



# Análisis de componentes principales aplicado a la prueba estatal Colombiana Saber 11

## Principal component analysis applied to the state Colombian test ICFES Saber 11

RUIZ ESCORCIA, Rafael Roberto [1](#); ARÉVALO MEDRANO, Jesús Bautista [2](#); MORILLO, Greisy Paola [3](#); ACOSTA-HUMÁNEZ, Primitivo Belén [4](#)

Recibido: 17/10/2017 • Aprobado: 10/11/2017

### Contenido

- [1. Introducción](#)
  - [2. Marco conceptual](#)
  - [3. Metodología](#)
  - [4. Resultados](#)
  - [5. Conclusiones](#)
- [Referencias bibliográficas](#)

#### RESUMEN:

El propósito de este artículo es evidenciar las variables que influyen en el cumplimiento de los estándares curriculares emitidos por el Ministerio de Educación Nacional (MEN) de Colombia para las instituciones educativas. Para tal efecto se utilizaron los resultados obtenidos por varias instituciones educativas en el distrito de Barranquilla, de la prueba estatal denominado ICFES Saber 11, realizada en el año 2012. Para el análisis estadístico se aplicó la técnica del análisis multivariado conocida como Análisis de Componentes Principales (ACP).

**Palabras clave:** Análisis de Componentes Principales, ICFES Saber 11, Estándares Curriculares, MEN

#### ABSTRACT:

The aim of this paper is to show which variables can influence in the achievement of high level criterion determined by Colombian National Education Ministry (MEN) to educational institutions in the country. To do this, it have been used the results obtained by some educational institutions located in the district of Barranquilla, corresponding to the national test known as ICFES Saber 11 during 2012. The statistical analysis was made through the multivariate analysis technique known as Principal Components Analysis.

**Keywords:** Principal Components Analysis, ICFES Saber 11, High level criterion, MEN

## 1. Introducción

De acuerdo con los requerimientos del Ministerio de Educación Nacional (MEN) de Colombia, establecidos en el artículo 3 del decreto 869 de 2010 (MEN [2010]), los estudiantes que en

Colombia pretendan continuar sus estudios hacia la educación superior colombiana tienen como requisito obligatorio la realización de las pruebas denominadas *Pruebas Saber 11*.

Uno de los propósitos de la prueba Saber 11 es medir la calidad de la educación en Colombia. En la actualidad el *Instituto Colombiano de Fomento a la Educación Superior (ICFES)* también realiza las pruebas Saber 3°, 5° y 9°, las cuales sirven como termómetro de la educación en Colombia previamente a la realización de las Pruebas Saber 11. De igual manera, el ICFES facilita a los estudiantes, por medio virtual, las pruebas realizadas anteriormente, Barón et al [2013], MEN [2010].

La Institución educativa forma un parte importante en este proceso dado que ésta aporta tanto los medios como la motivación, útiles en sentido pedagógico, para facilitar el proceso del cumplimiento de las metas establecidas por el Ministerio de Educación Nacional (MEN), las cuales consisten en el logro de mejores resultados en cada prueba (estándares curriculares), MEN [2007].

La evaluación es un factor esencial para incrementar la calidad, la cual es un aspecto relevante del proceso educativo debido a que provee variadas clases de información que permiten tomar decisiones con base a estar mejor informados y también entender procesos de enseñanza-aprendizaje que no se visualizan sin su aplicación, MEN [2008], González Zábala [2013].

Hacer un uso pedagógico de los resultados de la evaluación permite orientar el trabajo de las instituciones, los docentes, los estudiantes y los padres de familia para poder implementar las mejoras necesarias. Es allí donde radica la importancia de ver la evaluación como una herramienta potenciadora de los aprendizajes y los procesos que ocurren en el aula, dentro del ciclo de calidad que pretenden fortalecer las instituciones educativas conjugando estándares básicos de competencia, procesos de evaluación y diseño e implementación de planes de mejoramiento institucional, MEN [2008].

Los resultados de las evaluaciones del ICFES son un indicador indispensable, pero no único, del aprendizaje de las personas y de la calidad de las instituciones o de la educación en general, ICFES [2011], López [2011].

En aras de retroalimentarse con respecto a la calidad de la educación en todos sus niveles y adelantar investigación sobre los factores que inciden en la calidad educativa, el MEN ha establecido evaluaciones censales desde 1968 al implementar el denominado Examen de Estado de la Educación Media. Dicho examen a través del tiempo ha tenido varias modificaciones hasta la creación de la prueba SABER en 1991, que posteriormente evolucionó en 2010 al implementarse como la prueba ICFES SABER 11, véase ICFES [1999], ICFES [2010], ICFES [2012] y MEN [2010].

Estos aspectos también han sido considerados en otros países tales como España a través de las pruebas denominadas Selectividad, Martí Recober et. Al. [1997], EURYDICE [2008], lo cual indica que no es solamente una problemática colombiana.

El objetivo de este artículo es evidenciar el cumplimiento de dichos estándares tomando como fuente para la investigación una base de datos del ICFES con los resultados de la prueba Saber 11 realizada en el año 2012 por los estudiantes del distrito Barranquilla. Tales resultados se analizaron estadísticamente mediante la técnica de Análisis Multivariado denominada Análisis de Componentes Principales (León et. Al. [2008] Peña [2002], Padilla et. al. 2012), con el propósito de asociar las variables que influyen en el rendimiento de los estudiantes, ante la prueba ICFES SABER 11, de algunas instituciones del distrito de Barranquilla realizada en el año 2012.

---

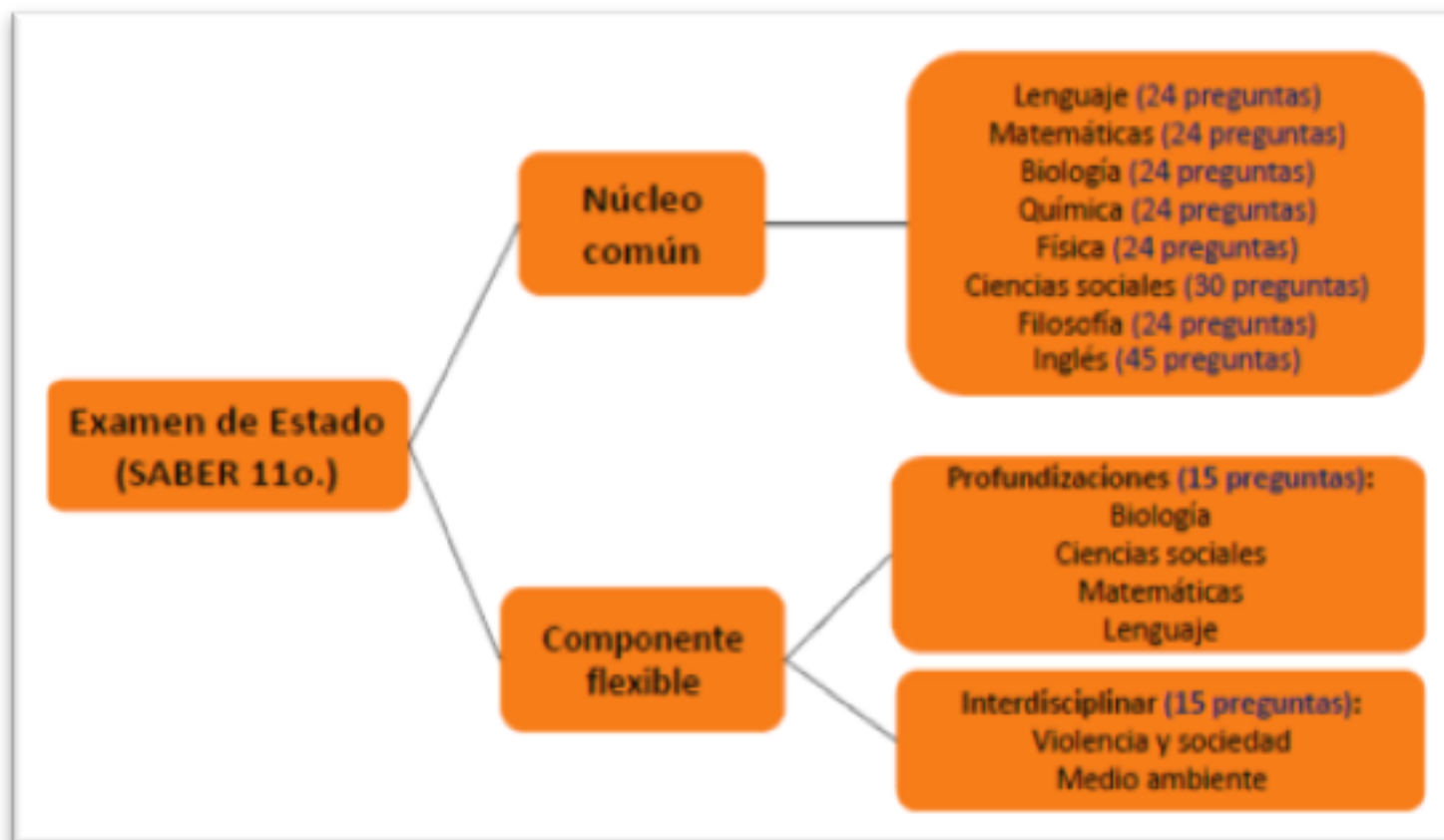
## 2. Marco conceptual

En esta sección se establecen los aspectos teóricos necesarios para comprender el artículo y por ende los resultados.

# Pruebas Saber 11

La evaluación ICFES SABER 11 presenta la siguiente estructura, Figura 1., 249 preguntas distribuidas de la siguiente forma:

**Figura 1**  
Examen de validación del bachillerato académico; Qué se evalúa



Fuente: ICFES SABER 11° (2012)

**Competencias:** son los procesos cognitivos que el estudiante debe realizar para resolver una pregunta.

**Componentes:** son las categorías conceptuales o los tópicos propios del área o la disciplina.

Todas las pruebas del núcleo común y las de profundización, presentadas en la Figura 1., evalúan competencias y componentes del área.

A continuación se muestra los aspectos que el examen saber 11 evalúa en cada prueba de núcleo común, por componente y competencia, ver Tabla 1.

**Tabla 1**  
Componentes

Núcleo común	Componentes	Competencias
Lenguaje	1-funcion semántica de la información local	1- Acción interpretativa
	2-Configuración del sentido global	2- Acción argumentativa
	3-sentido del texto en relación con otros textos.	3- Acción propositiva
Matemáticas	1-Numérico – Variacional	1- Comunicación
	2-Geométrico – métrico	2- Razonamiento
	3-Aleatorio	3- Solución de problemas
Química	1- Aspectos analíticos de sustancias.	1- Uso comprensivo del conocimiento
	2- Aspectos fisicoquímico de sustancias	
	3- Aspectos analíticos de mezclas	
	4- Aspectos fisicoquímicos de mezclas	
Física	1- Mecánica clásica	2- Explicación de fenómenos
	2- termodinámica	
	3- Eventos ondulatorios	
	4- Eventos Electromagnéticos	
Biología	1- Celular	3- Indagación
	2- Organísmico	
	3- Ecosistémico	
Ciencias sociales	1- El espacio, el territorio, el ambiente y la población.	1- Interpretativa 2- Argumentativa 3- Propositiva
	2- El poder, la economía y las organizaciones sociales.	
	3- El tiempo y las culturas.	
Filosofía	1- La pregunta del hombre frente a su mundo social y cultural.	3- Propositiva
	2- La pregunta por el ser	
	3- La pregunta por el conocimiento	

Fuente: Elaboración propia

El núcleo común de inglés presenta otra estructura en su prueba; en el año 2006 el MEN formuló los estándares básicos de competencia en lengua extranjera – inglés. Estos estándares están alineados con el Marco Común Europeo, razón por lo cual esta prueba consta de 45 preguntas, divididas en siete partes y sus resultados hacen referencia a las bandas del Marco Común Europeo:

- A1 y A2: usuario básico
- B1 y B2: usuario independiente

### Partes de la prueba de inglés

- Ante cinco avisos, el estudiante debe decidir en dónde los vería.
- Relación entre palabras y descripciones que dan cuenta de su definición.
- Completar cinco conversaciones cortas.
- Completar un texto (8 palabras).
- Comprensión de un texto (7 preguntas).
- Comprensión de un texto (5 preguntas).
- Completar un texto (10 palabras).

### Análisis de componentes principales (ACP)

Es una técnica estadística del Análisis Multivariado que nos permite explicar la mayor variabilidad de los datos con un menor número de variables.

## 3. Metodología

Esta investigación es de carácter descriptiva cuasi-experimental, lo cual permite la identificación de variables clave que pueden ser determinantes en el proceso, y mejorar la implementación de el examen Saber 11 en Colombia, para este proceso se realizó lo siguiente.

La base de datos utilizada para el desarrollo de esta investigación fue suministrada por el Instituto Colombiano para la Evaluación de la Educación (ICFES) mediante el portal virtual

www.icfes.gov.co. Se tomó inicialmente la base de datos que contenía 12274 Instituciones a nivel nacional y luego de depurarla se tomaron solo 385 Instituciones que correspondían al distrito de Barranquilla.

Además de lo anterior, se realizó el conteo de las instituciones de acuerdo con su naturaleza (oficial o no oficial), de manera que 220 Instituciones eran de naturaleza oficial y 165 Instituciones eran de naturaleza no oficial.

Con la base de datos depurada se establecieron las siguientes variables:

NOM = Nombre del municipio al cual pertenece el colegio

CAL = Calendario académico regido en el colegio

JOR = Jornada académica a la cual pertenece los estudiantes que representa a las Instituciones en el ICFES

NAT = Naturaleza del colegio

EVAL = Número de estudiantes que se presentaron por institución

NUCLEO COMUN: Explicado en la Tabla 2, que aparece a continuación:

**Tabla 2**  
Núcleo común

<i>Variable de desempeño academico</i>	<i>Codigo</i>
Promedio del porcentaje acumulado de los estudiantes, en el area de Matematicas	MAT
Promedio del porcentaje acumulado de los estudiantes, en el area de Quimica	QUI
Promedio del porcentaje acumulado de los estudiantes, en el area de Fisica	FISICA
Promedio del porcentaje acumulado de los estudiantes, en el area de Biologia	BIO
Promedio del porcentaje acumulado de los estudiantes, en el area de filosofia	FILO
Promedio del porcentaje acumulado de los estudiantes, en el area de ingles	INGLES
Promedio del porcentaje acumulado de los estudiantes, en el area de lenguaje	LENG
Promedio del porcentaje acumulado de los estudiantes, en el area de sociales	SOC

Fuente: Elaboración propia

Para el análisis de los datos se realizó un previo bosquejo de las medidas básicas estadísticas tales como media, desviación y boxplot, teniendo en cuenta a las variables NUCLEO COMUN, JOR, CAL, NAT, EVAL como variables activas, mientras que la variable NOM (Nombre del municipio) es ilustrativa o suplementaria. Se buscaron diferencias significativas en las variables de estudio para establecer las primeras conclusiones necesarias para sugerir mejoras en el proceso de enseñanza aprendizaje en las instituciones analizadas.

Posterior al análisis descriptivo, se desarrolló el análisis estadístico mediante la técnica de Componentes Principales (ACP) tomando las variables centradas-reducidas, es decir, se partió de la matriz de correlaciones o un ACP normado. Observando las tablas de salida del ACP (matriz de correlación, matriz de grafica de dispersión, grafica de sedimentación, grafica del círculo unitario), nos interesamos en las matrices de correlación y de dispersión buscando si las variables presentaban alguna interacción significativa.

La gráfica de sedimentación fue útil para la selección de la cantidad de componentes que explican la mayor variabilidad de los datos, mientras que la gráfica del círculo unitario se utilizó para observar que tan representadas se encuentran las variables.

Por último, el diagrama de dispersión de los datos en el primer plano factorial ayudó a evidenciar su comportamiento con respecto a las componentes establecidas.

Todos los análisis antes mencionados se efectuaron con el software libre R Foundation for Statistical Computing, versión 2.14.2 (2012-02-29) y se requirieron tablas del programa de Windows office Excel versión 2007.

## 4. Resultados

Por medio de la aplicación de la metodología desarrollada (ACP) se alcanzaron los siguientes resultados:

Se realizó un análisis descriptivo de las variables activas cuantitativas, el cual se resume en la Tabla 3 y en la Figura 2.

**Tabla 3**  
Análisis de los datos de variables cuantitativas

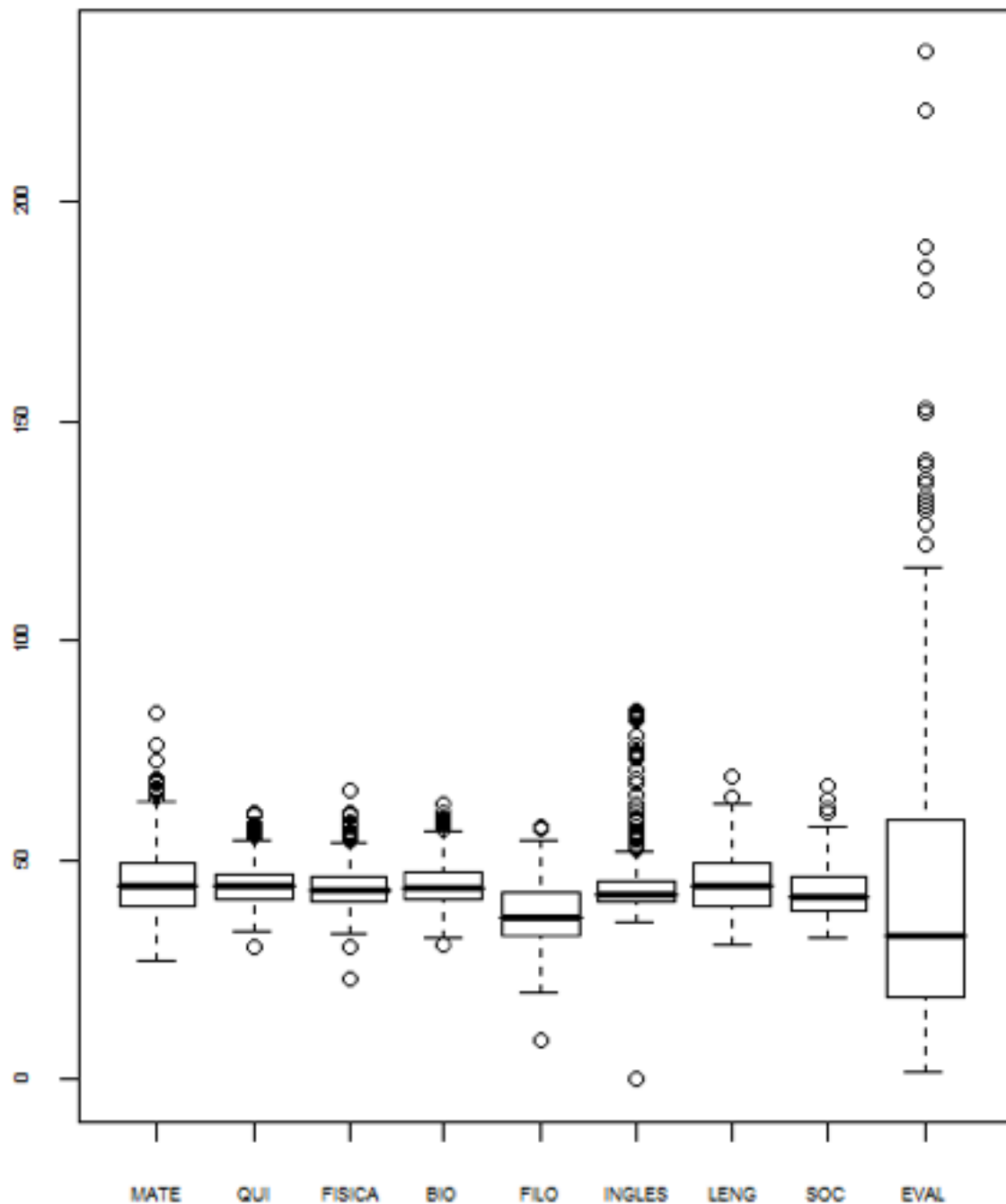
	<i>Promedio</i>	<i>Desviación Estándar</i>	<i>Coeficiente de Variación</i>
MAT	45,0397	8,12943	0,180495
QUI	44,247	4,4781	0,101207
FISICA	43,8919	4,85058	0,110512
BIO	44,225	5,13645	0,116143
FILO	37,8871	6,86038	0,181075
INGLES	44,4046	7,97755	0,179656
LENG	44,8358	6,67174	0,148804
SOC	42,7595	6,02325	0,140864
EVAL	44,3792	36,3544	0,819176

Fuente: SPSS

----

**Figura 2**  
Rendimiento - Boxplot





Fuente: SPSS

Se observa que la mayor variabilidad la presentan la variable EVAL (Numero de evaluados) con un coeficiente de variabilidad del 81.9% con respecto a las demás variables, debido a la variedad de egresados de las diversas entidades educativas del distrito de barranquilla, siendo que el 66.06% de las entidades presentaron menos de 50 estudiantes a la prueba y solo 1.81% presentaron más de 150 estudiantes.

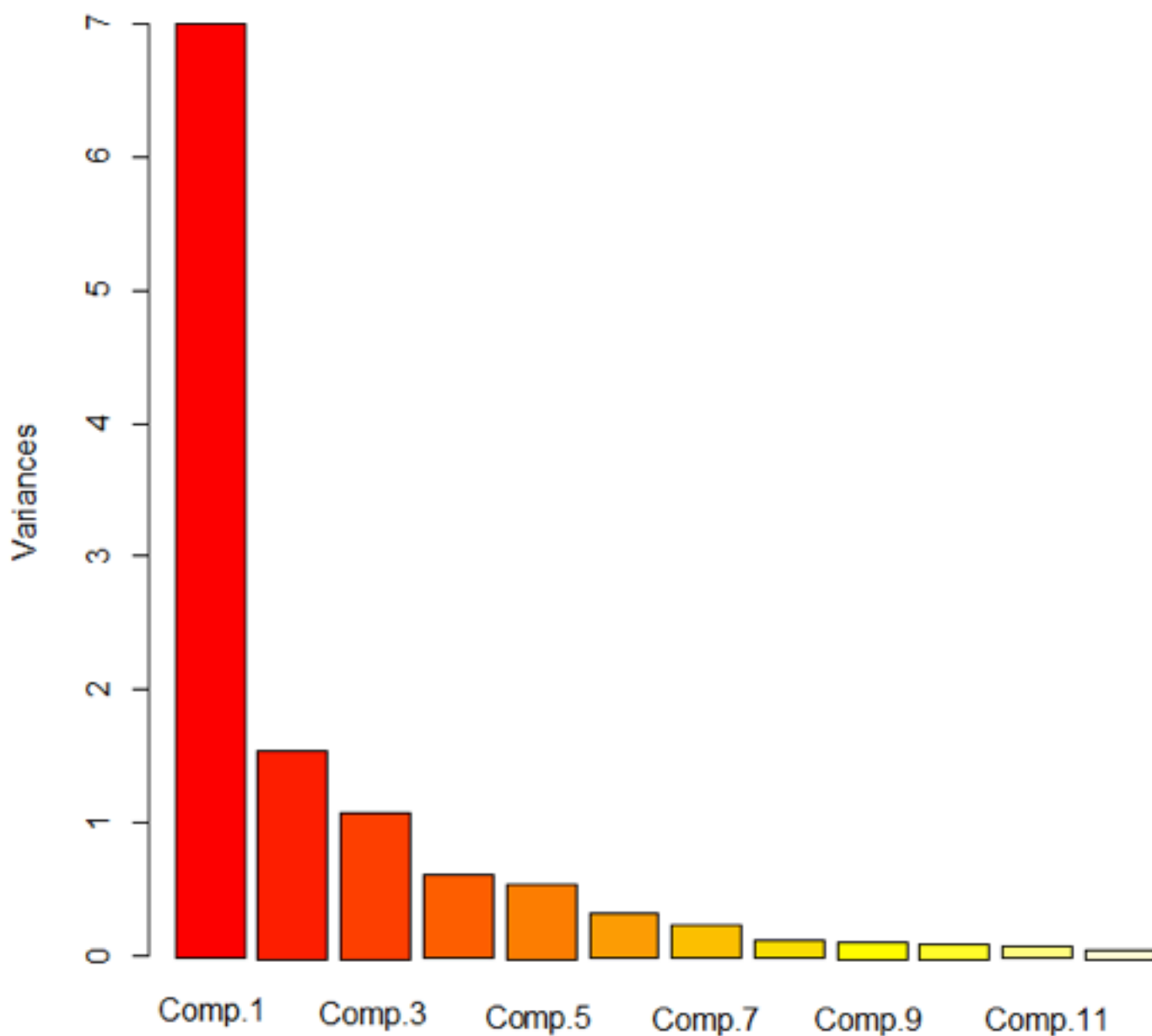
Se puede apreciar que las variables QUI (núcleo común química) y BIO (núcleo común biología) presentan una variabilidad semejante, según la Figura 2., correspondiente al boxplot, también se destaca que los desempeños promedios más altos se ubican en las variables MAT (núcleo común matemáticas) y LENG (núcleo común lenguaje), lo cual es consecuencia de las políticas gubernamentales en el fortalecimiento de estas asignaturas. Además, la variable MATE (núcleo común Matemáticas), después de la variable EVAL, es la que presenta una variación ligeramente superior a las demás del 18.04%.

La variable INGLES (núcleo común Ingles) presenta un mayor número de datos atípicos que se alejan de la media que es 44.40. Se observa que una institución no presento esta prueba debido a que se trata de una institución de rehabilitación auditiva y lenguaje oral, determinando un valor incierto en este punto, designándole una valoración cero por concepto de cálculos.

Al realizar este análisis descriptivo previo se decide que los posibles valores atípicos deben continuar en el análisis de la base de datos debido a que el estudio se interesa en esas particularidades que evidencia una diferencia entre las instituciones educativas del distrito de Barranquilla.

Continuando con el análisis de los datos de la prueba ICFES SABER 11, se aplicó la técnica de componentes principales obteniendo la Figura 3., donde se observa que la variabilidad de los datos se puede explicar en 81.36% con el uso de tres componentes principales. Ver Tabla 4.

**Figura 3**  
Análisis de Componentes Principales



Fuente: SPSS

-----

**Tabla 4**  
Valores propios de las componentes



	<i>inertia</i>	<i>cum</i>	<i>Ratio</i>
1	7,2504	7,2504	0,6042
2	1,4195	8,6699	0,7225
3	1,0935	9,7634	0,8136
4	0,6607	10,4241	0,8687
5	0,5529	10,9770	0,9148
6	0,3144	11,2914	0,9410
7	0,2444	11,5359	0,9613
8	0,1208	11,6567	0,9714
9	0,1094	11,7661	0,9805
10	0,0928	11,8589	0,9882
11	0,0768	11,9358	0,9946
12	0,0642	12,0000	1,0000

Fuente: SPSS

En el primer componente, ver Tabla 5., los aportes significativos son los del núcleo común. Entonces se puede afirmar que la primera componente se refiere a las asignaturas académicas por lo cual se denomina *academia*, siendo el de menor aporte la variable calendario (CAL) de -0,0729. La segunda componente es dominada por las variables naturaleza (NAT) y calendario (CAL) siendo sus aportes de 0,5820 y 0,5558 respectivamente. Por la afinidad de las variables que lo conforman este eje se denomina *diferenciador*, ya que claramente clasifica entidades oficiales y no oficiales. Para el tercer componente, los mayores aportes los ofrecen las variables número de evaluados (EVAL) y jornada (JOR), con un aporte de 0,5553 y 0,5590 respectivamente, siendo el menor aporte dado por la asignatura de sociales (SOC) con tan solo 0,0049. Este último eje se denomina *población estudiantil*, ya que se visualiza en este eje la diferencia entre entidades educativas con menor número de egresados de las de mayor número de egresados, lo cual además se puede notar en la Figura 4 del círculo unitario del primer plano factorial.

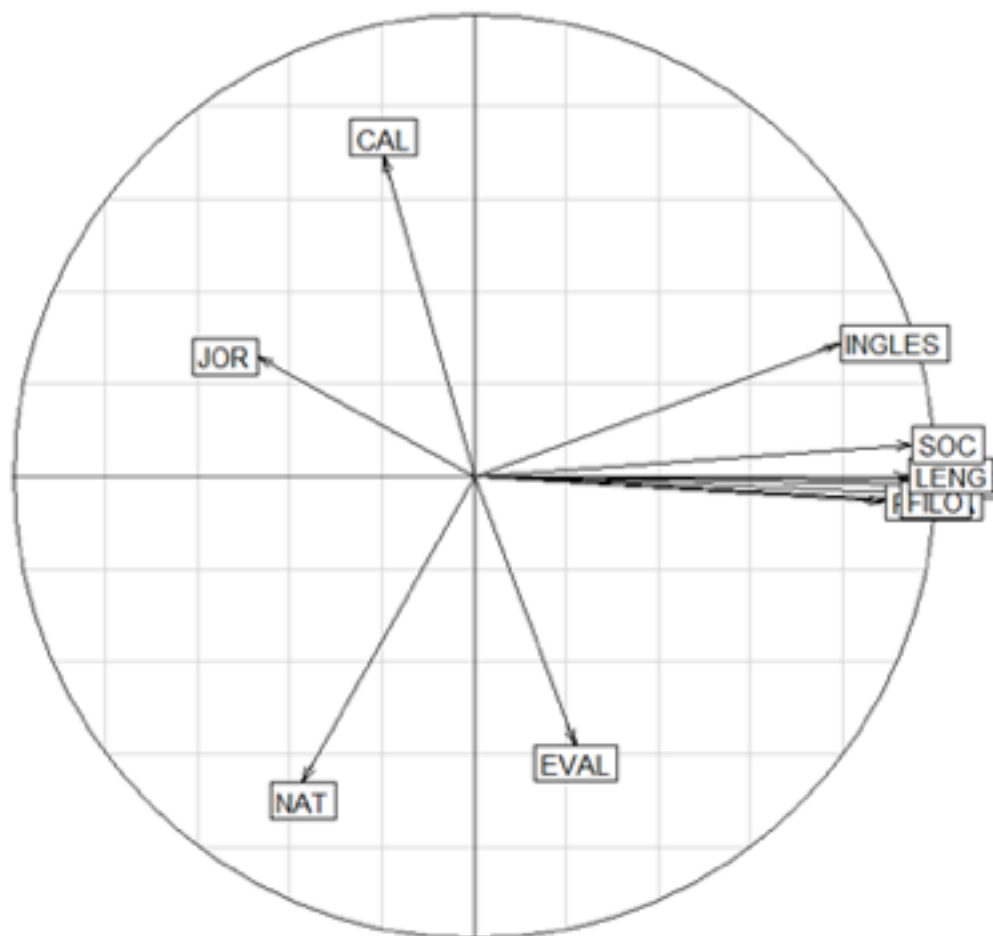
**Tabla 5**  
Aportes de las variables a las componentes

	CS1	CS2	CS3
<b>CAL</b>	-0,0729	0,5820	-0,5135
<b>JOR</b>	-0,1748	0,2181	-0,5590
<b>NAT</b>	-0,1380	-0,5558	-0,2960
<b>MATE</b>	0,3584	-0,0066	-0,0419
<b>QUI</b>	0,3484	-0,0306	-0,0812
<b>FISICA</b>	0,3312	-0,0467	-0,0857
<b>BIO</b>	0,3542	-0,0146	-0,0436
<b>FILO</b>	0,3442	-0,0431	-0,0184
<b>INGLES</b>	0,2952	0,2419	-0,0950
<b>LENG</b>	0,3507	0,0014	-0,0311
<b>SOC</b>	0,3532	0,0574	0,0049
<b>EVAL</b>	0,0816	-0,4876	-0,5553

Fuente: SPSS

-----

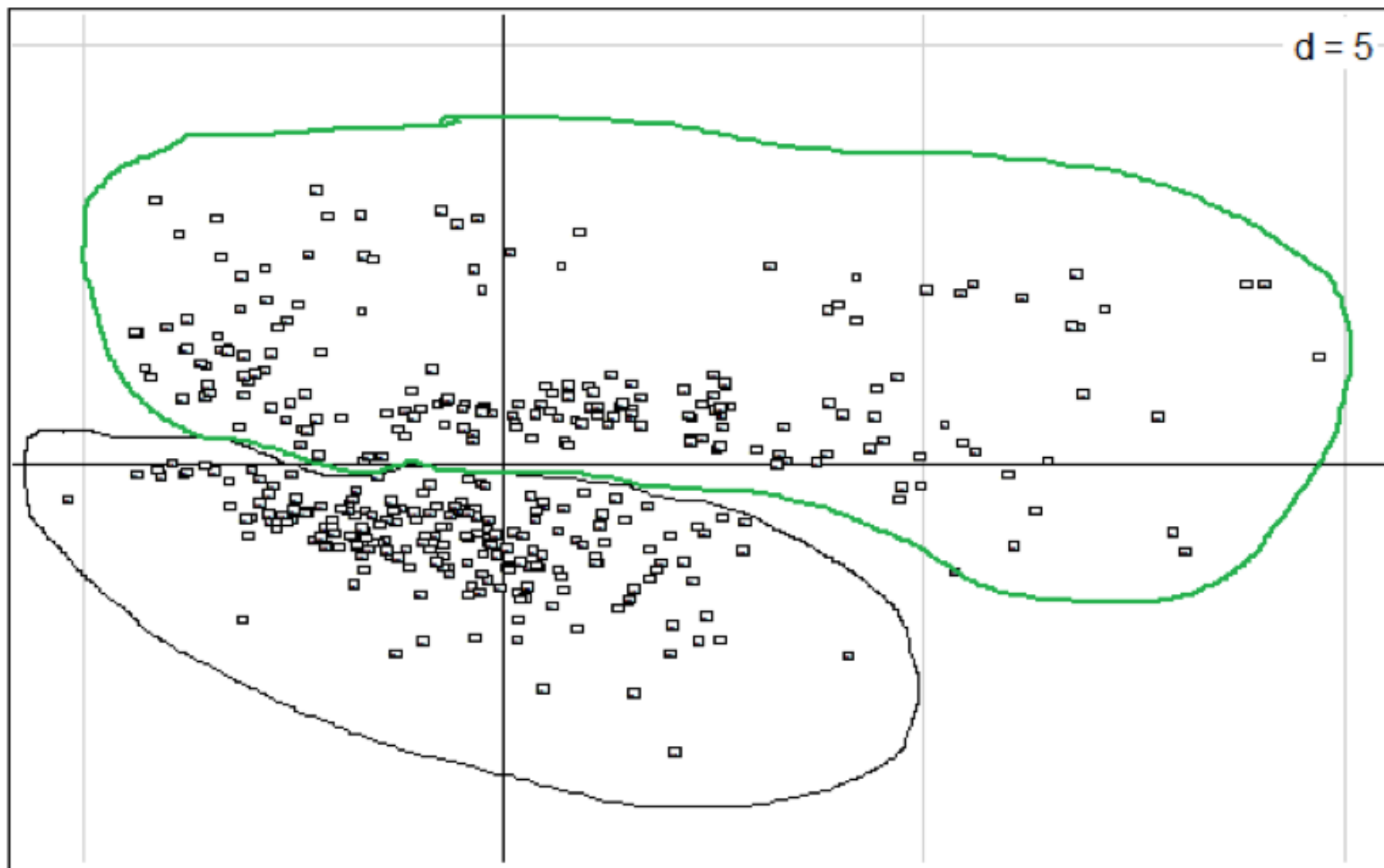
**Figura 4**  
Gráfica del círculo unitario



Fuente: SPSS

En la Figura 5 se visualizan dos nubes de puntos que explican la diferencia entre las instituciones oficiales y las no oficiales. Además, se observa una mayor concentración de las entidades educativas oficiales que se centran del lado izquierdo del primer componente, indicando un menor desempeño de las asignaturas del núcleo común, en contraste a las entidades no oficiales que se ubican en su mayoría del lado derecho, asociándose esto último a un mejor desempeño académico por parte de las entidades no oficiales.

Figura 5  
Entidades oficiales vs entidades no oficiales.



## 5. Conclusiones

Observando los valores propios de las variables de estudio y el porcentaje de varianza acumulada, se optó por disminuir de doce dimensiones a tres dimensiones, con una pérdida de información poco menos del 18%, lo cual se observa en la Tabla 2.

De acuerdo con las variables estudiadas y observando en la Tabla 3 los aportes de cada variable por componentes, se observa que en la primera componente se asocian con mayor peso las variables MAT, FILO, BIO, QUI, FISICA, INGLES, SOC, LENG, explicando el 60,42% de la variabilidad total. Ver la Tabla 2 y la Tabla 3.

Al segundo componente principal le corresponden las variables NAT y CAL que explican el 8,6699% de variabilidad total, mientras que al tercer componente se asocian las variables JOR Y EVAL explicando 9,7634% de la variabilidad total. Observar la Tabla 2 y la Tabla 3.

Tomando en cuenta que al primer componente se le asociaron las variables MAT, FILO, BIO, QUI, FISICA, INGLES, SOC, LENG, que mide el desempeño académico, se observa que las instituciones oficiales se ubicaron en el tercer cuadrante en su mayoría indicando un desempeño por debajo de 45 puntos en promedio de cada asignatura, no obstante a medida que se alejan a la izquierda su desempeño es menor. En contraposición, las instituciones educativas no oficiales en su mayoría se ubicaron en el primer cuadrante denotando un mayor desempeño en las pruebas con un promedio de más de 55 puntos en cada asignatura.

## Referencias bibliográficas

Barón, J., Bonilla, L., Cardona-Sosa, L., Ospina, M. (2013) ¿Quiénes eligen la disciplina de la

educación en Colombia? Caracterización desde el desempeño en las Pruebas Saber 11. Borradores de Economía Banco de la República, No. 785, 42pp

EURYDICE (2008). Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de los resultados. Disponible en el enlace [http://eacea.ec.europa.eu/education/eurydice/documents/thematic\\_reports/109ES.pdf](http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/109ES.pdf)

ICFES (1999), "Antecedentes Y Marco Legal Del Examen De Estado" ICFES.

González Zabala M. (2013). Propuesta de un modelo para evaluar la e-inclusión: el caso colombiano. Tesis Doctoral. Universidad Nacional.

ICFES (2010), Qué se evalúa; cómo se interpretan los resultados individuales, ICFES SABER 11°

ICFES (2012), Examen de validación del bachillerato académico; Qué se evalúa, ICFES SABER 11°

León, A., Llinás H., Tilano J. (2008) Análisis multivariado aplicando componentes principales al caso de los desplazados. Ingeniería y Desarrollo. no.23

López A., Roper Pacheco J., César Peralta J. (2011). Estudio de validez del examen de Estado Saber 11 de inglés. FOLIOS. N° 34. pp. 77-91

Martí Recober M., Ferrer Juliá F., Cuxart Jardí A., (1997) El desarrollo de la LOGSE: las nuevas pruebas de acceso a la universidad. Revista de Educación. No. 314. pp. 89-114.

MEN (2007), Al tablero No 42 septiembre-noviembre. Disponible en el enlace <http://www.mineducacion.gov.co/1621/article-137440.html>

MEN (2008), Al tablero No 44 enero-marzo. Disponible en el enlace <http://www.mineducacion.gov.co/1621/article-162342.html>

MEN (2010), Ministerio de Educación Nacional, Decreto N° 869: Por el cual se reglamenta el Examen de Estado de la Educación Media, ICFES SABER 11°.

Padilla, V., Concepción, A., y Ceja, J., (2012). El análisis de componentes principales aplicado a la revisión de los estados financieros de las empresas del índice habita de la Bolsa Mexicana de Valores. Observatorio de la Economía, 16 pp. Disponible en [http://www.eumed.net/cursecon/ecolat/mx/2012/bolsa\\_mexicana\\_valores.pdf](http://www.eumed.net/cursecon/ecolat/mx/2012/bolsa_mexicana_valores.pdf)

Peña D. (2002). Análisis de Datos Multivariados. Mc Graw Hill.

---

1. Doctorando en Ingeniería Industrial, Universidad Nacional Lomas de Zamora (Argentina). Docente de la Corporación Universitaria del Caribe-CECAR. [rafael.ruize@cecar.edu.co](mailto:rafael.ruize@cecar.edu.co)

2. Docente de la Institución Educativa Distrital Santa Bernardita. [jesus2009arevalo@gmail.com](mailto:jesus2009arevalo@gmail.com)

3. Docente Investigadora de la Corporación Universitaria Minuto de Dios, Barranquilla, Colombia. [gmorillo@uniminuto.edu.co](mailto:gmorillo@uniminuto.edu.co)

4. Docente Investigador de la Universidad Simón Bolívar, Barranquilla, Colombia.. [primitivo.acosta@unisimonbolivar.edu.co](mailto:primitivo.acosta@unisimonbolivar.edu.co)

---

Revista ESPACIOS. ISSN 0798 1015  
Vol. 39 (N° 10) Año 2018

[Índice]

[En caso de encontrar algún error en este website favor enviar email a [webmaster](mailto:webmaster)]

©2018. revistaESPACIOS.com • Derechos Reservados