

# Metodología Índice de Progreso Social Medellín

Samuel Agudelo Gamboa - Mateo Tuberquia Giraldo  
Universidad Nacional de Colombia, Sede Medellín

Noviembre, 2020

## Resumen

El Índice de Progreso Social fue creado para medir el éxito de los países más allá de los indicadores económicos como el PIB. El progreso social se trata de la satisfacción de las necesidades básicas como alimento, agua potable, vivienda y seguridad; de vivir vidas largas y saludables, de proteger el medio ambiente, de libertad y oportunidades. Cada vez más, el progreso social se ha convertido en el punto crítico en las agendas de los líderes de gobierno, sector privado y la sociedad civil, por esta razón impulsar el progreso social requiere de un nuevo modelo que permita medir sus niveles. Aquí se explica detalladamente la metodología del cálculo del IPS para la dimensión FUNDAMENTOS DEL BIENESTAR en la ciudad de Medellín

## 1. Introducción

El índice de Progreso Social (IPS) es un modelo holístico que busca medir el desempeño social de un país. Es holístico porque entiende que las condiciones de vida de una persona no pueden ser explicadas solo a partir de una dimensión por el contrario se deben considerar distintos aspectos que se encuentran interrelacionados, por ejemplo, el acceso a servicios básicos y la salud infantil.

El IPS entiende el progreso social como la 'Capacidad de una sociedad para satisfacer las necesidades humanas fundamentales de sus ciudadanos, establecer los elementos básicos que permitan a los ciudadanos y comunidades mejorar y mantener su calidad de vida, y crear las condiciones para que todas las personas alcancen su pleno potencial' SPI 2014, por consiguiente, un país debe tener la capacidad de brindarle a sus ciudadanos el acceso a servicios básicos, darle las herramientas para su propio desarrollo; para lo cual es necesario un contexto de respeto a las libertades

y derechos de sus ciudadanos.

Este informe describe y analiza la metodología utilizada para calcular el índice de Progreso Social en la ciudad de Medellín para la dimensión FUNDAMENTOS DEL BIENESTAR para 2019 dado que todos sus índices provienen de la Encuesta de Calidad de Vida. Comienza describiendo los principios que establecen la arquitectura conceptual del índice, y brinda una descripción general del marco de este. Se detalla y analiza los pasos tomados para seleccionar datos y calcular el índice. Finalmente, se describen las limitaciones y las mejoras respecto a los datos utilizados para el cálculo. Todos los cálculos se realizan usando el software R en su versión 4.0.3.

## 2. Principios Progreso Social

El modelo conceptual del índice de Progreso Social (IPS) propone medir y entender el progreso social a partir de tres dimensiones:

1. ¿Tienen todos cubiertas las necesidades de supervivencia: comida, agua, refugio, seguridad?.
2. ¿Tienen todos acceso a los bloques de construcción de la mejora de sus vidas: educación, información, salud y ambiente sustentable?
3. ¿Tienen todos acceso a la posibilidad de cumplir sus objetivos, sueños y ambiciones sin obstáculos?.

De este modo, los 12 componentes forman el marco de Progreso Social y para cada uno de estos componentes se tienen indicadores para medir el rendimiento en una sociedad, puesto que no son indicadores de esfuerzo o intención, sino de logro verdadero. Con este índice no se mide el gasto de una sociedad en sanidad, se mide longevidad y calidad de vida de las personas, no mide si los gobiernos aprueban leyes contra la discriminación, mide si las personas son discriminadas.



Figura 1: Marco del Progreso Social

Estos elementos interrelacionados se combinan para producir un nivel determinado de progreso social. La metodología del índice de Progreso Social permite medir cada componente y cada dimensión, produciendo una puntuación y clasificación general.

## **2.1. Niveles de Progreso Social**

El Índice de Progreso Social toma valores que van desde 0, siendo el menor nivel posible, hasta 100, evidenciando el mayor nivel posible. De acuerdo con esta escala, la metodología establecida por el IPS define seis niveles de desempeño. Los territorios con puntajes mayores o iguales a 85 tienen un nivel de Progreso Social muy alto, los que están entre 75 puntos y 85 puntos, un nivel de Progreso Social alto; los que están entre 65 y 75 puntos, un nivel de Progreso Social Medio alto; entre 55 y 65, medio bajo; entre 45 y 55, bajo; y por debajo de 45, muy bajo.

## **3. IPS en Medellín**

Desde el año 2010, Medellín cuenta con el Índice Multidimensional de Condiciones de Vida (IMCV), esta una medida que evalúa las características de los hogares, evaluadas mediante la Encuesta de Calidad de Vida en quince dimensiones: Entorno y calidad de la vivienda, acceso a servicios públicos, medio ambiente, escolaridad, desescolarización, movilidad, capital físico del hogar, participación, libertad y seguridad, vulnerabilidad, salud, trabajo, recreación, ingresos y percepción de la calidad de vida.

Este índice ha permitido evidenciar las diferencias en materia de calidad de vida entre comunas y su evolución (Medellín Cómo Vamos, 2017), mostrando mejores condiciones de calidad de vida en las comunas de El Poblado, Laureles-Estadio y La América, y condiciones relativamente peores en tres de las comunas de la zona nororiental de la ciudad, específicamente en Popular, Santa Cruz y Manrique, en Villa Hermosa, que hace parte de la zona centro-oriental, en El Doce de Octubre, que pertenece a la zona noroccidental de la ciudad, y en San Javier en el centro-occidente. Asimismo, el IMCV ha mostrado que pese al avance de las condiciones de vida promedio en los corregimientos, en los últimos años, ninguno de ellos logra superar el valor promedio del IMCV de Medellín en todo el periodo 2010-2017 (IMCV, 2018).

El IMCV es muy útil para identificar las diferencias en las condiciones de vida, de las comunas de la ciudad. En efecto, dentro de los análisis que realiza Medellín Cómo Vamos, la revisión de las brechas entre comunas de acuerdo con el IMCV constituye la principal herramienta para identificar desigualdades territoriales. Sin embargo, carece de dos atributos esenciales. En primer lugar, aunque está construido con una metodología sumamente rigurosa, no permite una lectura sencilla y directa de cuales son los fenómenos a los que debe prestárseles atención en cada territorio. En segundo lugar, el IMCV incluye exclusivamente variables provenientes de la Encuesta de Calidad de Vida, por lo cual deja por fuera indicadores clave como, por ejemplo, las incidencias de enfermedades o la tasa de homicidios en la comuna, lo que va en contra de su accionabilidad.

En ese entorno, el Índice de Progreso Social constituye, más que una medida que pueda reemplazar al IMCV, un indicador complementario a esta medida multidimensional de bienestar, aportando al conocimiento de la calidad de vida en la ciudad, la claridad y la accionabilidad como elementos centrales.

### **3.1. Dimensiones**

En el nivel más alto del marco, se sintetiza tres preguntas distintas relacionadas que, tomadas en conjunto, ofrecen información sobre el nivel de progreso social:

- ¿Están satisfechas las necesidades más esenciales de la sociedad?  
La dimensión de Necesidades Humanas Básicas entrega información que pretende responder a esta pregunta. Procura evidenciar el grado en el que una sociedad puede ofrecer las condiciones básicas para la subsistencia de los ciudadanos. Dichas necesidades incluyen nutrición adecuada y cuidado médico básico, saneamiento, refugio adecuado y seguridad personal
- ¿Existen los elementos fundamentales para que las personas y las comunidades puedan mantener y aumentar su bienestar?  
La dimensión de Fundamentos del Bienestar busca responder a esta pregunta. Cubre temas como la posibilidad de los ciudadanos de acceder a educación básica, acceder a información que les permita desplegar todo su potencial, beneficiarse de un sistema de salud moderno y vivir en un ambiente saludable que asegure una vida prolongada.

- ¿Existen oportunidades para que todas las personas puedan lograr su pleno potencial tomando decisiones autónomas?

La dimensión Oportunidades busca responder a esta pregunta. Incluye indicadores asociados a los derechos personales, la libertad personal y de elección, la existencia de un ambiente de tolerancia e inclusión, y el acceso a educación avanzada. Esta dimensión es la que ofrece mayores dificultades en la definición y cálculo de indicadores apropiados.

El índice se construye como el puntaje promedio de las tres dimensiones, con lo cual las tres dimensiones que lo componen tienen el mismo peso relativo (33.3 %)

### 3.2. Componentes

Cada dimensión del índice de Progreso social incluye cuatro componentes, que, a su vez, están compuestos por entre 3 y 5 indicadores. Cada componente en una dimensión se diseña para resaltar un aspecto distinto de los resultados que dan lugar a una dimensión, y fue contruido tanto a partir de la literatura académica como de literatura empleada para la formulación de políticas públicas.

Los indicadores dentro de cada componente son escogidos de acuerdo con la disponibilidad de información y con tres criterios:

- Capturar lo que se propone medir
- Estar disponible como información pública
- Existir información para todas las comunas y corregimientos

## 4. Selección indicadores

En el nivel más granular del marco del índice de Progreso Social, se identifican múltiples medidas de resultados independientes (indicadores) relacionados con cada componente. Cada conjunto de indicadores, agrupados por componentes, define y mide el mismo aspecto del progreso social. Según la disponibilidad de datos y la investigación en curso sobre los resultados sociales, los indicadores pueden cambiar con cada edición del índice de Progreso Social. Sin embargo, los conceptos capturados por cada conjunto de indicadores (es decir, componentes) siguen siendo los mismos. El índice de Progreso Social 2020 incluye 43 indicadores, con 3-4 indicadores por componente (ver figura 2), tomados en su mayoría de la Encuesta de Calidad de vida de Medellín, aplicado anualmente por la administración municipal a las 16 comunas y cinco corregimientos.

| INDICADORES PROGRESO SOCIAL EN MEDELLÍN  |   |   |
|--|---|---|
| <b>Necesidades Humanas Básicas</b>   | <b>Fundamentos de Bienestar</b>   | <b>Oportunidades</b>  |
| <b>Nutrición y cuidados básicos de salud</b> <ul style="list-style-type: none"> <li>* Mortalidad 5 años</li> <li>** Inseguridad alimentaria severa</li> <li>* Incidencia de bajo peso al nacer a término</li> <li>* Inseguridad alimentaria (cualquier grado)</li> </ul> | <b>Acceso al conocimiento básico</b> <ul style="list-style-type: none"> <li>* Asistencia escolar 5 - 11 años</li> <li>* Asistencia escolar 12 - 15 años</li> <li>* Asistencia escolar 16 - 17 años</li> <li>* Porcentaje de analfabetismo mayores de 15 años</li> </ul> | <b>Derechos personales</b> <ul style="list-style-type: none"> <li>* Libertad de expresar opiniones</li> <li>* Libertad de movimiento</li> <li>* Derechos de propiedad</li> <li>* Derechos de propiedad</li> </ul> |
| <b>Agua y Saneamiento Básico</b> <ul style="list-style-type: none"> <li>* Calidad del acueducto</li> <li>* Cobertura de alcantarillado</li> <li>* Calidad de alcantarillado</li> </ul>   | <b>Acceso a la información y comunicaciones</b> <ul style="list-style-type: none"> <li>* Conectividad a internet</li> <li>* Uso de teléfonos celulares</li> <li>* Conectividad a televisión por cable</li> </ul>  | <b>Libertad personal y de elección</b> <ul style="list-style-type: none"> <li>* Vulneración de derechos a infancia</li> <li>* Embarazo adolescente</li> <li>* Desplazamiento forzado intraurbano</li> </ul>       |
| <b>Vivienda y servicios públicos</b> <ul style="list-style-type: none"> <li>* Calidad de la energía eléctrica</li> <li>* Porcentaje viviendas con buena calidad aseo</li> <li>* Cobertura de gas natural domiciliario</li> </ul>   | <b>Salud y Bienestar</b> <ul style="list-style-type: none"> <li>* Satisfacción calidad de vida</li> <li>* Satisfacción acceso a salud</li> <li>* Hábitos saludables</li> <li>* Satisfacción con la calidad del servicio de salud</li> </ul>                             | <b>Tolerancia e inclusión</b> <ul style="list-style-type: none"> <li>* Violencia intrafamiliar</li> <li>* Discriminación</li> <li>* Convivencia</li> <li>* Equidad en el acceso a educación superior</li> </ul>   |
| <b>Seguridad Personal</b> <ul style="list-style-type: none"> <li>Tasa de homicidios</li> <li>Mortalidad por accidentes de tránsito</li> <li>Hurto a personas</li> </ul>  | <b>Sostenibilidad ambiental</b> <ul style="list-style-type: none"> <li>* Calidad del aire</li> <li>* Nivel de ruido</li> <li>* Arborización</li> <li>* Basuras en las calles</li> </ul>   | <b>Acceso a la educación superior</b> <ul style="list-style-type: none"> <li>* Acceso a educación superior</li> <li>* Deserción de educación superior</li> <li>* Asistencia 18-24 años</li> </ul>                 |

Figura 2: Indices por componente

## 5. Cálculo índice

Hay cuatro pasos básicos para calcular el índice de Progreso Social:

- Calcular los indicadores seleccionados en cada componente con las fuentes de información
- Invertir y estandarizar los indicadores para que sean comparables.
- Usar Análisis de Componentes Principales(ACP) para agregar indicadores en un puntaje.
- Calcular las puntuaciones de las dimensiones y del índice de Progreso Social general promediando componentes y dimensiones, respectivamente.

En el presente informe se describe la metodología del IPS de Medellín para la dimensión **Fundamentos del Bienestar** en 2019 usando datos desde 2013 dado que todos sus indicadores se extraen de la Encuesta de Calidad de Vida (ver figura 3). A continuación se mostrará el procedimiento a detalle y como se usa la herramienta R para tal fin.

| Dimensión                 | Componente                               | Variable  | Indicador  | Fuente                      | Tiempo | Relación con IPS |
|---------------------------|--|---|--|-----------------------------|--------|------------------|
| Fundamentos del Bienestar | Acceso al conocimiento básico            | Asistencia escolar 5-11 años                                  | Porcentaje de niños entre 5 y 11 años que asisten a alguna institución educativa formal, de acuerdo con ECV                    | Encuesta de Calidad de Vida | T      | Directa          |
|                           |  | Asistencia escolar 12-15 años                                 | Porcentaje de niños entre 12 y 15 años que asisten a alguna institución educativa formal, de acuerdo con ECV                   | Encuesta de Calidad de Vida | T      | Directa          |
|                           |  | Asistencia escolar 16-17 años                                 | Porcentaje de niños entre 16 y 17 años que asisten a alguna institución educativa formal, de acuerdo con ECV                   | Encuesta de Calidad de Vida | T      | Directa          |
|                           |  | Porcentaje de analfabetismo en personas mayores de 15 años    | Porcentaje de personas de quince años o más que no saben leer o escribir más de un párrafo                                     | Encuesta de Calidad de Vida | T      | Inversa          |
|                           | Acceso a la información y comunicaciones | Conectividad a internet                                       | Hogares que afirman tener servicio de internet como porcentaje del total de hogares.   | Encuesta de Calidad de Vida | T      | Directa          |
|                           |  | Uso de teléfonos celulares                                    | Hogares que afirman tener al menos un teléfono celular en el hogar como porcentaje del total de hogares.                       | Encuesta de Calidad de Vida | T      | Directa          |
|                           |  | Conectividad a televisión por cable                           | Hogares que afirman tener servicio de tv por suscripción como porcentaje del total de hogares.                                 | Encuesta de Calidad de Vida | T      | Directa          |
|                           | Salud y Bienestar                        | Satisfacción con la calidad de vida de los miembros del hogar | Porcentaje de jefes de hogar que consideran que en los miembros de su hogar tienen un nivel malo o muy malo de calidad de vida | Encuesta de Calidad de Vida | T      | Inversa          |
|                           |  | Satisfacción con el acceso a servicios de salud               | Porcentaje de jefes de hogar que consideran malo o muy malo el acceso a salud en los últimos doce meses                        | Encuesta de Calidad de Vida | T      | Inversa          |
|                           |  | Hábitos saludables  | Porcentaje de jefes de hogar que afirman practicar actividades deportivas  | Encuesta de Calidad de Vida | T      | Directa          |
|                           |  | Satisfacción con la calidad del servicio de salud             | Porcentaje de jefes de hogar que consideran malo o muy malo el servicio de salud en los últimos doce meses                     | Encuesta de Calidad de Vida | T      | Inversa          |
| Fundamentos del Bienestar | Sostenibilidad ambiental                 | Calidad del aire  | Porcentaje de jefes de hogar insatisfechos con el aire en su barrio o vereda   | Encuesta de Calidad de Vida | T      | Inversa          |
|                           |  | Nivel de ruido  | Porcentaje de jefes de hogar insatisfechos con el ruido en su barrio o vereda  | Encuesta de Calidad de Vida | T      | Inversa          |
|                           |  | Arborización  | Porcentaje de jefes de hogar insatisfechos con la arborización en su barrio o vereda   | Encuesta de Calidad de Vida | T      | Inversa          |
|                           |  | Basuras en las calles   | Porcentaje de jefes de hogar insatisfechos con el estado de escombros y basuras en las calles                                  | Encuesta de Calidad de Vida | T      | Inversa          |

Figura 3: Detalle índices Fundamentos del Bienestar

## 5.1. Cálculo índices de componentes

Dado que los índices en la dimensión de **Fundamentos del Bienestar** se extraen de la Encuesta de Calidad de Vida; se mostrará como abordar estos datos para llegar al conjunto de índices que posteriormente se usan para el cálculo de esta dimensión.

La Encuesta de Calidad de vida de Medellín es una encuesta anual que consta de 342 preguntas las cuales cubren 15 dimensiones sociales. Esta encuesta es realizada tomando una muestra representativa de cada barrio de las 16 comunas y 5 corregimientos de Medellín, con el objetivo de realizar un seguimiento y medición de las condiciones socio-económicas de la ciudad, entre las dimensiones que

aborda la encuesta se encuentra el entorno y calidad de la vivienda, acceso a servicios públicos, medio ambiente, escolaridad, desescolarización, movilidad, capital físico del hogar, participación, libertad y seguridad, vulnerabilidad, salud, trabajo, recreación, ingresos y percepción de la calidad de vida. A continuación se muestra como se hace la lectura de los datos además de como se extraen los hogares únicos entrevistados en las encuestas ya que estos se usan para el cálculo de algunos índices.

```
#=====
#   IPS MEDELLIN - FUNDAMENTOS DEL BIENESTAR
#=====

# Librerías que se utilizan
require(dplyr)
require(tidyverse)
require(psych)

# Lectura encuesta de calidad de vida
ecv <- read.csv("Encuesta_Calidad_Vida.csv", sep = ",", encoding = "UTF-8")

# Se leccionan los años necesarios para el calculo
ecv <- subset(ecv, YEAR >= 2013 & YEAR <= 2019)

# Se convierte la variable COMUNA en una categorica (16 comunas, 5 corregimientos)
ecv$P_6 <- as.factor(ecv$P_6)

# Hogares únicos entrevistados (FORM: indicador punico del hogar entrevistado)
ecv_hogares <- ecv[!duplicated(ecv$Form),]
```

La dimensión de **Fundamentos del Bienestar** incluye los indicadores que se resumen en la tabla. En total, esta dimensión incluye 15 indicadores en cuatro componentes. Entre ellos, siete están invertidos (a mayor valor del indicador, representa una situación de menor Progreso Social) y los demás se toman en su sentido original (ver Relación con IPS Figura 3).

A continuación se muestra el procedimiento para calcular cada uno de los 15 índices en cada componente:



Se construyen 4 conjuntos de datos donde cada uno se forma por los componentes en la dimensión y se les asigna un nombre, como se muestra en la siguiente tabla:

| Componente                               | Nombre Asignado |
|--|-----------------|
| Acceso al conocimiento básico            | db_BA           |
| Acceso a la información y comunicaciones | db_BB           |
| Salud y Bienestar                        | db_BC           |
| Sostenibilidad ambiental                 | db_BD           |

### • COMPONENTE ACCESO AL CONOCIMIENTO BÁSICO

```
#-----
# Acceso al conocimiento básico
#-----

# Asistencia escolar 5-11 años
Ind_b1 <- ecv_2019 %>% filter(P_18 >= 5 & P_18 <= 11) %>% group_by(P_6) %>%
  dplyr::summarize(Ind_B1 = sum(P_36 == 1)/sum(table(P_36))) %>%
  ungroup()

# Asistencia escolar 12-15 años
Ind_b2 <- ecv_2019 %>% filter(P_18 >= 12 & P_18 <= 15) %>% group_by(P_6) %>%
  dplyr::summarize(Ind_B2 = sum(P_36 == 1)/sum(table(P_36))) %>%
  ungroup()

# Asistencia escolar 16-17 años
Ind_b3 <- ecv_2019 %>% filter(P_18 >= 16 & P_18 <= 17) %>% group_by(P_6) %>%
  dplyr::summarize(Ind_B3 = sum(P_36 == 1)/sum(table(P_36))) %>%
  ungroup()

# Porcentaje de analfabetismo en personas mayores de 15 años
Ind_b4 <- ecv_2019 %>% filter(P_18 > 15) %>% group_by(P_6) %>%
  dplyr::summarize(Ind_B4 = sum(P_35 == 2)/sum(table(P_35))) %>%
  ungroup()

# Indices del componente Acceso al conocimiento básico por comuna
db_BA <- cbind(Ind_b1, Ind_b2[, -1], Ind_b3[, -1], Ind_b4[, -1])

### RESULTADO db_BA
COMUNA   Ind_B1   Ind_B2   Ind_B3   Ind_B4
1      0.9947644 0.9126214 0.8235294 0.037833828
```

```

2      0.9215686 0.9310345 0.8679245 0.023152271
3      0.9111111 0.8907563 0.8305085 0.030214991
4      0.9354839 0.9021739 0.8125000 0.015343562
5      0.9836066 0.9753086 0.8684211 0.024464832
6      0.9824561 0.9583333 0.8372093 0.023726852
7      0.9672131 0.9520548 0.9066667 0.020608440
8      0.9405941 0.8913043 0.7746479 0.032953865
9      0.9631902 0.9625000 0.9047619 0.016988869
10     0.8906250 0.9210526 0.8750000 0.016891892
...     ...         ...         ...         ...

```

## • COMPONENTE ACCESO A LA INFORMACIÓN Y COMUNICACIONES

```

#-----
# Acceso a la información y comunicaciones
#-----

```

```

db_BB <- ecv_2019 %>% group_by(P_6) %>%
  dplyr::summarize(Ind_B5 = sum(P_178 == 1)/sum(table(P_178)),
                  Ind_B6 = sum(P_207 >= 1)/sum(table(P_207)),
                  Ind_B7 = sum(P_206 >= 1)/sum(table(P_206))) %>%
  ungroup()

```

```

### RESULTADO db_BB
# A tibble: 21 x 4
COMUNA  Ind_B5 Ind_B6 Ind_B7
  <fct>   <dbl> <dbl> <dbl>
1 1      0.433 0.948 0.385
2 2      0.539 0.963 0.491
3 3      0.572 0.970 0.472
4 4      0.629 0.972 0.598
5 5      0.751 0.972 0.758
6 6      0.672 0.974 0.762
7 7      0.737 0.978 0.739
8 8      0.547 0.966 0.496
9 9      0.753 0.977 0.717
10 10     0.749 0.989 0.708
# ... with 11 more rows

```

## • COMPONENTE SALUD Y BIENESTAR

```
#-----
# Salud y Bienestar
#-----

db_BC <- ecv_2019_hogares %>% group_by(P_6) %>%
  dplyr::summarize(Ind_B8 = sum(P_323 <= 2)/sum(table(P_323)),
                  Ind_B9 = sum(P_265 <= 2)/sum(table(P_265)),
                  Ind_B10 = sum(P_134 == 1)/sum(table(P_134)),
                  Ind_B11 = sum(P_266 <= 2)/sum(table(P_266)))

### RESULTADO db_BC
# A tibble: 21 x 5
  COMUNA    Ind_B8 Ind_B9 Ind_B10 Ind_B11
  <fct>    <dbl>  <dbl>  <dbl>  <dbl>
1 1      0.0404  0.311  0.147  0.247
2 2      0.0182  0.312  0.117  0.247
3 3      0.0199  0.328  0.159  0.25
4 4      0.0193  0.284  0.145  0.242
5 5      0.0273  0.325  0.189  0.236
6 6      0.0202  0.332  0.172  0.249
7 7      0.0217  0.313  0.172  0.252
8 8      0.0202  0.343  0.177  0.241
9 9      0.0197  0.337  0.202  0.222
10 10     0.0222  0.277  0.274  0.233
# ... with 11 more rows
```

## • COMPONENTE SOSTENIBILIDAD AMBIENTAL

```
#-----
# Sostenibilidad ambiental
#-----

db_BD <- ecv_2019_hogares %>% group_by(P_6) %>%
  dplyr::summarize(Ind_B12 = sum(P_312 <= 2)/sum(table(P_312)),
                  Ind_B13 = sum(P_314 <= 2)/sum(table(P_314)),
                  Ind_B14 = sum(P_317 <= 2)/sum(table(P_317)),
                  Ind_B15 = sum(P_315 <= 2)/sum(table(P_315)))

### RESULTADOS db_BD
# A tibble: 21 x 5
  COMUNA    Ind_B12 Ind_B13 Ind_B14 Ind_B15
  <fct>    <dbl>  <dbl>  <dbl>  <dbl>
```

```

1 1      0.340  0.415  0.334  0.330
2 2      0.388  0.479  0.406  0.323
3 3      0.361  0.449  0.281  0.373
4 4      0.319  0.397  0.228  0.280
5 5      0.284  0.364  0.15   0.248
6 6      0.288  0.387  0.197  0.295
7 7      0.261  0.334  0.118  0.246
8 8      0.315  0.407  0.168  0.335
9 9      0.240  0.299  0.118  0.214
10 10     0.537  0.546  0.177  0.332
# ... with 11 more rows

```

Así, juntando los 4 componentes, se construye el total de indicadores para la dimensión **FUNDAMENTOS DE BIENESTAR** de esta forma:

```

### INDICADORES DE DIMENSION FUNDAMENTOS DEL BIENESTAR
db_B <- cbind(db_BA, db_BB[, -1], db_BC[, -1], db_BD[, -1])

```

```

### RESULTADOS db_DB (Juntos todos los indicadros)

```

| COMUNA | Ind_B1    | Ind_B2    | Ind_B3    | Ind_B4     | Ind_B5    | Ind_B6    | Ind_B7    |
|--------|-----------|-----------|-----------|------------|-----------|-----------|-----------|
| 1 1    | 0.9947644 | 0.9126214 | 0.8235294 | 0.03783383 | 0.4328955 | 0.9480297 | 0.3854940 |
| 2 2    | 0.9215686 | 0.9310345 | 0.8679245 | 0.02315227 | 0.5387275 | 0.9626556 | 0.4910097 |
| 3 3    | 0.9111111 | 0.8907563 | 0.8305085 | 0.03021499 | 0.5721625 | 0.9701074 | 0.4717422 |
| 4 4    | 0.9354839 | 0.9021739 | 0.8125000 | 0.01534356 | 0.6286645 | 0.9723127 | 0.5982628 |
| 5 5    | 0.9836066 | 0.9753086 | 0.8684211 | 0.02446483 | 0.7509506 | 0.9721166 | 0.7579214 |
| 6 6    | 0.9824561 | 0.9583333 | 0.8372093 | 0.02372685 | 0.6720075 | 0.9736098 | 0.7620170 |
| 7 7    | 0.9672131 | 0.9520548 | 0.9066667 | 0.02060844 | 0.7372439 | 0.9779028 | 0.7388509 |
| 8 8    | 0.9405941 | 0.8913043 | 0.7746479 | 0.03295386 | 0.5468604 | 0.9662605 | 0.4957826 |

|   | Ind_B8     | Ind_B9    | Ind_B10   | Ind_B11   | Ind_B12   | Ind_B13   | Ind_B14   | Ind_B15   |
|---|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1 | 0.04042553 | 0.3106383 | 0.1468085 | 0.2468085 | 0.3404255 | 0.4148936 | 0.3340426 | 0.3297872 |
| 2 | 0.01822917 | 0.3125000 | 0.1171875 | 0.2473958 | 0.3880208 | 0.4791667 | 0.4062500 | 0.3229167 |
| 3 | 0.01986755 | 0.3278146 | 0.1589404 | 0.2500000 | 0.3609272 | 0.4486755 | 0.2814570 | 0.3725166 |
| 4 | 0.01934236 | 0.2843327 | 0.1450677 | 0.2417795 | 0.3191489 | 0.3965184 | 0.2282398 | 0.2804642 |
| 5 | 0.02727273 | 0.3250000 | 0.1886364 | 0.2363636 | 0.2840909 | 0.3636364 | 0.1500000 | 0.2477273 |
| 6 | 0.02020202 | 0.3316498 | 0.1717172 | 0.2491582 | 0.2878788 | 0.3872054 | 0.1969697 | 0.2946128 |
| 7 | 0.02170963 | 0.3134328 | 0.1723202 | 0.2523745 | 0.2605156 | 0.3337856 | 0.1180461 | 0.2455902 |
| 8 | 0.02020202 | 0.3434343 | 0.1767677 | 0.2407407 | 0.3148148 | 0.4074074 | 0.1683502 | 0.3350168 |

De esta misma manera se deben construir los índices para las otras dos dimensiones con el fin de aplicar el Análisis de Componentes Principales donde los sujetos corresponden a las COMUNAS y las variables a los ÍNDICES

## 5.2. Estandarización y transformación índices

La transformación y estandarización de los índices se hace con el fin de obtener indicadores en la misma escala para que puedan ser comparables. Este proceso se realiza de la siguiente manera.

- Invertir índices que tienen relación negativa con el Progreso Social, cuando se utilizan este tipo de indicadores, se invierten para incluirlos en el índice, es decir, se multiplican todos los valores del indicador por (-1).

```
#=====
# ESTANDARIZACION Y TRANSFORMACION INDICADORES
#=====

# Se adiciona la variable de COMUNA a el nombramiento de cada observacion
comunas <- db_B$P_6
db_B <- db_B[,-1]
row.names(db_B) <- comunas

#-----
# Invertir indicadores negativos
#-----
db_B$Ind_B4 <- db_B$Ind_B4*(-1)
db_B$Ind_B8 <- db_B$Ind_B8*(-1)
db_B$Ind_B9 <- db_B$Ind_B9*(-1)
db_B$Ind_B11 <- db_B$Ind_B11*(-1)
db_B$Ind_B12 <- db_B$Ind_B12*(-1)
db_B$Ind_B13 <- db_B$Ind_B13*(-1)
db_B$Ind_B14 <- db_B$Ind_B14*(-1)
db_B$Ind_B15 <- db_B$Ind_B15*(-1)
```

- Establecer utopías y distopías para cada indicador, que configuran los valores máximo y mínimo en los que puede oscilar cada uno de los indicadores. Estos valores son identificados a partir de la muestra de observaciones, es decir que la utopía es el valor máximo observado para un indicador y la distopía, el mínimo observado.

```
#-----
# Calculo de utopias (Maximo) y distopias(Minimo)
```

```
#-----
distopias <- apply(db_B, MARGIN = 2, min)
utopias <- apply(db_B, MARGIN = 2, max)

## RESULTADOS
> utopias
      Ind_B1      Ind_B2      Ind_B3      Ind_B4      Ind_B5
1.000000000 1.000000000 1.000000000 -0.005734767 0.953917051
      Ind_B6      Ind_B7      Ind_B8      Ind_B9      Ind_B10
1.000000000 0.949967084 0.000000000 -0.222816399 0.394174757
      Ind_B11     Ind_B12     Ind_B13     Ind_B14     Ind_B15
-0.167557932 -0.027397260 -0.052631579 -0.026315789 -0.026315789

> distopias
      Ind_B1      Ind_B2      Ind_B3      Ind_B4      Ind_B5      Ind_B6
0.89062500 0.83333333 0.66666667 -0.05737705 0.24852071 0.92749245
      Ind_B7      Ind_B8      Ind_B9      Ind_B10     Ind_B11     Ind_B12
0.15976331 -0.04042553 -0.44736842 0.11718750 -0.31128405 -0.53739612
      Ind_B13     Ind_B14     Ind_B15
-0.54570637 -0.40625000 -0.37251656
```

Los valores de utopías y distopías se incluyen con el conjunto de datos de las comunas cuando se aplica el PCA.

```
db_B <- (rbind(db_B, utopias = utopias, distopias = distopias))
```

- Se procede a estandarizar las variables, restando la media y dividiendo entre la desviación estándar, de tal forma que cada indicador diera lugar a una variable centrada en 0 con una varianza de 1.

```
#-----
# Estandarizacion Variables
#-----
db_B_estandar <- scale(db_B)
```

## 5.3. Análisis de Componentes Principales

### 5.3.1. Alpha de Cronbach

Antes de usar Análisis de Componentes Principales a los índices se debe comprobar que estos índices identificados inicialmente, sean los más idoneos para incluirse al IPS, para esto se usa la prueba de Alpha de Cronbach que debe tener un valor superior a 0.7 (parámetros metodología) para probar la confiabilidad de los índices, el presente análisis estadístico se hace para cada componente y se usa la librería psych de R. El Alpha de Cronbach nos da una medida de la consistencia interna que tienen los reactivos que forman una escala. Si esta medida es alta, suponemos tener evidencia de la homogeneidad de dicha escala, es decir, que los items están apuntado en la misma dirección.

$$\alpha = \frac{K}{K-1} \left( \frac{\sum_{i=1}^K \sigma_{Y_i}^2}{\sigma_X^2} \right)$$

donde,

- K = Número de items en la escala
- $\sigma_{Y_i}^2$  = Varianza del índice i
- $\sigma_X^2$  = Varianza de las puntuaciones observadas de los individuos

```
#=====
#   Confiabilidad Alpha Cronbach
#=====
alpha_BA <- psych::alpha(db_B_estandar[,1:4])
alpha_BA$total[1]

alpha_BB <- psych::alpha(db_B_estandar[,5:7])
alpha_BB$total[1]

alpha_BC <- psych::alpha(db_B_estandar[,8:11])
alpha_BC$total[1]

alpha_BD <- psych::alpha(db_B_estandar[,12:15])
alpha_BD$total[1]

## RESULTADOS
Alpha_BA Alpha_BB Alpha_BC Alpha_BD
      0.87      0.87      0.92      0.95
```

Se puede notar que las cuatro componentes tienen un Alpha de Cronbach por encima de 0.7, por lo tanto se procede a realizar el Análisis de Componentes Principales.

### 5.3.2. Componentes Principales

Para calcular los puntajes de los componentes, agregamos el conjunto de indicadores dentro de cada componente en un factor usando ACP y los 7 años de datos. ACP combina los índices anteriormente calculados de una manera que captura la cantidad máxima de varianza en los datos mientras reduce la redundancia entre índices. Básicamente, asigna a cada índice una ponderación, para garantizar contribuyan de manera significativa a la puntuación de un componente, al tiempo que tienen en cuenta las similitudes entre ellos. A las componentes que extrae la técnica se les llama con la letra  $Z_i$  donde  $i$  hace referencia a el número de la componente, para el IPS de Medellín se usa la primera componente para extraer los pesos o sea la componente  $Z_1$ .

Cada componente principal ( $Z_i$ ) se obtiene por combinación lineal de los índices calculados con la Encuesta de Calidad de Vida. Se pueden entender como nuevas variables obtenidas al combinar de una determinada forma las variables originales. **La primera componente principal** del grupo de Índices (IndB1, IndB2, ... IndB3) es la combinación lineal normalizada de dichos índices que tiene mayor varianza:

$$Z_1 = \phi_{1,1}IndB1 + \phi_{2,1}IndB2 \dots \phi_{15,1}IndB15$$

Los términos  $\phi_{1,1}, \dots, \phi_{15,1}$  reciben el nombre de *cargas* y son los que definen a la componente. Por ejemplo  $\phi_{1,1}$  es la *carga* del Indicador B1 que es la Asistencia escolar 5-11 años de la primera componente principal. Las cargas pueden interpretarse como el peso/importancia que tiene cada índice en cada componente y, por lo tanto, ayudan a conocer que tipo de información recoge cada una de las componentes.

El procedimiento del cálculo y sus resultados se describen a continuación:

```
#=====
# Analisis de Componentes principales
#=====
ips_B_escalado <- prcomp(db_B_estandar)
ips_B_escalado$rotation # Cargas

# Varianza explicada de la primer componente
varianza_explicada <- ips_B_escalado$sdev^2/sum(ips_B_escalado$sdev^2)
```



| FUNDAMENTOS DEL BIENESTAR |   |               |
|---------------------------|---|---------------|
| Componente                | Indicador   | Peso Escalado |
| Acceso a educación básica | Asistencia escolar 5-11 años                                  | 0.21          |
|                           | Asistencia escolar 12-15 años                                 | 0.27          |
|                           | Asistencia escolar 16-17 años                                 | 0.27          |
|                           | Porcentaje de analfabetismo en personas mayores de 15 años    | 0.28          |
| Acceso a la información   | Conectividad a internet                                       | 0.33          |
|                           | Uso de teléfonos celulares                                    | 0.30          |
|                           | Conectividad a televisión por cable                           | 0.29          |
| Salud y Bienestar         | Satisfacción con la calidad de vida de los miembros del hogar | 0.24          |
|                           | Satisfacción con el acceso a servicios de salud               | 0.28          |
|                           | Hábitos saludables  | 0.30          |
|                           | Satisfacción con la calidad del servicio de salud             | 0.21          |
| Sostenibilidad Ambiental  | Calidad del aire  | 0.11          |
|                           | Nivel de ruido  | 0.11          |
|                           | Arborización  | 0.23          |
|                           | Basuras en la calle   | 0.27          |

### 5.3.3. Prueba KMO

Se debe realizar una prueba de Kaiser-Meyer-Olkin (KMO) dado que este informa sobre la idoneidad de una matriz de correlaciones para aplicar el Análisis de componentes principales, es decir, permite valorar el grado en que cada una de las variables es predecible a partir de las demás. En la metodología de IPS para Medellín esta medida debe estar por encima de 0.5 para cada una de sus componentes para ser aceptada. Este es el calculo en R:

$$KMO = \frac{\sum_{i \neq j} \sum r_{ij}^2}{\sum_{i \neq j} \sum r_{ij}^2 + \sum_{i \neq j} \sum a_{ij}^2}$$

donde,  $r_{ij}$  : Correlación lineal de Pearson,  $a_{ij}$  : Correlación parcial

```
#=====
#          PRUEBA KMO
#=====
KMO_BA <- psych::KMO(db_B_estandar[,1:4])$MSA
KMO_BB <- psych::KMO(db_B_estandar[,5:7])$MSA
KMO_BC <- psych::KMO(db_B_estandar[,8:11])$MSA
KMO_BD <- psych::KMO(db_B_estandar[,12:15])$MSA

# RESULTADOS
print(c("KMO_BA" = round(KMO_BA,2), "KMO_BB" = round(KMO_BB,2),
      "KMO_BC" = round(KMO_BC,2), "KMO_BD" = round(KMO_BD,2)))

## RESULTADOS
KMO_BA KMO_BB KMO_BC KMO_BD
  0.71   0.53   0.76   0.70
```

Se puede notar que el estadístico KMO para cada uno de los componentes en la dimensión de **FUNDAMENTOS DEL BIENESTAR** es mayor a 0.5, así que se acepta

#### 5.3.4. Valor y Scores componentes

Dado que los estadísticos (Alpha y KMO) están por encima de los valores parámetro, se procede a calcular cada componente del índice como el promedio de los indicadores que la componen, ponderado por los pesos escalados del análisis de componentes principales.

$$ValorComponente_c = \sum_i (\phi_{1,i} \times indicador_i)$$

donde,

- c = Componente Social
- i = indicador del componente

```
#=====
#  Valores componentes
#=====
# Se extrae la primer componente principal
Z1 <- ips_B_escalado$rotation[,1]*(-1)

# Se calcula el valor de cada componente para cada comuna
```

```

Componente_BA <- rowSums(db_B[,1:4]*Z1[1:4])
Componente_BB <- rowSums(db_B[,5:7]*Z1[5:7])
Componente_BC <- rowSums(db_B[,8:11]*Z1[8:11])*(-1)
Componente_BD <- rowSums(db_B[,12:15]*Z1[12:15])*(-1)

```

Después de obtenidos los ponderadores, el valor de cada componente será el promedio ponderado de los indicadores que lo integran, estandarizado a un valor entre 1 y 100 teniendo como puntos de referencia la utopia como valor máximo y la distopía como valor mínimo, de la siguiente forma

$$Componente_c = 100 \times \frac{valor_j - distopia}{utopia - distopia}$$

donde *valor* : Valor componente, *j* : Comuna

```

#-----
#   IPS Componentes FUNDAMENTOS DEL BIENESTAR por comuna
#-----
IPS_BA <- (Componente_BA - 0.01)/(1-0.01)
IPS_BB <- (Componente_BB - 0.01)/(1-0.01)
IPS_BC <- ((1 - Componente_BC) - 0.78)/(0.92-0.75)
IPS_BD <- ((1 - Componente_BD) - 0.60)/(0.98-0.60)

print(list("Aceso a educación basica" = IPS_BA, "Acceso a la información" = IPS_BB,
          "Salud y Bienestar" = IPS_BC, "Sostenibilidad ambiental" = IPS_BD))

## RESULTADOS
$`Aceso a educación basica`
      1      2      3      4      5      6      7      8
0.6501231 0.6818273 0.6882260 0.7552493 0.7085358 0.7083334 0.7085269 0.7579595
      9     10     11     12     13     14     15     16
0.6966330 0.7210590 0.7630131 0.8127595 0.6774995 0.7667161 0.7172729 0.8000795
      50     60     70     80     90  utopias distopias
0.7056448 0.6942697 0.6876293 0.7469597 0.6665657 0.7746894 0.6629464

$`Acceso a la información`
      1      2      3      4      5      6      7      8
0.4550208 0.5137853 0.4896482 0.5678217 0.6365796 0.6014710 0.6177130 0.5383105
      9     10     11     12     13     14     15     16
0.6313093 0.6469654 0.7321532 0.7042771 0.5688832 0.7510298 0.6549085 0.6802700
      50     60     70     80     90  utopias distopias
0.4206029 0.5022174 0.4782581 0.5650236 0.4223314 0.7569952 0.3884255

```

\$'Salud y Bienestar'

| 1         | 2         | 3         | 4         | 5         | 6         | 7         | 8         |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0.3808651 | 0.2780530 | 0.4235199 | 0.3410711 | 0.6456410 | 0.2427067 | 0.5359620 | 0.2551108 |
| 9         | 10        | 11        | 12        | 13        | 14        | 15        | 16        |
| 0.5897227 | 0.4223392 | 0.8020937 | 0.4540503 | 0.5809205 | 0.7331379 | 0.7524418 | 0.5645779 |
| 50        | 60        | 70        | 80        | 90        | utopias   | distopias |           |
| 0.4454729 | 0.2364257 | 0.3214807 | 0.3732904 | 0.4249233 | 0.7427421 | 0.2926010 |           |

\$'Sostenibilidad ambiental'

| 1         | 2         | 3         | 4         | 5         | 6         | 7         | 8         |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0.2915520 | 0.3525979 | 0.4349426 | 0.4438785 | 0.5312574 | 0.5399139 | 0.6269562 | 0.4211030 |
| 9         | 10        | 11        | 12        | 13        | 14        | 15        | 16        |
| 0.5508827 | 0.4222603 | 0.6425873 | 0.6697786 | 0.5284839 | 0.7555209 | 0.5135347 | 0.5821402 |
| 50        | 60        | 70        | 80        | 90        | utopias   | distopias |           |
| 0.9589382 | 0.6999218 | 0.6007920 | 0.6730007 | 0.8793019 | 0.9611506 | 0.2082342 |           |

### 5.3.5. IPS Dimensión FUNDAMENTOS DEL BIENESTAR

El valor de la dimensión se calcula como el promedio aritmético de los valores de las componentes, y el valor del índice, como el promedio aritmético de las dimensiones, en este caso calcularemos el IPS de la dimensión FUNDAMENTOS DE BIENESTAR por comuna con el promedio de las cuatro componentes que calculamos en la sección anterior.

$$Dimensión_d = \frac{1}{4} \sum_c Componente_c$$

, donde,  $d$  : Dimensión y  $c$  : Componente

```
#-----
```

```
# IPS dimensión FUNDAMENTOS DEL BIENESTAR por comuna
```

```
#-----
```

```
BIENESTAR <- cbind(IPS_BA, IPS_BB, IPS_BC, IPS_BD)
```

```
IPS_BIENESTAR <- apply(BIENESTAR, MARGIN = 1, mean)[1:21]
```

```
## RESULTADOS
```

| 1         | 2         | 3         | 4         | 5         | 6         | 7         | 8         |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0.4443902 | 0.4565659 | 0.5090842 | 0.5270051 | 0.6305035 | 0.5231063 | 0.6222895 | 0.4931210 |
| 9         | 10        | 11        | 12        | 13        | 14        | 15        | 16        |
| 0.6171369 | 0.5531560 | 0.7349618 | 0.6602164 | 0.5889468 | 0.7516012 | 0.6595395 | 0.6567669 |
| 50        | 60        | 70        | 80        | 90        |           |           |           |

0.6326647 0.5332087 0.5220400 0.5895686 0.5982806

## 6. Análisis descriptivo Resultado

En la figura 4, se muestra un breve resumen de los resultados del Indicador de Progreso Social para la dimensión **FUNDAMENTOS DEL BIENESTAR** calculado en este informe, si se comparan con los resultados obtenidos en el 2017 es muy similar, recordar que en este informe no se calculo el IPS general por las comunas sino solo el de la dimensión de Fundamentos del Bienestar.

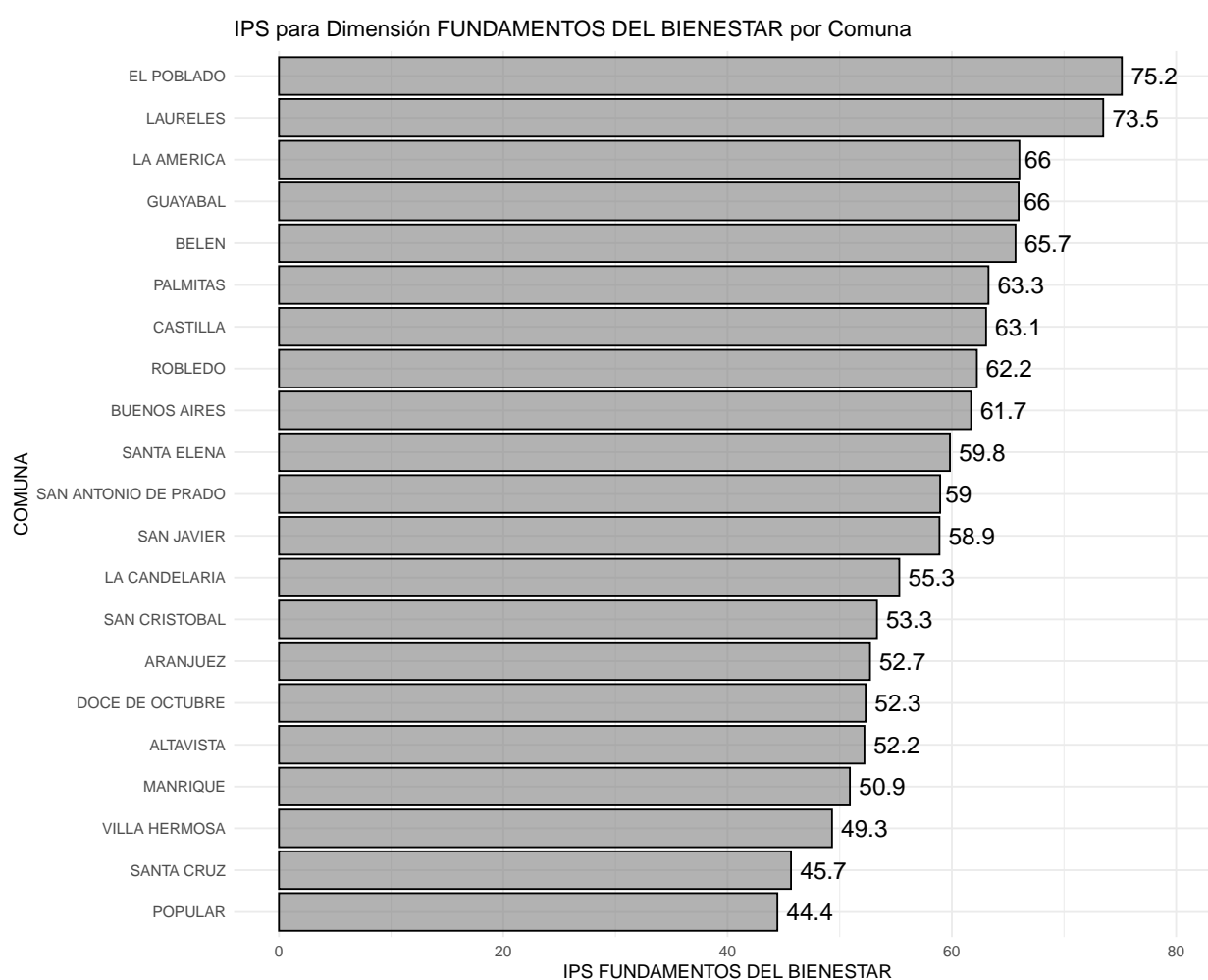


Figura 4: Resultados IPS para Fundamentos del Bienestar

## 7. Conclusiones y sugerencias

- La elección de Análisis de Componentes Principales como base de la agregación a nivel de componente está influenciada por la calidad y cantidad de datos disponibles sobre el progreso social.
- Para que el ACP sea válido, cada indicador debe estar relativamente libre de errores de medición
- Si los estadísticos Alpha de Cronbach y KMO están por debajo de los valores parámetro 0.7 y 0.5 respectivamente, se deben retirar los indicadores hasta tener especificaciones técnicas
- El conjunto de datos para usar la metodología ACP debe contener las comunas en las filas y los índices en las columnas. La recomendación es organizar como se mostró en este informe para la dimensión FUNDAMENTNOS DEL BIENESTAR las otras dos dimensiones e ir actualizando constantemente.
- En el anexo Metodología de cálculo del índice del informe del índice de Progreso Social Comunas y Corregimientos de Medellín 2013-2017 publicado en Abril de 2019, no se especifica muy claro algunos pasos que se necesitan para construir el índice, la recomendación es desplegar un poco más cada paso.
- En la bibliografía no se encontró como extraer datos de otras fuentes como la subdirección de planeación y la secretaría de salud, la recomendación es anexar a los informes el sitio de extracción y el ejemplo para calcular índices que son usados posteriormente en la metodología.
- El propósito de este informe no es replicar la metodología y obtener resultados iguales, el propósito es detallar el método y dar un panorama a la utilización de fuentes de datos como la Encuesta de Calidad de Vida para llegar a valores que caractericen los barrios con la metodología encontrada en las bibliografías.

## Referencias

- [1] WARES, A; HELLMAN, T. (2020). Social Progress Index 2020 - Methodological report. Washington
- [2] MEDELLÍN COMO VAMOS (2019). índice de Progreso Social Comunas y corregimientos de Medellín 2013-2017 Medellín
- [3] PORTER, M; STERN, S, GREEN, M (2017). Social Progress Index 2016 Washington
- [4] JAMES, G; WITTEN, D; HASTIE, T; TIBSHIRANI, R (2019). An Introduction to Statistical Learning with Applications in R Springer
- [5] LÓPEZ, J (2012). Análisis de Datos con R en Psicología y Educación
- [6] AIZEN, I; FISHBEIN, M (1980). Understanding attitudes and predicting social behavior. Englewood, NJ: Prentice-Hall
- [7] GRISALES, H; ARBELÁEZ M (2008). Metodología para el diseño de un índice de condiciones de vida para los adolescentes jóvenes. Medellín