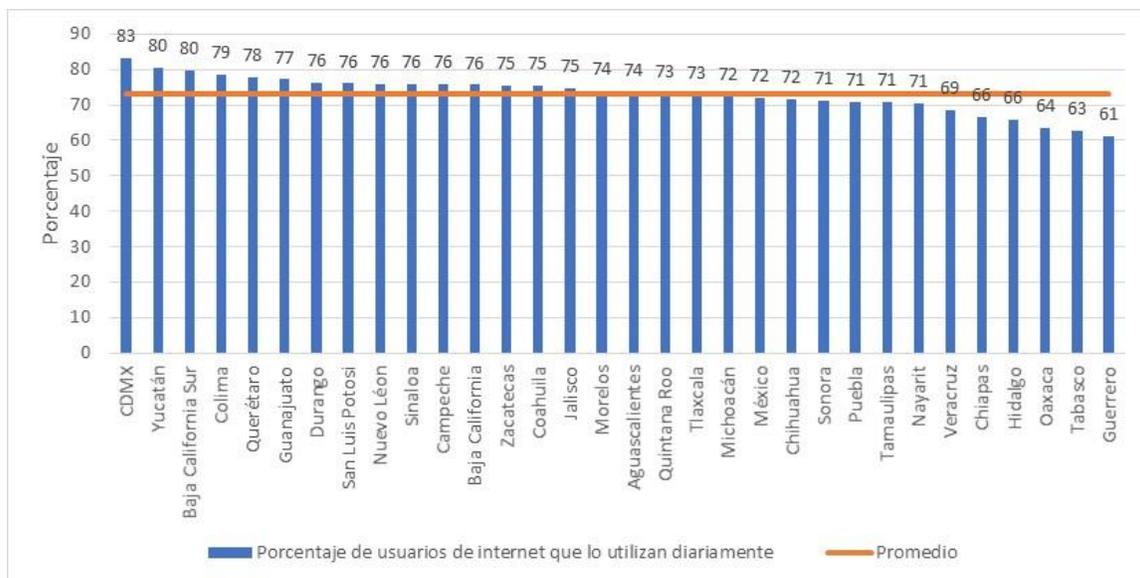
La Gráfica A10 muestra al porcentaje de usuarios de telefonía celular con plan tarifario (postpago) siendo CDMX la entidad federativa que ocupa el primer lugar con un 25 por ciento, seguido de Nuevo León con un 24%. En el otro extremo se encuentran Oaxaca y Chiapas con 8% y 6%, respectivamente. El promedio a nivel nacional es del 15%.

Gráfica A10. Porcentaje de usuarios de telefonía celular con plan tarifario (postpago)



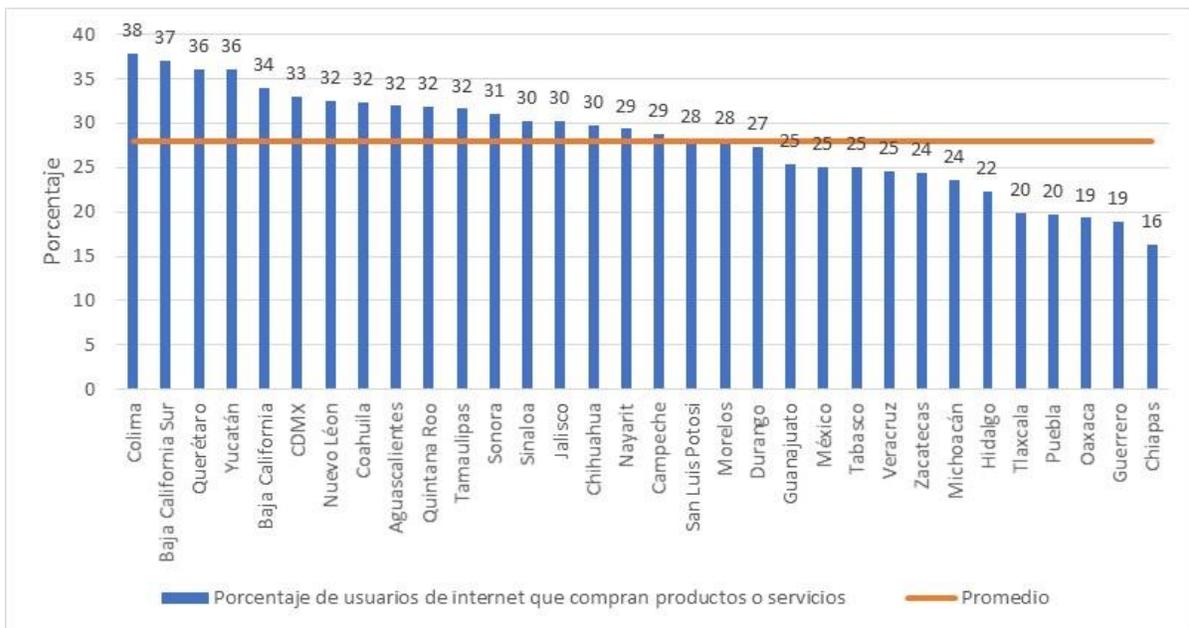
La Gráfica A11 presenta al porcentaje de usuarios de internet que lo utilizan diariamente. CDMX ocupa el primer sitio con 83 por ciento, seguido de Yucatán con un 80 por ciento. En las posiciones más rezagadas se encuentran Oaxaca, Tabasco y Guerrero con 64%, 63% y 61%, respectivamente. El promedio de los treinta y dos estados es de 73%.

Gráfica A11. Porcentaje de usuarios de internet que lo utilizan diariamente



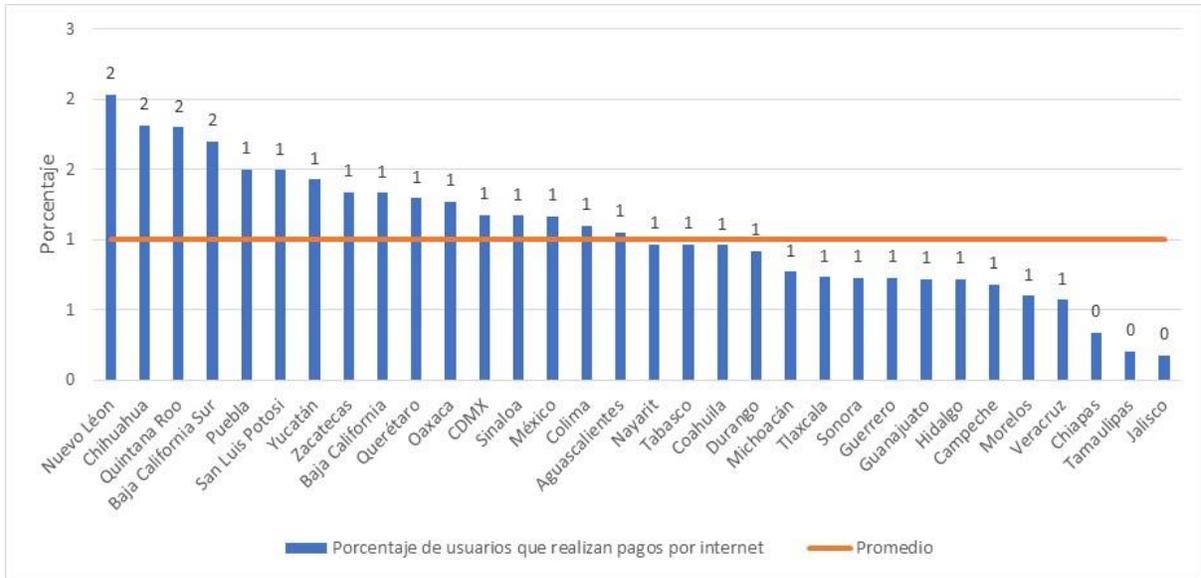
El diagnóstico sobre la inclusión digital para la actividad económica inicia con el porcentaje de usuarios que realizan compras por internet (Gráfica A12). El primer lugar es para Colima con un 38%, seguido de Baja California Sur con un 37% y Querétaro con 36%. Por su parte Oaxaca, Guerrero y Chiapas reportan los menores porcentajes con 19%, 19% y 16%, respectivamente. La media nacional es de 28%.

Gráfica A12. Porcentaje de usuarios que realizan compras por internet



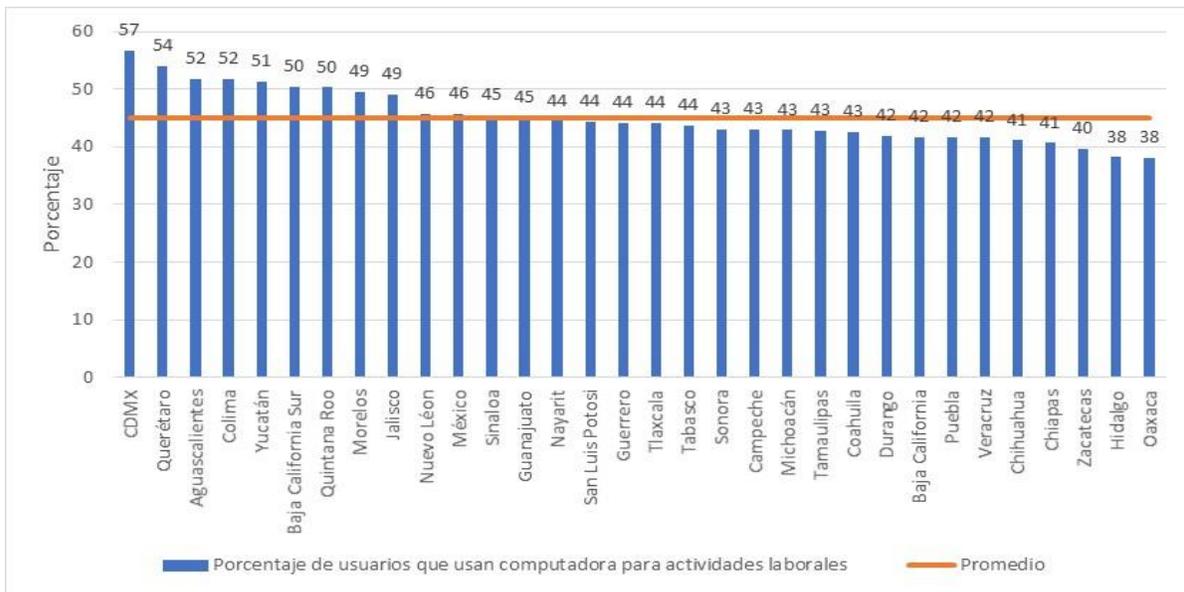
Por su parte, la Gráfica A13 muestra al porcentaje de usuarios que realizan pagos por internet con sistema de cobro digital (CoDi). En este caso los estados de Nuevo León, Chihuahua, Quintana Roo y Baja California Sur reportan el mayor porcentaje, las tres con el 2 por ciento. Los que tienen menor aplicación de pagos en línea son Chiapas, Tamaulipas y Jalisco con 0 por ciento. A nivel nacional el promedio es de 1 por ciento.

Gráfica A13. Porcentaje de usuarios que realizan pagos por internet



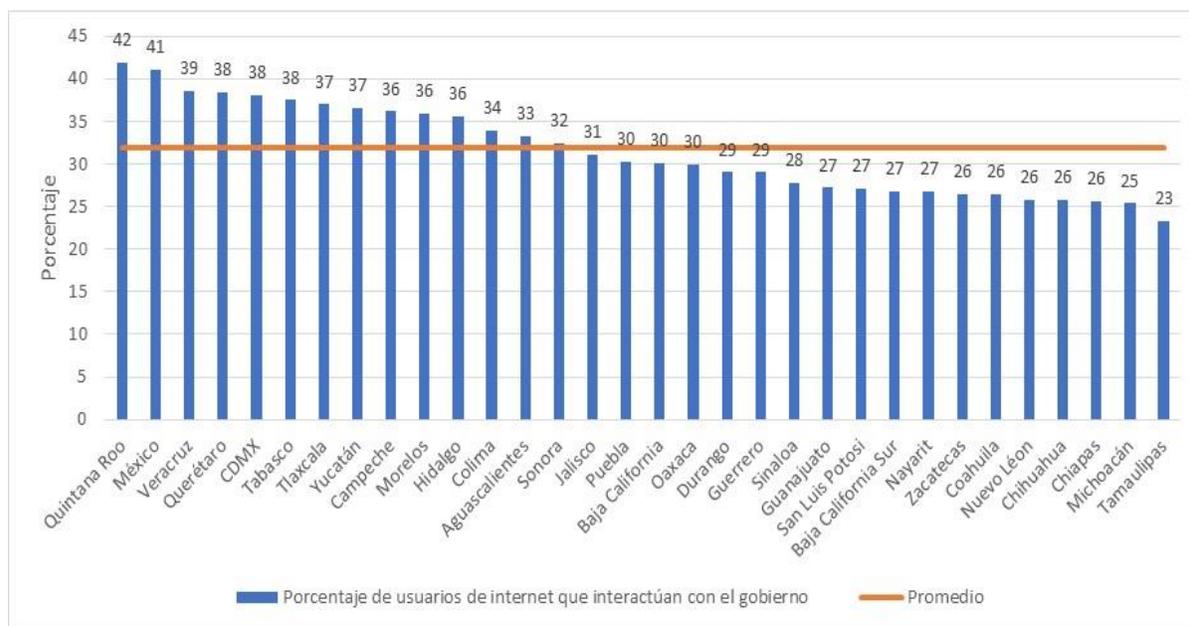
En la Gráfica A14 se observa al porcentaje de usuarios que utilizan computadora para realizar actividades laborales. El primer sitio es para CDMX con el 57 por ciento, seguido de Querétaro y Aguascalientes con 54 por ciento y 52 por ciento, respectivamente. Las posiciones más rezagadas las ocupan Zacatecas, Hidalgo y Oaxaca con 40 por ciento, 38 por ciento y 38 por ciento, respectivamente. Por su parte la media nacional se encuentra al 45%.

Gráfica A14. Porcentaje de usuarios que usan computadora para actividades laborales.



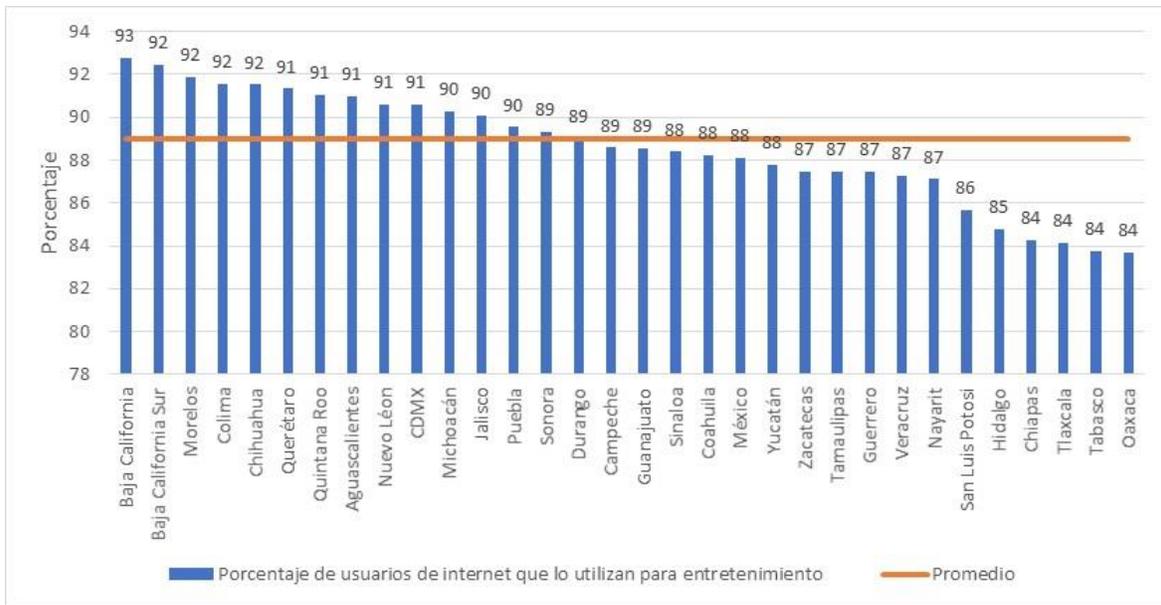
La Gráfica A15 presenta al porcentaje de usuarios de internet que realizan interacciones con el gobierno. El primer sitio es para Quintana Roo con 42 por ciento, seguido por el Estado de México con el 41 por ciento y Veracruz con el 39 por ciento. En los lugares más rezagados se encuentran Chiapas, Michoacán y Tamaulipas con 26 por ciento, 25 por ciento y 23 por ciento, respectivamente. A nivel nacional el promedio alcanza el 32 por ciento.

Gráfica A15. Porcentaje de usuarios que realizan interacciones con el gobierno



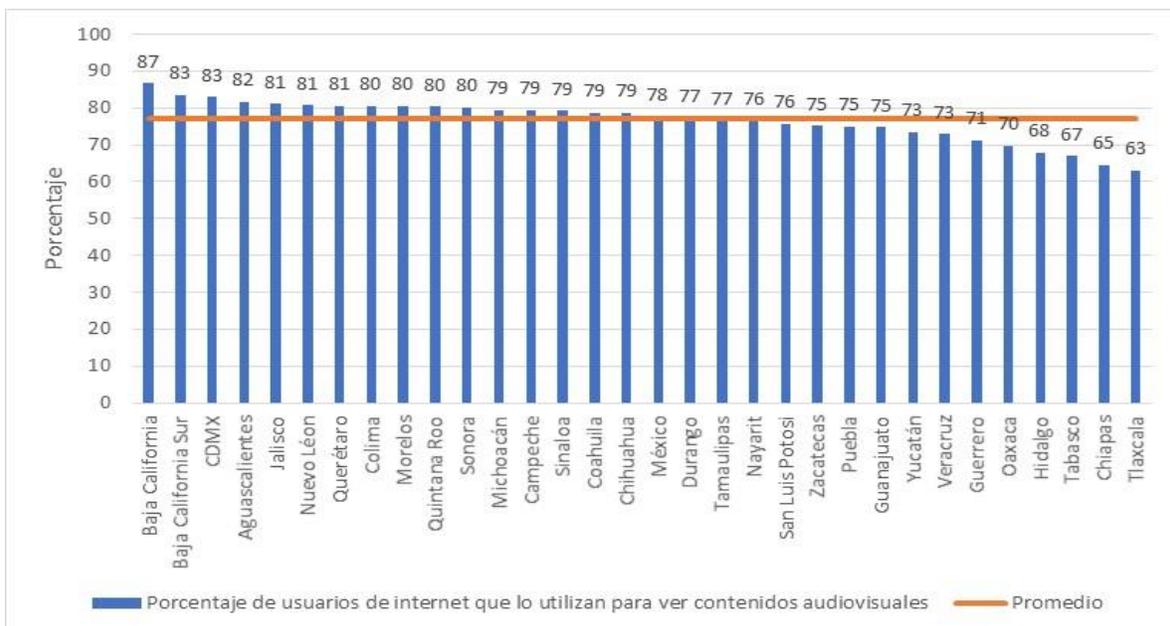
En la Gráfica A16 se observa al porcentaje de usuarios de internet que lo utilizan para entretenimiento. Esta es una modalidad de uso que cuenta ya con una elevada aceptación entre los usuarios en todas las entidades federativas. El mayor uso lo reporta Baja California con 93 por ciento, seguido Baja California Sur y Morelos con el 92 por ciento cada uno. El menor uso se reporta en Tlaxcala, Tabasco y Oaxaca con 84 por ciento. La media obtenida entre las treinta y dos entidades federativas es de 89 por ciento.

Gráfica A16. Porcentaje de usuarios internet que lo utilizan para entretenimiento



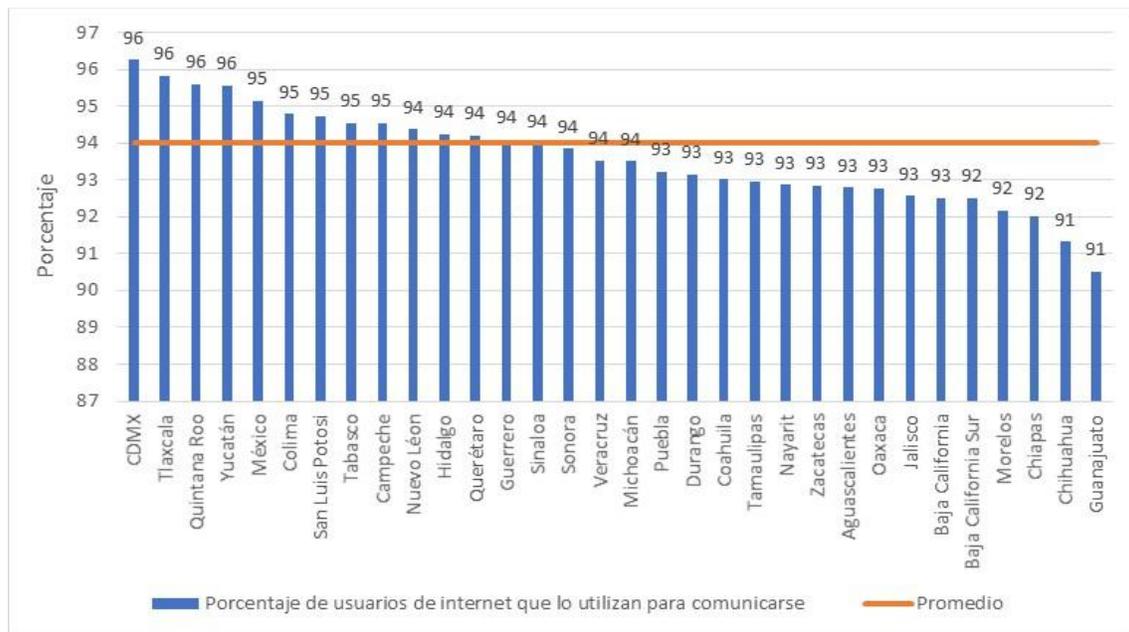
En la Gráfica A17 se observa al porcentaje de usuarios de internet que lo utilizan para ver contenidos audiovisuales. Baja California se encuentra a la cabeza con un 87 por ciento, seguido de Baja California Sur y CDMX con 83 por ciento. En el otro extremo se encuentran Chiapas y Tlaxcala con 65 por ciento y 63 por ciento, respectivamente. El promedio nacional es de un 77 por ciento

Gráfica A17. Porcentaje de usuarios de internet que lo utilizan para ver contenidos audiovisuales



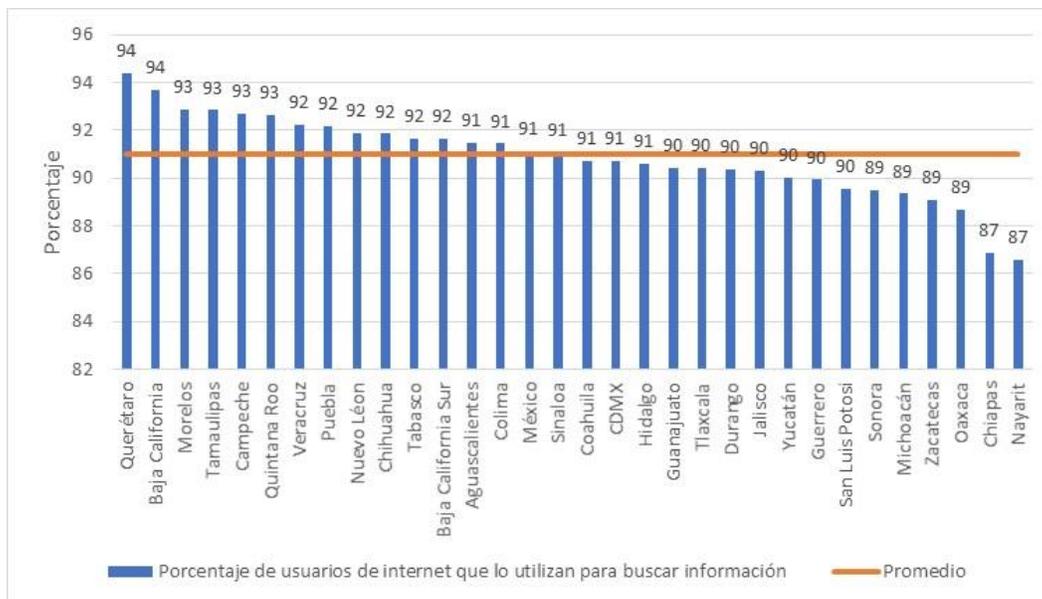
La Gráfica A18 presenta el porcentaje de usuarios de internet que lo utilizan para comunicarse. En el primer sitio se encuentra CDMX con 96 por ciento. En los sitios más rezagados se encuentran los estados de Chihuahua y Guanajuato con 91 por ciento cada uno. La media nacional corresponde a un 94%.

Gráfica A18. Porcentaje de usuarios de internet que lo utilizan para ver comunicarse



En la Gráfica A19 se muestra el porcentaje de usuarios de internet que lo utilizan para obtener información. El primer sitio es para Querétaro y Baja California con 94 por ciento. Chiapas y Nayarit reportan las menores cifras con 87 por ciento. En cuanto al promedio nacional, éste es igual a un 91 por ciento.

Gráfica A19. Porcentaje de usuarios de internet que lo utilizan para obtener información



Anexo 2

Metodología y Análisis de *clústers*.

En esta sección se presenta el análisis de clúster que es utilizado en este reporte analítico, iniciando con un gráfico Screeplot²⁶ (Gráfica A20) el cual es de utilidad para determinar el número de grupos a utilizar. La forma de interpretar la gráfica Screeplot²⁷ consiste en observar en qué parte se rompe la curva formándose una especie de “codo”, lo cual es indicativo de que la calidad del modelo no se incrementa conforme van aumentando la cantidad de clústers. Es así que la gráfica sugiere trabajar con seis clústers, tal y como lo señala el círculo de color rojo, debido a que un séptimo grupo no mejora el modelo.

Una vez determinado el número de *clústers* con el cual se va a trabajar se inicia con la aplicación de un método jerárquico conocido como el método Ward²⁸ Este método construye una estructura conocida como dendograma, en la cual muestra en qué orden se han unido los *clústers* y cuál es su grado de proximidad. En este sentido se obtienen dos dendogramas para cada *clúster* en virtud de que

²⁶ *Screeplot*. Diagrama en el cual se grafican los eigenvalores de los factores o de los componentes principales. Se utiliza para determinar el número de factores o componentes principales a mantener. Tan, P.-N., Steinbach, M., and Kumar, V. *Introduction to Data Mining*. Addison-Wesley, 2006.

²⁷ <https://campus.datacamp.com/courses/unsupervised-learning-in-r/unsupervised-learning-in-r?ex=9>

²⁸ Peña, D., 2002. Análisis de Datos Multivariantes, McGraw Hill

utilizamos dos tipos de distancias en el algoritmo para darle robustez a los resultados: la Euclidiana y la Manhattan. Por una parte, la distancia Euclidiana se refiere a que la distancia más corta que une a dos puntos es una línea recta; en tanto que la distancia Manhattan consiste en que la distancia más corta viene dada por la suma de los dos catetos de un triángulo rectángulo²⁹.

Al usar seis *clústers* (Gráficas A21 y A22) se observa que la Ciudad de México queda aislada cuando se utiliza la distancia Euclidiana; en tanto que bajo la distancia Manhattan se agrupa junto a las entidades federativas de Baja California, Baja California Sur, Quintana Roo y Nuevo León. El resto de *clústers* varían para ambas distancias, salvo para el caso del que conforman los estados de Chiapas, Guerrero y Oaxaca. Otros dos grupos son muy similares, y sólo se diferencian porque se intercambia el estado de Veracruz. Estos grupos son los que conforman bajo la distancia Euclidiana, por una parte: Campeche, Durango, Guanajuato, Michoacán, Nayarit, San Luis Potosí y Zacatecas; y, por otra, Hidalgo, Puebla, Tabasco, Tlaxcala y Veracruz.

Para dar robustez al análisis técnico, se sustituye la técnica jerárquica por la técnica de partición mediante el algoritmo de K-Medias repitiendo la estimación para seis *clústers* (Gráfica A23). La idea básica detrás del algoritmo de K-medias consiste en definir los agrupamientos de tal manera que la variación intra-grupo sea minimizada. El algoritmo de K-medias se puede resumir de la siguiente forma:

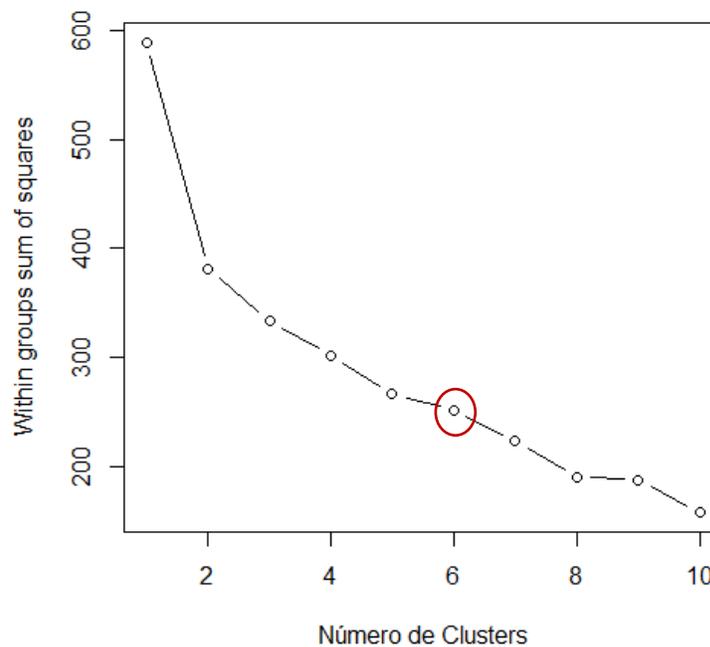
- i. Se especifica el número de *clústers* (k) que serán utilizados.
- ii. A partir de los datos se seleccionan de forma aleatoria k puntos como los centros iniciales (centroide³⁰) o medias.
- iii. Cada observación es asignada al centroide más cercano, basado en la distancia Euclidiana, o Manhattan, entre el punto y el centroide.
- iv. Para cada uno de los k *clústers* se actualiza el centroide del *clúster* por medio del cálculo del nuevo promedio para todos los puntos en el *clúster*. El centroide de un clúster K es un vector de longitud p que contiene las medias de todas las variables para las observaciones del K -ésimo clúster, p se refiere al número de variables.
- v. Se minimiza de forma iterativa el total de la suma de cuadrados. Esto es, se iteran los pasos iii y iv hasta que el *clúster* deje de cambiar o se alcance el número máximo de iteraciones.

²⁹ <https://xlinux.nist.gov/dads/HTML/manhattanDistance.html>

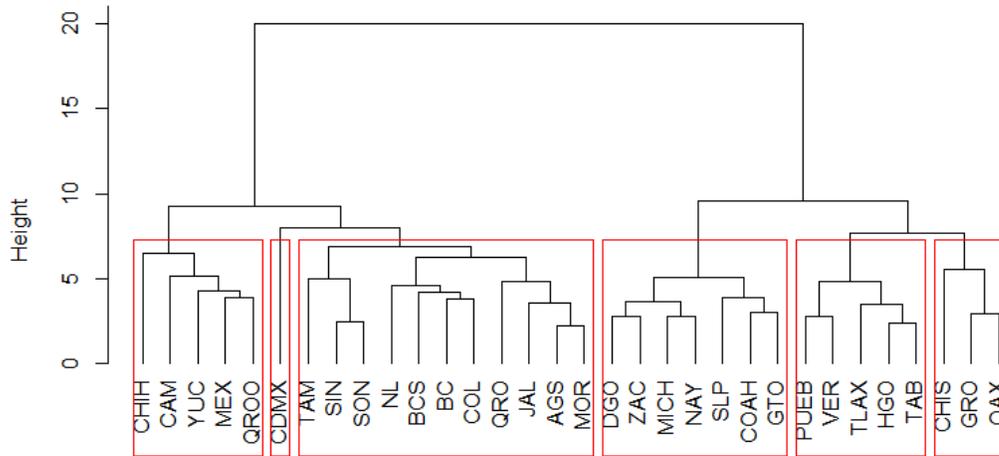
³⁰ El centroide de un clúster se define como el punto equidistante de los objetos pertenecientes a dicho clúster. Tan, P.-N., Steinbach, M., and Kumar, V. *Introduction to Data Mining*. Addison-Wesley, 2006.

Aplicando este método para definir seis *clústers* (Gráfica A23) se observa que el primer grupo está conformado exclusivamente por la Ciudad de México. En el segundo grupo se encuentran Baja California, Baja California Sur, Quintana Roo, Nuevo León y Sonora. En el tercer grupo se encuentra únicamente Campeche. En tanto que el cuarto grupo tiene a Aguascalientes, Coahuila, Colima, Chihuahua, Jalisco, Estado de México, Morelos, Querétaro, Sinaloa, Tamaulipas y Yucatán. El quinto *clúster* tiene a los estados de Durango, Guanajuato, Michoacán, Nayarit, San Luis Potosí, Veracruz y Zacatecas. Por último, en el sexto grupo están Chiapas, Guerrero, Hidalgo, Oaxaca, Puebla, Tabasco y Tlaxcala.

Los resultados se resumen en el Cuadro A1. El ejercicio estadístico completo considera que seis es el número óptimo de *clústers* a considerar.

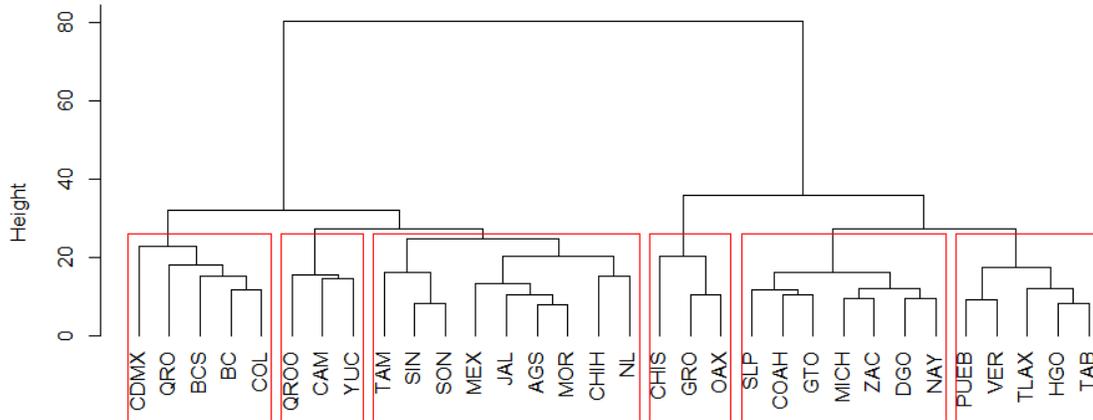


Cluster Dendrogram

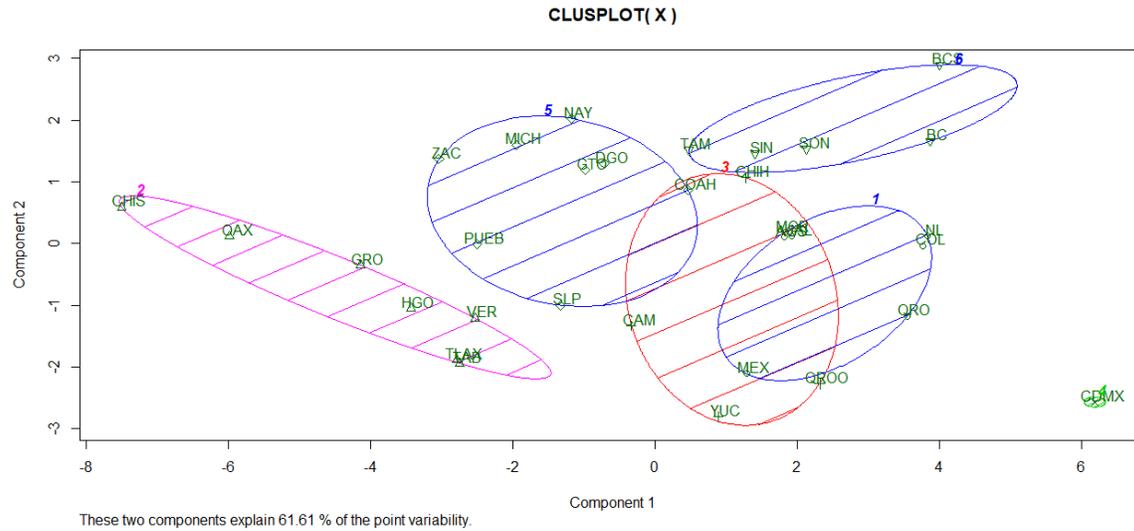


d
hclust(*, "ward.D2")

Cluster Dendrogram



d
hclust(*, "ward.D2")



Metodología aplicada al diseño del índice de desarrollo digital

El diseño del índice de bienestar digital conlleva la aplicación de la técnica de los componentes principales³¹, la cual es de gran utilidad cuando se trabaja con bases de datos multivariadas, como lo es en este caso. El análisis de los componentes principales va a permitir observar la “forma” de los datos en su totalidad.

En suma, los componentes principales son un tipo de transformación lineal de un set de datos, el cual ajusta dichos datos a un nuevo sistema de coordenadas, de tal forma que la parte más significativa de la varianza se encuentra en la primera coordenada, y cada coordenada subsecuente es ortogonal a la anterior y tiene una menor varianza. De esta forma, lo que se hace es transformar a un set de x

³¹ Peña, D., 2002. Análisis de Datos Multivariantes, McGraw Hill.

variables correlacionadas sobre y sujetos (en este caso las treinta y dos entidades federativas) en un set de p componentes principales no correlacionados sobre los mismos sujetos.

En virtud de que el set de datos cuenta con 19 indicadores se puede decir que éste tiene 19 dimensiones. Por lo tanto, el número de eigenvectores y sus eigenvalores asociados será también de 19 ya que siempre es igual al número de dimensiones del set de datos. Ahora, un eigenvector es algo similar a una dirección, en tanto que un eigenvalor es un número que dice cuánta varianza existe en los datos en dicha dirección. Es así que el eigenvector con el eigenvalor más alto será el primer componente principal.

En función de lo anterior es que el diseño del índice de desarrollo digital implica la ponderación de cada una de las cuatro dimensiones y su suma para obtener el resultado final, el cual se presentó en el Cuadro 3 del reporte.