

# Índices de Desarrollo: Correlación con la Población Colombiana

Juan Sebastián Cortázar<sup>1</sup>, María Alejandra Restrepo<sup>2</sup>, and Juan Camilo Rueda<sup>3</sup>

<sup>1,2,3</sup>Universidad de los Andes,  
js.cortazar533,ma.restrepot,jc.rueda169@uniandes.edu.col

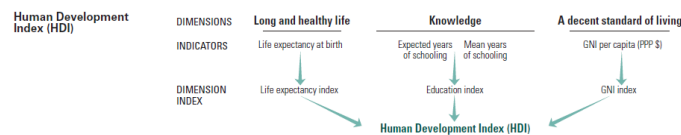
7 de Julio de 2018

## Abstract

En este trabajo se muestra el **Índice de Desarrollo Humano** (IDH) de Colombia por departamento, donde a través de estadística poblacional se observa el nivel de educación, nivel de salud e ingresos per cápita de los distintos departamentos para poder determinar los departamentos más vulnerables y los que tienen un mejor índice. Adicional, se identifica una correlación entre la población y el IDH, por lo cual se hace una regresión entre el IDH y las poblaciones de cabecera por departamento y el total de población de cada departamento con el IDH.

## Introducción

El Índice de Desarrollo Humano es una medida utilizada para determinar el crecimiento y el desarrollo de las zonas y los países teniendo no solo en cuenta el crecimiento económico. Este índice busca medir no solo los PIB per cápita de las personas, sino que busca entender el acceso a salud y a la educación que determinan la posibilidad de crecimiento de las sociedades y sus capacidades de generar unas mejores condiciones de vida. El valor del IDH es la media geométrica entre los índices normalizados de las tres dimensiones (Salud, Educación y Nivel de Vida) como se muestra en la imagen a continuación.



Para el cálculo de cada uno de los índices se tienen unos límites inferiores y superiores que en conjunto con los valores del sector (ya sea país o departamento) generan cada uno de los índices. En la tabla 1 se puede ver lo siguiente.

Dimensiones	Indicador	Min	Max
Salud	Expectancia de Vida (an os)	20	85
Educacio´n	Escolaridad Adultos	0	18
	Esperanza educativa nin os	0	15
Nivel de Vida	PIB per Capita (USD ctes 2011)	100	75,000

Table 1: Rango de Dimensiones IDH

La variable Salud se genera a trave´s de un i´ndice compuesto que refleja condiciones de salud en los hogares: proteccio´n de salud, a trave´s del IGSS o de un seguro, nu´mero de personas por dormitorio, tipo de acceso a agua y saneamiento y tipo de piso en la vivienda. Todos estos factores influyen la expectantica de vida y se calculan de la siguiente manera.

$$Salud = \frac{LE - 20}{85 - 20}$$

Donde  $LE = ExpectativadeVida$  La variable Educacio´n es un indicador compuesto que incluye la escolaridad alcanzada por adultos mayores de 25 an os y la esperanza educativa en nin os. En el primer indicador se mide la tasa de alfabetizacio´n de adultos en el segundo se mide la tasa bruta combinada de matriculacio´n en educacio´n primara, secundario y superior, asi´ como los an os de duracio´n de la educacio´n obligatoria. El ca´lculo del i´ndice de educacio´n se define de la siguiente manera

$$Educacion = \frac{EA + EN}{2}$$

Donde

$$EA = \frac{Promdean osdeeducacionadultos}{18}$$

$$EN = \frac{Promdean osdeeducacionnin os}{15}$$

La variable del nivel de vida mide el PIB per ca´pita de una zona o pai´s teniendo en cuenta un mi´nimo esperado y un maximo. La formula es la siguiente

$$NiveldeVida = \frac{Ln(PIBx) - Ln(100)}{Ln(75,000) - Ln(100)}$$

Comencemos viendo que hay en la seccio´n 1 en la pa´gina 3.

# 1 Exploración Univariada

En esta sección nos interesa explorar cada índice (IDH), para esto se realizan varias estadísticas con la información obtenida. En primer lugar, se evalúa el número de datos y la mediana de cada uno de los tipos de población. En la tabla 2 en la página 3.

Table 2: Medidas estadísticas

Statistic	N	Median
cabecera	32	717,197
resto	32	268,111.5

Para resaltar lo anterior, tenemos la Figura 1 en la página 3.

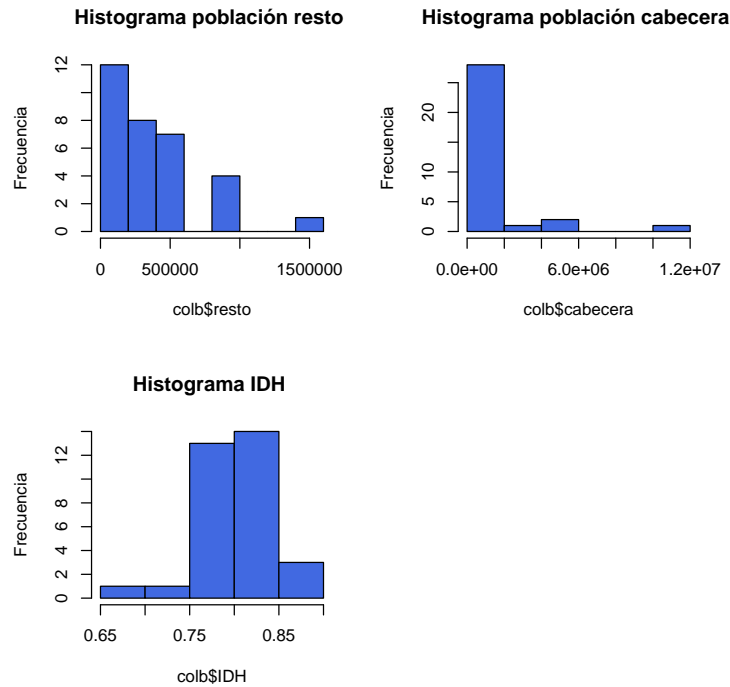


Figure 1: Histograma del IDH

Como las poblaciones tienen un sesgo se normalizan los datos con log, el histograma de estos nuevos datos se muestra en la Figura 2 en la página 4.

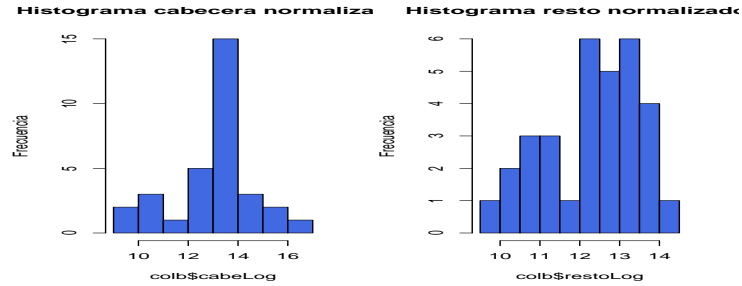


Figure 2: Histograma de poblaciones

## 2 Exploración Bivariada

en esta sección nos interesa ver el impacto que tiene la población en el IDH, para esto se presenta en la tabla refcorrDem en la página 4. la correlación de las variables normalizadas con respecto al IDH  $i+$

Table 3: Correlación de Democracia con las demás variables

total	cabeLog
0.399	0.487

Ademas, se muestra la correlacion entre todas las variables independientes en la tabla 4 en la página 4

Table 4: Correlación entre variables independientes

	total	cabeLog
total	1	
cabeLog	0.71	1

los datos anteriores los puede ver visualmente en la figura 3 en la página 6

## 3 Modelos de Regresión

Finalmente, vemos los modelos propuestos. En cada una se evalua la variable independiente DIH con cada una de las categorias de la poblacion. Los resultados se muestran en la Tabla 5 de la página 5.

Table 5: Modelos de Regresión

	<i>Dependent variable:</i>		
	IDH		
	(1)	(2)	(3)
cabeLog	0.013*** (0.004)		
restoLog		0.007 (0.007)	
totalLog			0.013** (0.005)
Constant	0.634*** (0.055)	0.722*** (0.082)	0.629*** (0.068)
Observations	32	32	32
R <sup>2</sup>	0.238	0.031	0.179
Adjusted R <sup>2</sup>	0.212	−0.001	0.152
Residual Std. Error (df = 30)	0.037	0.042	0.039
F Statistic (df = 1; 30)	9.347***	0.974	6.561**
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01		

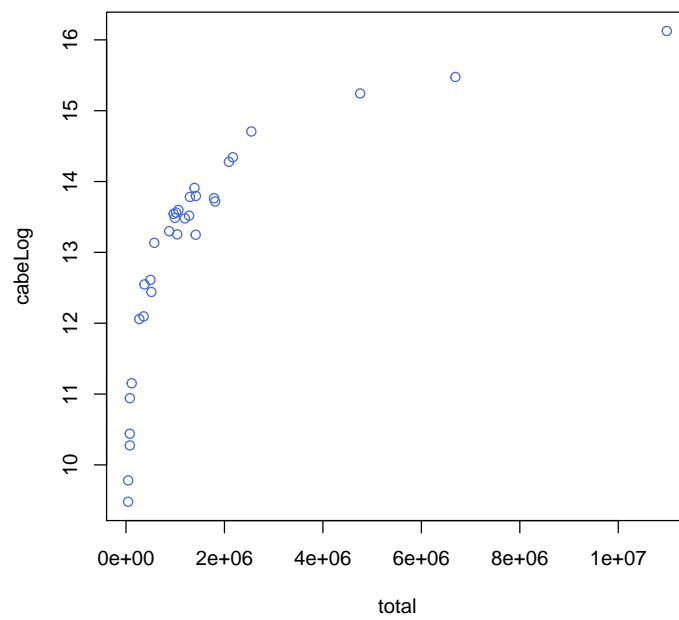


Figure 3: correlacion entre cabelog y restolog

## 4 Exploración Espacial

El siguiente mapa muestra el impacto que tiene la población de cada departamento sobre su respectivo IDH:

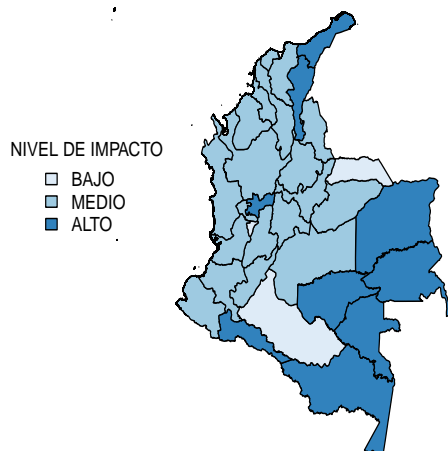


Figure 4: Impacto de población en IDH por departamento

Entre mayor es el impacto de la población sobre el IDH del departamento, más oscuro se muestra en el mapa. Se dividió el impacto en 3 niveles: BAJO, MEDIO Y ALTO. Para lograr esta escala se implementó el siguiente procedimiento:

Primero se obtuvieron los datos de IDH, población de cabecera, el resto de la población y el total de la población de cada departamento del país.

Se limpiaron los datos reemplazando caracteres no reconocibles tales como la letra "ñ" y tildes. Para evitar un sesgo significativo por variaciones amplias en número de habitantes, primero se utilizaron valores logarítmicos y luego se normalizaron. Finalmente se crearon las 3 agrupaciones por medio de la técnica K-means.