

# INDICES POBLACIONALES EN COLOMBIA POR DEPARTAMENTO

Lorenzo Vicini<sup>1</sup>

<sup>1,2</sup>Escuela de Ingenieria, Universidad de los Andes, Herramientas Computacionales  
le.vicini10@uniandes.edu.co

02 de Julio de 2018

## Abstract

Este es mi primer trabajo en exploracion y modelamiento de indices usando LATEX. Este trabajo lo he hecho bajo la filosofia de trabajo replicable a partir de el archivo provisto por el profesor del curso, el cual hacia uso de programas tales como: RStudio, Python, Zotero, entre otros.[1]

## Introduccion

El siguiente trabajo pretende mostrar El indice de desarrollo humano (IDH) para Colombia a partir de la densidad poblacional de cada departamento, la cual se divide en rural y urbana. Cabe recordar que el IDH es un indicador del desarrollo humano por pais, elaborado por el Programa de las Naciones Unidas para el Desarrollo , el cual tiene en cuenta: el tener una vida larga y saludable, adquirir conocimientos, esperanza de vida, periodos de educacion recibidos, indice del PIB, etc. En el desarrollo del trabajo se realizaran analisis estadisticos fundamentados con graficas y tablas. Se finalizara con un analisis de Clusterizacion el cual mostrara de manera grafica el desarrollo de los diferentes departamentos de Colombia.

# 1 Exploracion Univariada

Se denomina exploracion univariada al proceso en el que se obtiene las medidas estadisticas, la tabla de frecuencias y algun grafico de resumen de una variable en particular. Dichos datos y analisis se mostraran a continuacion. En esta seccion se exploran los diferentes datos (IDH, Poblacion Rural y Urbana)

Comencemos viendo que hay en la sección 1 en la página 2.

Table 1: Medidas estadisticas

Statistic	N	Median
IDH	32	0.804
Población.Cabecera	32	717,197
Población.Resto	32	268,111.5
Población.Total	32	1,028,429

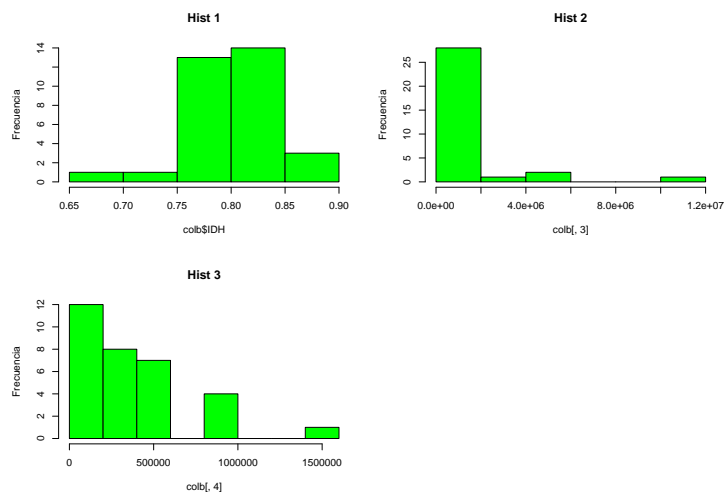


Figure 1: Histogramas Iniciales

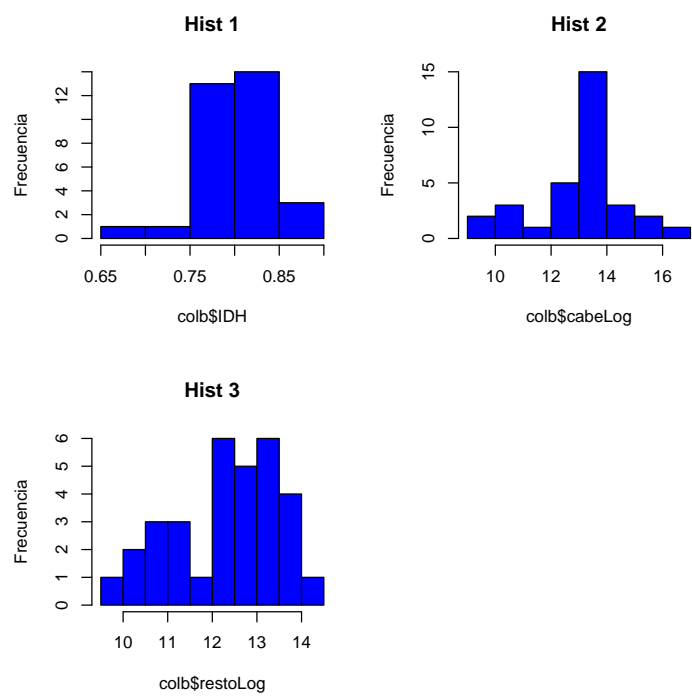


Figure 2: Histogramas despues de la transformacion logaritmica

## 2 Exploracion Bivariada

Las correlaciones bivariadas son herramientas comunes y se utilizan para estudiar como una variable influye con otra. En esta seccion se genera un analisis de la variable IDH con respecto a las otras dos variables (poblacion urabana y rural)

Table 2: Correlacion de IDH con las demas variables

cabeLog	restoLog
0.487	0.177

En probabilidad y estadistica, la correlacion indica la fuerza y la direccion de una relacion lineal y proporcionalidad entre dos variables estadisticas. Se considera que dos variables cuantitativas estan correlacionadas cuando los valores de una de ellas varian sistematicamente con respecto a los valores de la otra.

Estos son los resultados de la correlacion:

Table 3: Correlación entre variables independientes

	cabeLog	restoLog
cabeLog	1	
restoLog	0.84	1

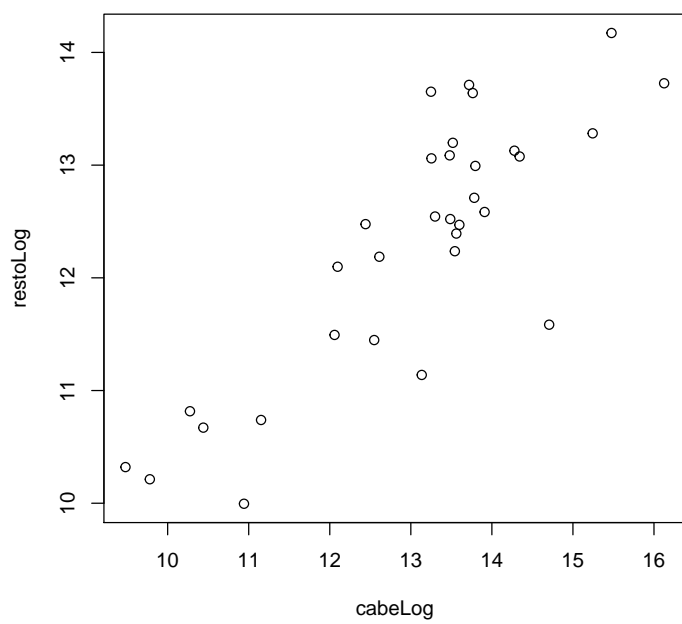


Figure 3: Grafico de dispersion

### 3 Modelos de Regresion

En estadística, el análisis de la regresión es un proceso estadístico para estimar las relaciones entre variables. Incluye muchas técnicas para el modelado y análisis de diversas variables, cuando la atención se centra en la relación entre una variable dependiente y una o más variables independientes. Para nuestro caso, la variable dependiente será el IDH.

Table 4: Modelos de Regresión

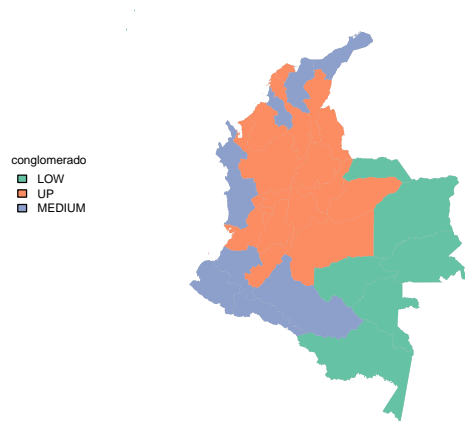
	<i>Dependent variable:</i>	
	IDH	
	(1)	(2)
cabeLog	0.013*** (0.004)	0.031*** (0.007)
restoLog		-0.030*** (0.010)
Constant	0.634*** (0.055)	0.766*** (0.065)
Observations	32	32
R <sup>2</sup>	0.238	0.425
Adjusted R <sup>2</sup>	0.212	0.385
Residual Std. Error	0.037 (df = 30)	0.033 (df = 29)
F Statistic	9.347*** (df = 1; 30)	10.706*** (df = 2; 29)

*Note:*

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## 4 Exploracion Espacial

Para realizar el siguiente mapa, se llevo a cabo la utilizacion de la tecnica de "Kmeans" la cual es un metodo de agrupamiento, que tiene como objetivo la particion de un conjunto de  $n$  observaciones en  $k$  grupos en el que cada observacion pertenece al grupo cuyo valor medio es mas cercano. Lo cual realizaremos con el mapa de Colombia para mostrar de manera grafica el desarrollo de cada departamento clasificado en tres categorias: Alto, medio y bajo.



## Bibliografia

- [1] J. MACQUEEN. SOME METHODS FOR CLASSIFICATION AND ANALYSIS OF MULTIVARIATE OBSERVATIONS. *MULTIVARIATE OBSERVATIONS*, page 17.