

Una herramienta para reportar la investigación en el aula desde el análisis multivariable

A tool to report classroom research from multivariate analyzes

Un outil pour rendre compte de la recherche en classe
à partir d'analyses multivariées

Gerardo R Lara Morel

laragerardo@gmail.com

U.E.N Luis Eduardo Egui Arocha.
San Antonio de Los Altos. Venezuela

Artículo recibido en septiembre y publicado en diciembre 2020

RESUMEN

El artículo que se presenta ofrece una pauta para abordar la investigación docente en el aula de clase como un proceso natural y enfocado a generar conclusiones que posteriormente puedan ser evaluada para su ratificación o no. El docente puede aplicarla cuando emprende su estudio basado en la metodología de la investigación acción donde se obtendría gran cantidad de variables. Razón por la cual, se propone para su manejo el análisis multivariable, con apoyo del programa estadístico SPSS 25 como herramienta para obtener componentes principales o los factores que expliquen la variable dependiente. La información presentada en este escrito no pretende ser un manual, es una propuesta para los docentes que quieran iniciarse en este campo tengan un recurso a su alcance que le apoye en sus investigaciones en el aula.

Palabras clave: Análisis multivariable, componentes principales, factores, distribución normal, significancia.

ABSTRACT

The article that is presented offers a guideline to approach teaching research in the classroom as a natural process and focused on generating conclusions that

later may or may not continue to be ratified. The teacher can apply it when he approaches his study based on the action research methodology where a large number of variables will be obtained. For this reason, the multivariate analysis is proposed for its management, with the support of the SPSS 25 statistical program as a tool to obtain main components, factors or clusters that explain the dependent variable. The information presented in this writing does not pretend to be a manual, it is a proposal for teachers who want to start in this field to have a resource at their disposal to support their research in the classroom.

Key words: *Multivariate analysis, principal components, factors, normal distribution, significance.*

RÉSUMÉ

L'article qui est présenté offre une ligne directrice pour aborder la recherche pédagogique en classe comme un processus naturel et axé sur la production de conclusions qui pourront ou non continuer à être ratifiées ultérieurement. L'enseignant peut l'appliquer lorsqu'il aborde son étude sur la base de la méthodologie de recherche-action où un grand nombre de variables sera obtenu. Pour cette raison, l'analyse multivariée est proposée pour sa gestion, avec le soutien du programme statistique SPSS 25 en tant qu'outil permettant d'obtenir les principaux composants, facteurs ou groupes expliquant la variable dépendante. Les informations présentées dans cet écrit ne prétendent pas être un manuel, c'est une proposition pour les enseignants qui souhaitent se lancer dans ce domaine d'avoir une ressource à leur disposition pour soutenir leurs recherches en classe.

Mots clés : *analyse multivariée, principales composantes, facteurs, distribution normale, signification.*

INTRODUCCIÓN

En los tiempos actuales todas las instituciones de formación universitaria en docencia apuntan a un docente/investigador manteniendo esquemas que en algunos casos son rígidos y casi inflexible, pero el entorno escolar se ha visto alienado por factores totalmente diferentes que apuntan a una sociedad conectada virtualmente, pero desconectada socialmente. El dúo estudiante/docente se enriquece con componentes nuevos desde lo tecnológico y por ende el desarrollo

del proceso enseñanza y aprendizaje se ve afectado de una manera significativa, que según sea el caso puede favorecer lo aprehendido o dejándolo solo como información sin aplicabilidad ni utilidad en los estudiantes.

Este trabajo tiene como finalidad dejar un marco de referencia para abordar la investigación desde un ámbito multifocal, ver la mayor cantidad de variables que se están involucrando en el aula y detectar aquellas ausentes que son variables exógenas, que pueden influir de una o de otra manera en el proceso de enseñanza y aprendizaje por medio del Análisis Multivariado, con apoyo del Programa de Estadística SPSS.

¿Por qué el Análisis Multivariado? porque para desarrollar una investigación con rigor debemos ubicarnos en algunos de estos escenarios, como por ejemplos: Los estudiantes siendo objetos de estudios o los alumnos/docente como agente de participación. Podemos hasta lograr precisar una información variada de éstos, que va desde situaciones básicas hasta detalles muy especiales o particulares que están latentes en un aula, siendo estos: Los hábitos alimentarios, Traslado a la institución, Tiempo de Ocio, Rendimiento académico, entre otros.

DIMENSIONES A DETERMINAR SEGÚN EL ESTUDIO

Para procesar la información se utilizará el SPSS25, cuya traducción es Programa de Estadística para Ciencias Sociales, desarrollado bajo ambiente Windows que es fácil de trabajar, en los subsiguientes párrafos se describirá una pauta para poder cargar la data e interpretar los resultados de una investigación.

En la actualidad desarrollar una investigación en el aula sin observar las interrelaciones de los agentes intervinientes es como introducir un sesgo a la investigación. Es por esta razón que se hace la sugerencia de realizar el análisis estadístico de datos con métodos y técnicas multivariante que permiten estudiar y tratar en bloque una o varias variables medidas u observadas en una colección de individuos. Sin embargo, existe la posibilidad de que estas variables sean sólo

cuantitativas, o cualitativas, pero también está la viabilidad de que sean simultáneamente de ambos tipos.

Esta forma de abordar la investigación unido a la diversidad de enfoques teóricos y prácticos que puede darse a un estudio multidimensional ha invitado a los desarrolladores de programas estadísticos a crear herramientas que facilitan el análisis de grandes volúmenes de datos y un mínimo de conocimiento matricial (matemáticas) para reflexionar sobre las posibles causas y consecuencias del objeto de estudio. Al conocer y manejar estos tipos de programas se logra obtener excelentes y variados resultados; sólo quedará por parte del investigador su análisis, sugerencias, recomendaciones y conclusiones, bajo el paradigma desarrollado.

¿Qué son las técnicas y métodos del análisis multivariado?

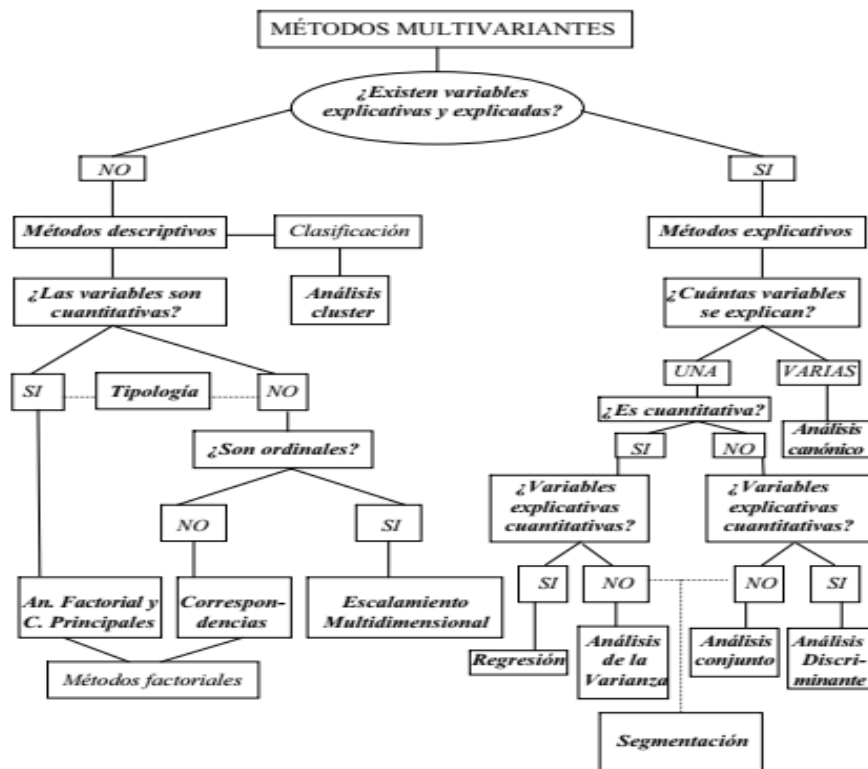
Parafraseando lo que expresa Pérez López (2004) es lo referido al conjunto de métodos y técnicas, así sean univariante o multivariante, que permiten estudiar y tratar en bloque una o varias variables medidas u observadas en una colección de individuos. Es decir, abordar la investigación con gran cantidad de información de los agentes involucrados para describir con una mayor exactitud nuestra pregunta de investigación.

El aula es una combinación de varios agentes que aportan una considerable fuente de información, están: los estudiantes, los profesores, los trabajadores sociales, el entorno familiar y muchos otros que en mayor o menor intensidad dan vida al proceso de enseñanza y aprendizaje. Por ejemplo: En un aula de 30 jóvenes se quiere conocer su rendimiento escolar; estaríamos ante un grupo de variables que podemos seleccionar, tales como: edad, género, número de hermanos, información sobre los padres, tiempo dedicado al estudio, si tiene celular, redes sociales, número de hora que dedica a las redes, número de hora que le dedica al estudio para realizar actividades asignadas o repaso, la preferencia de una(s)

asignatura(s) o el rechazo de otra(s) por parte del estudiante, número de hora en llegar al colegio, el manejo de su inteligencia emocional, estilo de aprendizaje, cómo describe a sus docentes, a sus amigos, estilo del docente en el aula y muchas otras variables que pueden presentarse a lo largo de la investigación.

Como se puede observar el volumen de datos para analizar es rico, variado y abundante, donde el docente/investigador puede abordarlo de forma independiente para obtener sus valores representativos para cada una de ellas y además tendría que realizar un análisis integrador para aproximarse al rendimiento escolar. Es aquí, donde el análisis multivariado permite agrupar variables y abordarlas a partir de estos componentes o factores y así obtener respuestas a la pregunta de investigación. Al emplear este tipo de análisis se debe elaborar una pauta para el éxito de la investigación. En el siguiente esquema N°1, observe los pasos a seguir según las interrogantes que usted como investigador desea estudiar.

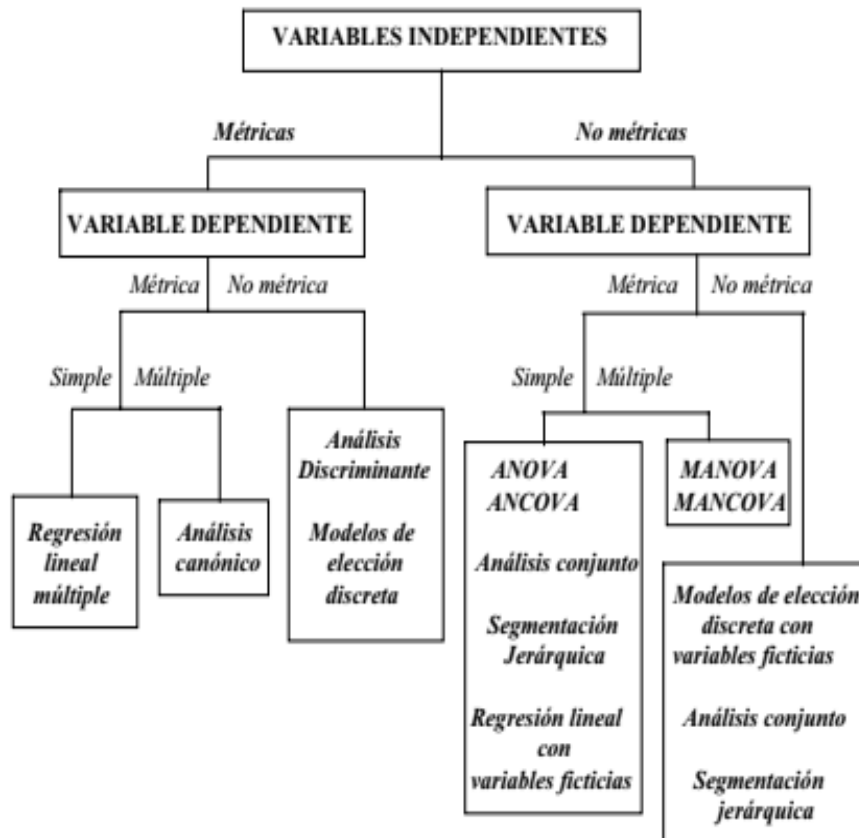
Esquema N°1 Análisis Multivariado



Tomado de Pérez López (2004)

Al emplear el anterior ejemplo sobre el curso de 30 alumnos, podríamos decir que *las variables explicativas* serían, por nombrar algunas, el género, la edad, el tiempo de desplazamiento al centro de estudio; y *la variable explicada*: el nivel de rendimiento, es decir; tendríamos que observar si esta relación existe entre ellas; entonces se emplearía para realizar el análisis, uno de los Métodos Explicativos. Más aún, se puede aproximar al análisis siguiendo la pauta del esquema N°2.

Esquema N°2 Variables



Tomado de Pérez López (2004)

A la luz de los esquemas anteriores, se puede establecer una visión desde las variables, la cual nos indica el camino a seguir para el desarrollo de la investigación, esto con la finalidad de reducir el número de variables, pero con un criterio basado

en el mayor número de información generada y una mayor variabilidad de los datos (Varianza). En el caso de que las variables sean numéricas se tendrían como opciones: Componentes Principales, Análisis Factorial, Análisis por conglomerado o Escalamiento multidimensional (en este artículo sólo se desarrollarán los dos primeros análisis), incluso, si nos orientamos por lo expuesto por Guardiola (2014) sobre la investigación multivariable, tendríamos tres técnicas más para desarrollar el análisis, las cuales las clasifica de la siguiente forma:

- Análisis de Dependencia: que trata de explicar la variable considerada independiente a través de otras variables independientes o explicativas.
- Análisis de Interdependencia: que es una técnica de clasificación y tiende a buscar las interrelaciones entre las variables y sus estructuras subyacente.
- Otras técnicas: que intenta superar lo tratado en las dos anteriores intentando explicar con procesos complejos los hallazgos.

El panorama para el investigador nobel parece tomar tramas difíciles, pero es aquí donde el programa estadístico entra en acción. Por eso, lo primero que debe saber el docente/investigador es que la descripción o explicación que se dará a continuación no indica un orden o jerarquía; primera, segunda,... sino como algo latente en el sistema de variables obtenido en una investigación.

¿Qué debe conocer el investigador? Él debe percatarse de la existencia de diferentes formas de abordar la interpretación de la data. Algunas de ellas son: el Análisis de Componentes Principales y el Factorial, entre otros. Estos tienen mucho en común, a la luz de lo que dice Calvo (2020), ellos buscan nuevas variables que expliquen los datos originales, estableciendo una cantidad de variables menores (reducción de las variables originales). Por un lado, las componentes principales, se argumenta en la explicación solo con variables originales, puede ser obtenido mediante expresiones matemáticas independientemente de su interpretaciones individuales o práctica. Y por el otro, el análisis factorial trabaja con las variables anteriores, pero además trata de encontrar variables condensadas ocultas,

inobservables y aún no medidas, cuya existencia se sospecha en las variables originales.

A continuación se presentaran algunos estudios y los diferentes gráficos que se pueden obtener de la aplicación de los diversos análisis e interpretaciones. El primer ejemplo estará referido al uso del *análisis de componentes principales* en un “estudio de nivel de rendimiento según las asignaturas y pruebas alternativas”, se graficaría de esta manera:

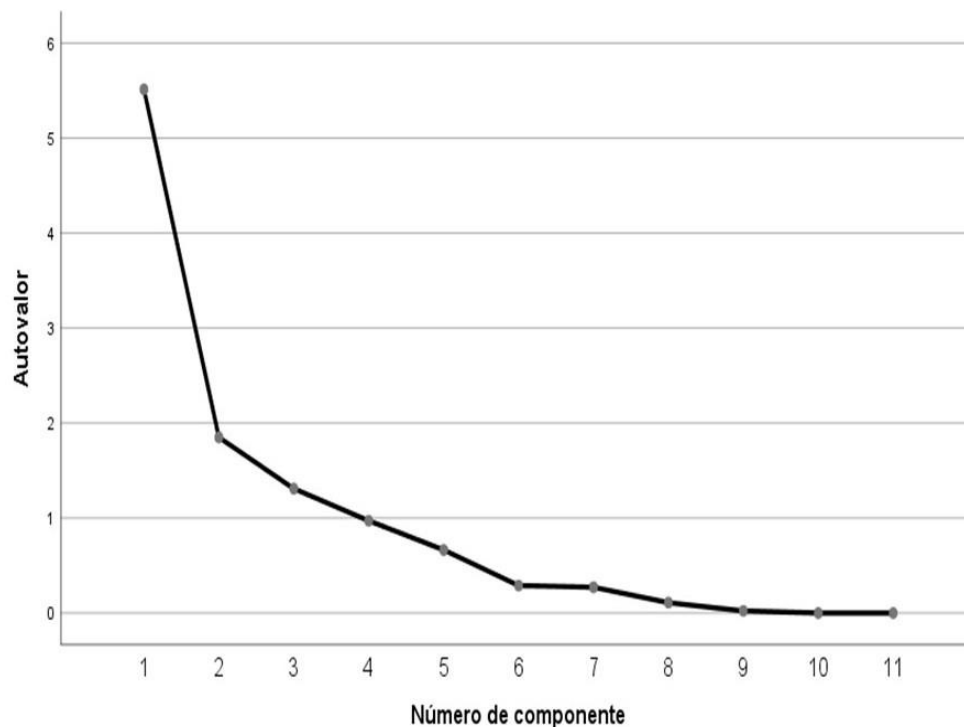


Gráfico N° 1

Línea de Sedimentación estudio sobre el “estudio de nivel de rendimiento según las asignaturas y pruebas alternativas”

Una posible interpretación del gráfico N° 1 sería: Se observa que los tres primeros componentes aportan la mayor información (valores superiores a 1 para los autovalores), para realizar el análisis de la variable por explicar.

Cuando el SPSS realiza análisis de factores, ya no se habla de componentes sino simplemente de factores y las relaciones presentes, en el siguiente gráfico N° 2 se representa como influyen las variables en los factores, además se puede observar el cambio de terminos en relación al gráfico anterior.

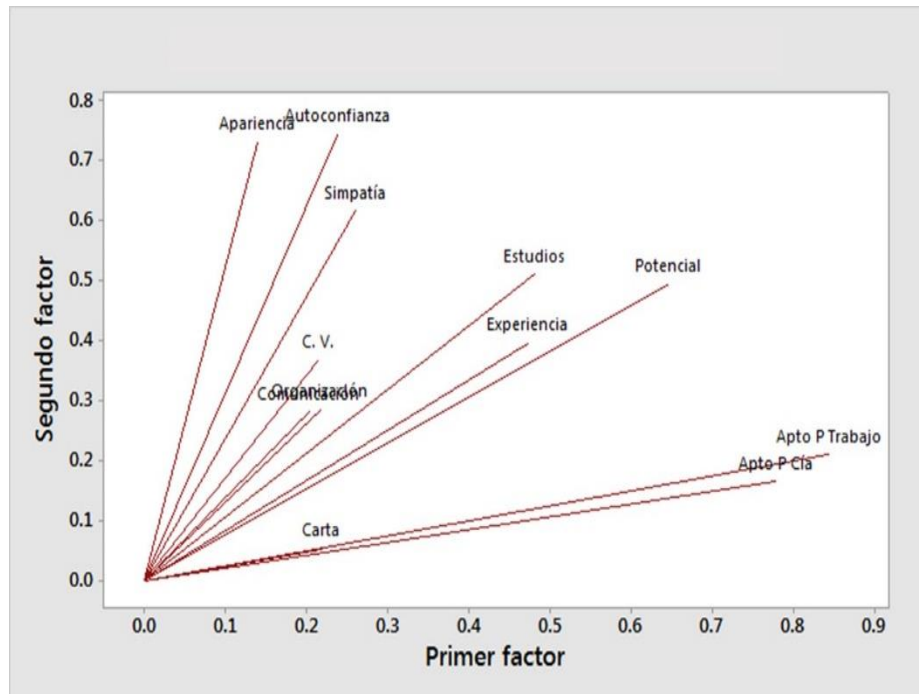


Gráfico N° 2
Influencia de las variables sobre los dos factores principales en el estudio sobre personalidad y los factores dados

Debemos tener presente que la escala va desde 0 hasta 1 y que los valores que se aproximen a estos 1 tienen peso en la construcción del factor, mientras que los valores cercanos al origen (cero) tienen poca influencia.

En circunstancias algo similares, nos encontramos cuando se utiliza *el análisis factorial*, porque esta es una técnica estadística de reducción de datos usada para explicar las correlaciones entre las variables observadas en términos de un número menor de variables no observadas llamadas factores. Si quisiéramos ahondar en la

investigación sobre el grupo de 30 estudiantes en un aula, podríamos correlacionar muchos de los agentes/variables mencionadas (relaciones personales, inteligencia emocional, sexo, edad, autoestima, ...).

Para ello lo ejemplificaremos de esta manera, en el gráfico N° 2 se puede observar como el segundo factor explica: Apariencia, Autoconfianza y Simpatía (parte superior izquierda). El primer factor explica casi solo (ver la parte inferior derecha) a carta, Apto para el Trabajo y Apto para Cuentas. Los dos factores juntos explican los elementos centrales tales como: Comprensión, Estudio, Potencial, etc.; con esta información arrojada, el investigador puede interpretar y explicar la variable dependiente en función de los dos factores obtenidos.

El programa estadístico

Después de haber leído la información anterior sobre algunas formas de abordar la interpretación de la data, nos corresponde ahora hablar del programa estadístico informático *Statistical Package for the Social Sciences* (SPSS) que es un software que le proporciona a los investigadores herramientas que permiten consultar datos y formular/comprobar hipótesis, así como desarrollar inferencias, ejecutar procedimientos para aclarar las relaciones entre variables, identificar tendencias y realizar predicciones. Un programa que fue desarrollado para el manejo de grandes volúmenes de datos donde, además de lo expuesto, es posible presentar tablas, gráficos, la validez de instrumentos, entre otras de sus cualidades.

Esta ambientado en el sistema operativo Windows muy semejante a una hoja de Excel, pero con mayor cantidad de aplicaciones. Estas serán descritas en las próximas líneas a través de ejemplos, para que el lector pueda seguir la secuencia en su quehacer docente/investigador y así poder visualizar las propiedades de este software, como también podrá percibir como se obtienen un resultado o varios, dependiendo de lo se este indagando en el estudio. Además, podrá observar que todas las técnicas que se emplean con el programa tienen un fuerte basamento

matemático, y será con apoyo del SPSS que abordaremos en los análisis e interpretación de los datos a la luz de los argumentos desarrollados en el misma.

Abordando el análisis de los datos

Lo primero que se debe hacer cuando tenemos una data nueva, es conocer el tipo de variable, para así poder determinar su estadístico más apropiado. Para ello se determinará según sea el caso: la media, moda, mediana, varianza, desviación, asimetría, gráfico Q-Q, etc. Lo antes mencionado se denomina análisis estadístico descriptivo, que es un tipo de análisis que tiene como finalidad conocer la distribución de la data recogida y detectar si la muestra proviene de una distribución normal o no. Aquí es cuando el SPSS en unos cuantos pasos logra responder todas estas preguntas o necesidades previas después de cargar la data.

Para ello es necesario seguir los siguientes pasos que se encontrará en la barra de herramienta, debe ubicar la palabra “Analizar”, se seleccione; luego despliega una ventana donde elige la “Estadística descriptiva” y más tarde se extiende otra opción que es la “Frecuencia”, ver figura N° 1; (a partir de aquí cada paso a seguir, seleccionar o elegir será identificado de esta manera, separado con /; Analizar/Estadística descriptiva/frecuencia); además se escogieran las variables y los diferentes estadísticos como: la Media, la Moda, la mediana, etc. y los gráficos (histograma+curva normal) para caracterizarlas, ver figura N° 2. Para mayor comprensión de lo descrito anteriormente, observe las siguientes figuras.

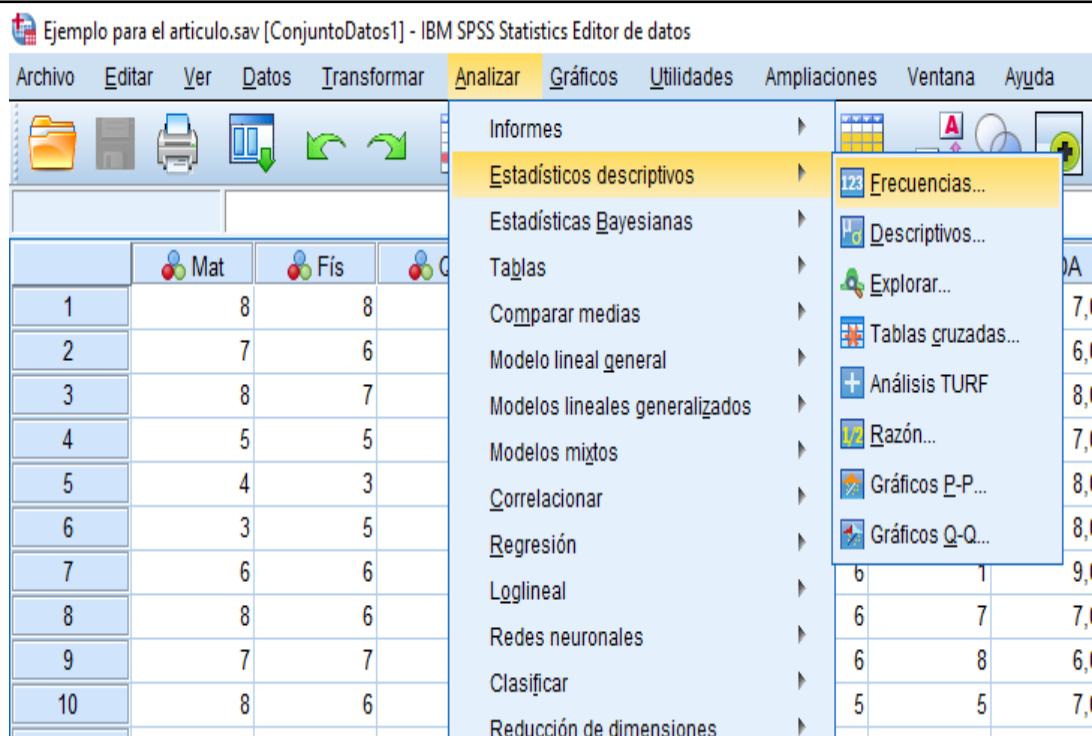


Figura N° 1
Lineas de comando para caracterizar variables

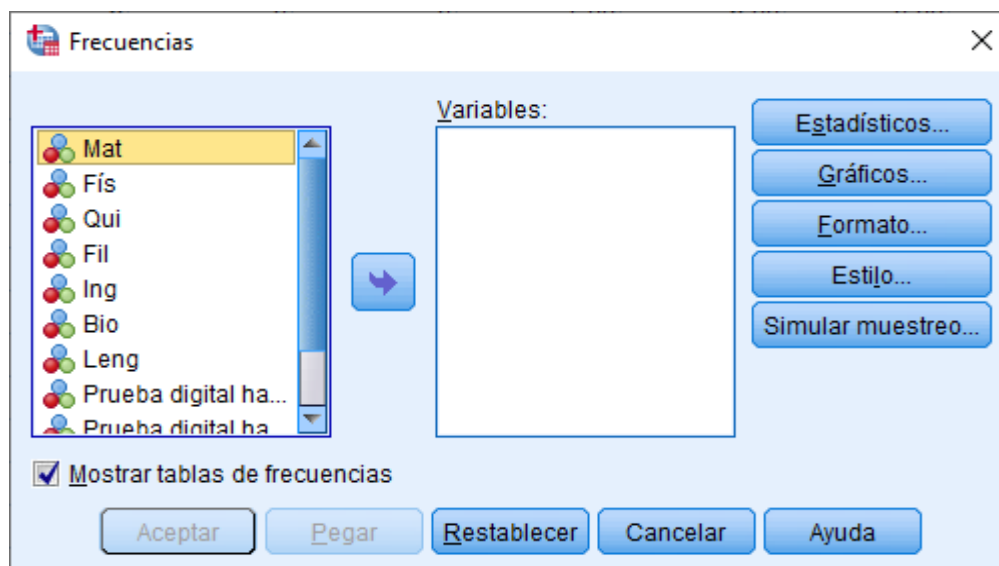


Figura N° 2
Lineas de comando para seleccionar estadísticos y gráficos

Si usted como investigador desea ahondar en la interpretación de su estudio, sólo podría indicarle al SPSS que haga otro procedimiento y elegiría la opción de: Analizar/ estadística descriptiva/Explorar, ver figura N° 3, todavía más, podría seguir seleccionando otras funciones: estadísticos descriptivos y gráficos (histograma+gráfico de prueba de normalidad).

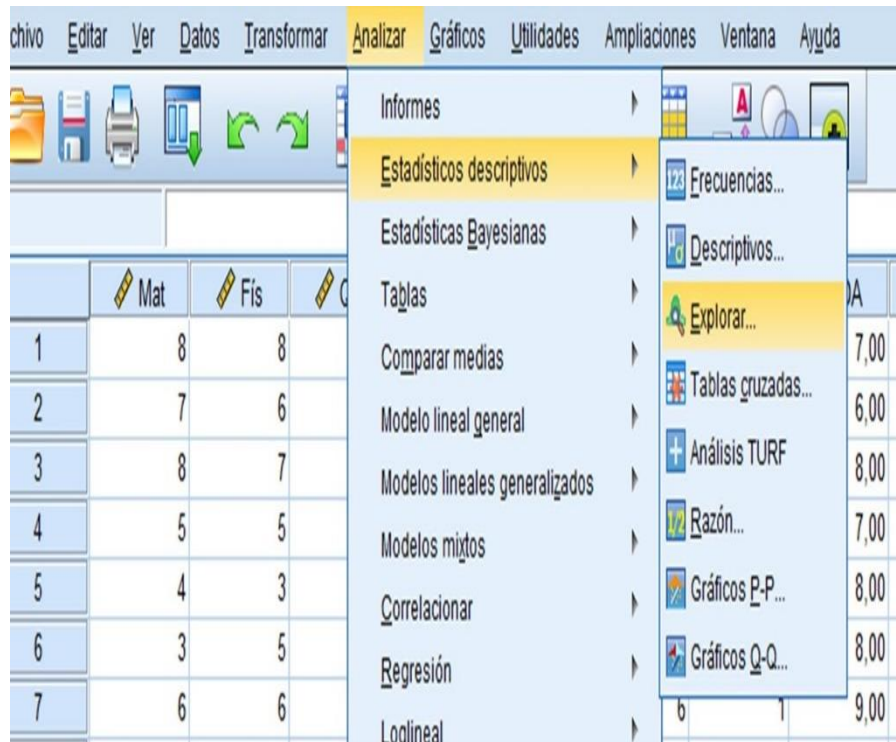


Figura N° 3
Línea de comando para seleccionar el análisis estadístico descriptivo

Estos gráficos de normalidad son complementarios a los probabilísticos. Así, el gráfico Q-Q, que explica los datos de la variable frente a los datos esperados si la distribución fuera normal. Si observamos el gráfico N° 3, sobre el estudio del promedio de notas de un curso de matemática, notaremos que los puntos estarán cerca de la diagonal y así podemos decir que la distribución es normal.

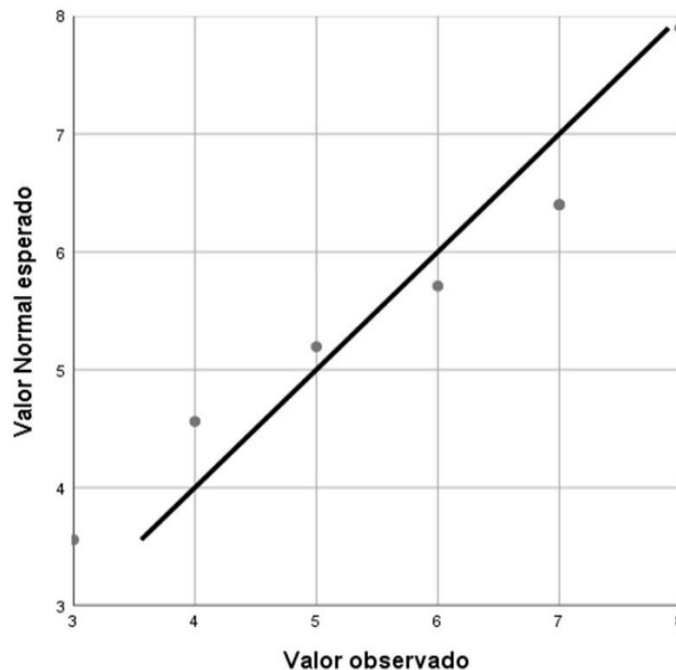


Gráfico N°3

Gráfico Q-Q Normal del promedio de un curso de Matemática

Observe que los valores obtenidos están muy cerca de la línea (Diagonal) bien sea por encima o por debajo de la misma, esto sugiere que la data recogida sigue una distribución normal para la variable Matemática.

¿Qué más nos ofrece este programa SPSS? Este software tiene la virtud de ejecutar otras representaciones como los Gráficos de Cajas y Bigotes, y es a través de esta graficación que se obtiene la mediana. La caja está definida por el segundo y tercer cuartil, mientras que los bigotes por el primero y el cuarto, por lo que dentro de la caja tenemos el 50 % de los datos de la muestra. Este tipo de representación es también útil para detectar valores atípicos, ver gráfico N° 4, sobre un estudio relacionado al promedio de un curso de física.

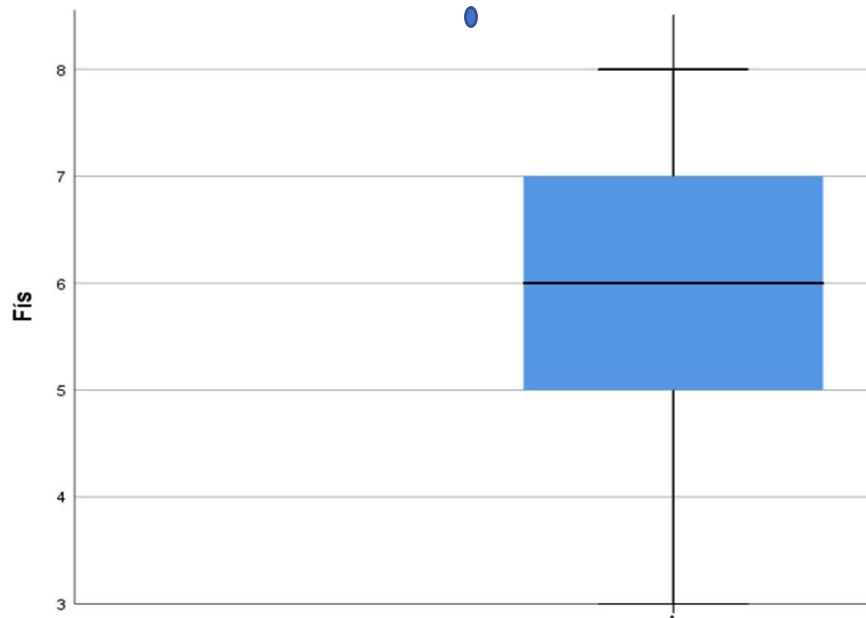


Gráfico N ° 4

Diagrama de caja y bigotes en el estudio sobre el promedio de un curso de física

El gráfico muestra como el primer cuartil (parte inferior del diagrama) se aleja dos unidades del segundo cuartil, mientras que el cuarto se aleja del tercero una unidad, lo que sugiere que la curva de valores se desplaza con valores altos a la derecha y descende hacia la izquierda. Además, se puede observar un caso atípico; que esta por encima del bigote.

Además de lo descrito anteriormente, en el SPSS existen diferentes tests estadísticos que nos evalúan la normalidad. Si seguimos los pasos anteriores obtenemos el test Kolmogorov-Smirnov (K-S) cuya hipótesis nula, H_0 , es que la distribución de la variable seleccionada proviene de una distribución normal. Por ejemplo, si el nivel de significación o p-valor (Sig.) obtenido en el test (K-S) es 0.20, entonces para un nivel de significación del 0.05 (lo que está fuera del 95 % de probabilidades) no rechazamos la hipótesis nula, ya que el p-valor es $0.20 > 0.05$. Por tanto, según este test, podemos considerar que la distribución de los datos es normal. En resumen:

En el test Kolmogorov-Smirnov (K-S):

Si Sig. (p-valor) > 0.05 aceptamos H_0 (hipótesis nula) > distribución normal.

Si Sig. (p-valor) < 0.05 rechazamos H_0 (hipótesis nula) > distribución no normal, ver cuadro N° 1 que corresponde a verificación de la normalidad de la muestra.

Cuadro N°1

Ejemplo sobre pruebas de normalidad en una investigación

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Mat	,228	10	,150	,847	10	,053

a. Corrección de significación de Lilliefors

Todo lo antes descrito tiene como finalidad: la descripción, comparación, reducción, clasificación, explicación o previsión de la data a través del ciclo de operaciones del programa SPSS de esas informaciones obtenidas de una serie determinada de objetos (Escobar, 2013). Entonces, se puede decir que el trabajo del investigador se afianza sobre estos pilares, como son: recoger, organizar, analizar e interpretar lo obtenido mediante el SPSS, como se verá a continuación.

El programa ya tiene los algoritmo listo, sólo necesita de los insumos suficiente para arrojar unos resultados que debemos interpretar. Como se ve en la figura N° 4.

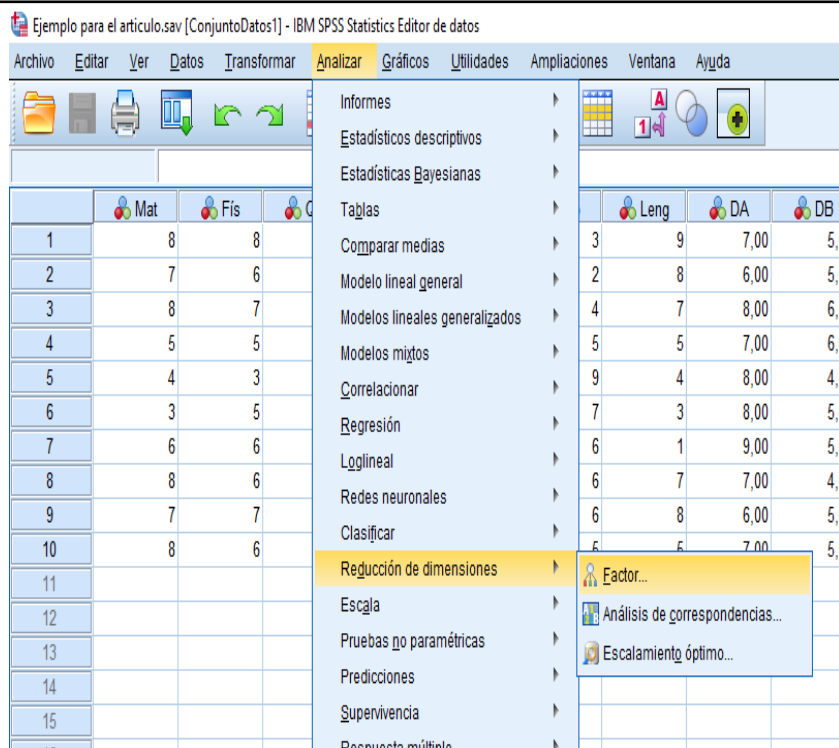


Figura N ° 4
Algoritmos del SPSS

Al seleccionar un tipo de análisis, por ejemplo, “*Reducción de dimensiones*”, cuya proceso nos permitirá condensar la información de un conjunto de variables en un nuevo conjunto de variables, con el menor coste de información posible; al ejecutar la opción de “*factor*” nos enviará a una nueva ventana donde debemos seleccionar lo que necesitamos para esa reducción, ver figura N° 5...; pongamos como caso la escogencia de factor, éste nos llevara a la opción de “Análisis factorial”.

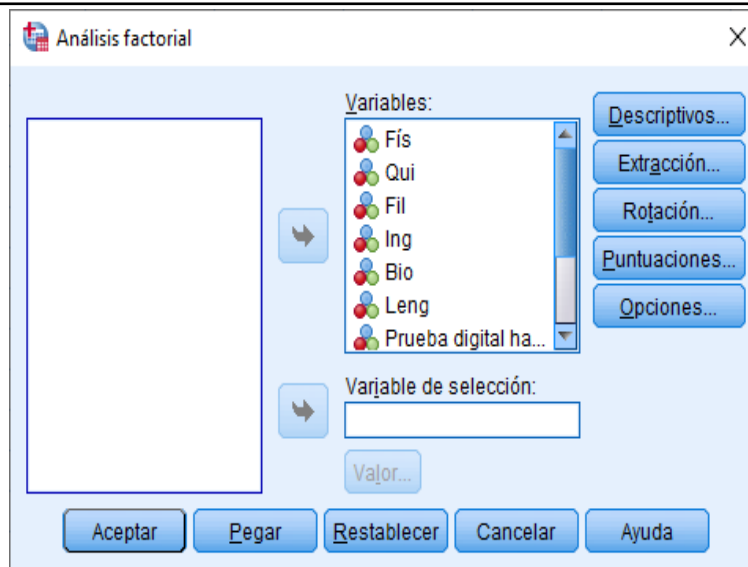


Figura N° 5
Ventana del análisis factorial

Los botones: Descriptivos, Extracción, Rotación, Puntuación y Opciones tienen funciones específicas; cuyas características se detallarán a continuación...

Al aparecer en la pantalla el análisis factorial, este nos ofrece una ventana que de nuevo nos indica lo que necesitamos hacer, en este caso escogeremos la pestaña de "Descriptivos", ver figura N° 6.

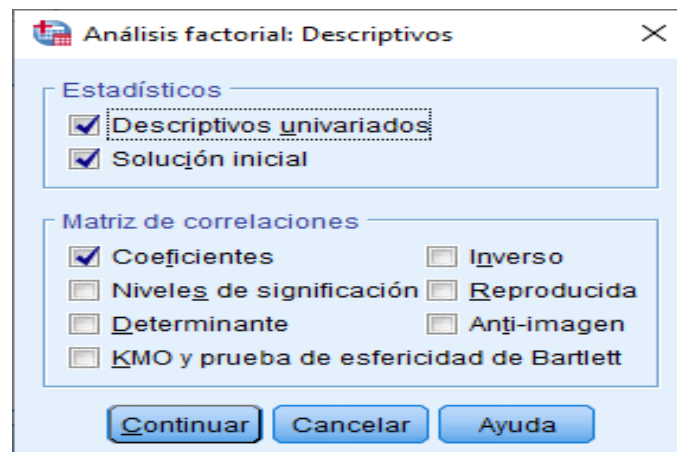


Figura N° 6
Ventana de selección de análisis factorial: Descriptivos

En este estadístico descriptivo univariado incluyen la media, desviación típica y el número de casos válidos para cada variable. La solución inicial muestra las communalidades iniciales, los autovalores y el porcentaje de varianza explicada que resaltan información necesaria para el análisis de componentes principales. Obsérvese que en la figura N° 6 en el recuadro de matriz de correlaciones Se dispone de la opción: coeficientes que muestra el aporte por variable a cada componente.

Al seleccionar estas tres casillas (continuar/aceptar) el programa SPSS reporta la información solicitada con sus respectivas identificaciones y leyendas en una sola hoja continua; ver figura N° 7.

Estadísticos descriptivos			
	Media	Desv. Desviación	N de análisis
Mat	6,40	1,838	10
Fis	5,90	1,370	10
Qui	4,40	2,221	10
Fil	6,20	1,932	10
Ing	5,80	1,317	10
Bio	5,30	2,003	10
Leng	5,70	2,541	10

Matriz de correlaciones								
Correlación	Mat	Fis	Qui	Fil	Ing	Bio	Leng	
	Mat	1,000	,768	,419	,682	,606	,640	
	Fis	,768	1,000	,270	,453	,813	,717	
	Qui	,419	,270	1,000	,487	,084	,030	
	Fil	,682	,453	,487	1,000	,498	,614	
	Ing	,606	,813	,084	,498	1,000	,910	
	Bio	,640	,717	,030	,614	,910	1,000	
	Leng	,647	,597	,055	,439	,651	,614	

Comunalidades		
	Inicial	Extracción
Mat	1,000	,815
Fis	1,000	,767
Qui	1,000	,911
Fil	1,000	,719
Ing	1,000	,862
Bio	1,000	,846
Leng	1,000	,689

Método de extracción: análisis de componentes principales.

Varianza total explicada						
Componente	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	4,313	61,620	61,620	4,313	61,620	61,620
2	1,295	18,495	80,115	1,295	18,495	80,115
3	,519	7,412	87,527			
4	,494	7,063	94,590			
5	,225	3,215	97,805			
6	,112	1,593	99,399			
7	,042	,601	100,000			

Método de extracción: análisis de componentes principales.

Matriz de componentes ^a		
	Componente	
	1	2
Mat	,875	,222
Fis	,874	,051
Qui	,309	,903
Fil	,742	,410
Ing	,885	,281
Bio	,882	,260
Leng	,759	,336

Método de extracción: análisis de componentes principales.
a. 2 componentes extraídos.

Figura N° 7
Reporte que arroja el programada de SPSS

Para facilitar la comprensión al lector de la anterior figura, en los siguientes cuadros se interpretará la información obtenida.

Cuadro N° 2A

Ejemplo sobre los estadísticos descriptivos sobre las notas del curso en las materias seleccionada

	Media	Desviación	N de análisis
Mat	6,40	1,838	10
Fís	5,90	1,370	10
Qui	4,40	2,221	10
Fil	6,20	1,932	10
Ing	5,80	1,317	10
Bio	5,30	2,003	10
Leng	5,70	2,541	10

En este cuadro se observa la media y la desviación estándar de cada asignatura, además de colocar los números de observaciones presente en el análisis. Con esta información el investigador puede hacer un acercamiento de explicación de cómo se comportan las variables en la investigación. Es decir, se puede observar como la desviación estándar nos explica que tan rápido se separa los datos de la media en cada asignatura.

Cuadro N° 2B

Ejemplo Matriz de correlación sobre las notas del curso en las materias seleccionada

		Mat	Fís	Qui	Fil	Ing	Bio	Leng
Correlación	Mat	1,000	,768	,419	-,682	-,606	-,640	,647
	Fís	,768	1,000	,270	-,453	-,813	-,717	,597
	Qui	,419	,270	1,000	-,487	-,084	-,030	-,055
	Fil	-,682	-,453	-,487	1,000	,498	,614	-,439
	Ing	-,606	-,813	-,084	,498	1,000	,910	-,651
	Bio	-,640	-,717	-,030	,614	,910	1,000	-,614
	Leng	,647	,597	-,055	-,439	-,651	-,614	1,000

Esta matriz muestra como se correlacionan entre pares, por ejemplo, Matemática/Física, al cruzar ambas filas y columnas se observa el valor 0.768 y así sucesivamente con cada uno de los pares. Otra forma de interpretar la matriz, sería hacer el cruce de columnas con filas, nótese que el resultado sigue siendo el mismo: Física/Matemática. Además, denotar, que los valores oscilan entre -1 y 1 y se interpreta de la siguiente manera: si se aproxima a 1 se dice que la correlación es fuerte y directa, en caso contrario, se dice que es fuerte pero inversa.

Cuadro N° 2C

Ejemplo Comunalidades obtenidas sobre las notas del curso en las materias seleccionada

	Inicial	Extracción
Mat	1,000	,815
Fís	1,000	,767
Qui	1,000	,911
Fil	1,000	,719
Ing	1,000	,862
Bio	1,000	,846
Leng	1,000	,689

El cuadro 2C muestra el Método de extracción: análisis de componentes principales. Esta información indica como sería el aporte de cada una de las variables a dicho proceso. Por ejemplo: El investigador podría interpretar con esta comunalidad que el grupo de estudiantes no tienen un buen rendimiento en lenguaje, filosofía y física.

Cuadro N° 2D

Ejemplo Varianza total explicada sobre las notas del curso en las materias seleccionada

Componente	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	4,313	61,620	61,620	4,313	61,620	61,620
2	1,295	18,495	80,115	1,295	18,495	80,115
3	,519	7,412	87,527			
4	,494	7,063	94,590			
5	,225	3,215	97,805			
6	,112	1,593	99,399			
7	,042	,601	100,000			

El cuadro N° 2D se observan los componentes; en el caso del componente 1 su porcentaje es de 61,62% de varianza. Si realizan el análisis con las dos primeras componentes (1-2) se explicaría con un 80,115% de varianza acumulada y así sucesivamente.

Aquí es donde el investigador debe elegir el camino, entre dos, tres o más componentes (recordando siempre que se quiere minimizar el número de los mismo).

Cuadro N° 2E

Matriz de componentes principales

	Componente	
	1	2
Mat	,875	,222
Fís	,874	-,051
Qui	,309	,903
Fil	-,742	-,410
Ing	-,885	,281
Bio	-,882	,260
Leng	,759	-,336

Método de extracción: análisis de componentes principales.

a. 2 componentes extraídos.

Esta matriz 2E nos indica el aporte de cada variable al componente. Es decir, si observamos el componente 1 en el cuadro, las matemáticas (0,875), física (0,874), Química (0,309) y lengua (0,759) aportan positivamente a dicha componente, sin embargo, el resto de las asignaturas reportan cargas negativas a dicho componente.

Al regresarnos al SPSS, nos encontramos que dentro del análisis factorial a parte del estadístico descriptivo tenemos el de Extracción, tiene por objetivo determinar el número mínimo de factores comunes capaces de reproducir, de un modo satisfactorio, las correlaciones observadas entre las variables; este se realiza seleccionando Componentes Principales cuyo objetivo es maximizar la varianza explicada. El factor que mejor explica la dimensión analizada será el primero y así sucesivamente, ver figura N° 8.

Como nuestro interés es el análisis de componentes principales, realizaremos la siguiente selección: Método/componentes principales. Ver el recuadro “Mostrar”, esta la opción: *Solución factorial sin rotar* y el *Gráfico de sedimentación de los*

autovalores. Luego se selecciona en el recuadro “Extraer”: se escoge la opción *Basado en autovalor*, esta propiedad permite elegir todos los factores cuyos autovalores excedan un valor especificado (por defecto 1) o elegir un número específico de factores. Por último, esta Número máximo de iteraciones para convergencia en donde se puede especificar el número máximo de pasos que el algoritmo sigue para estimar la solución.

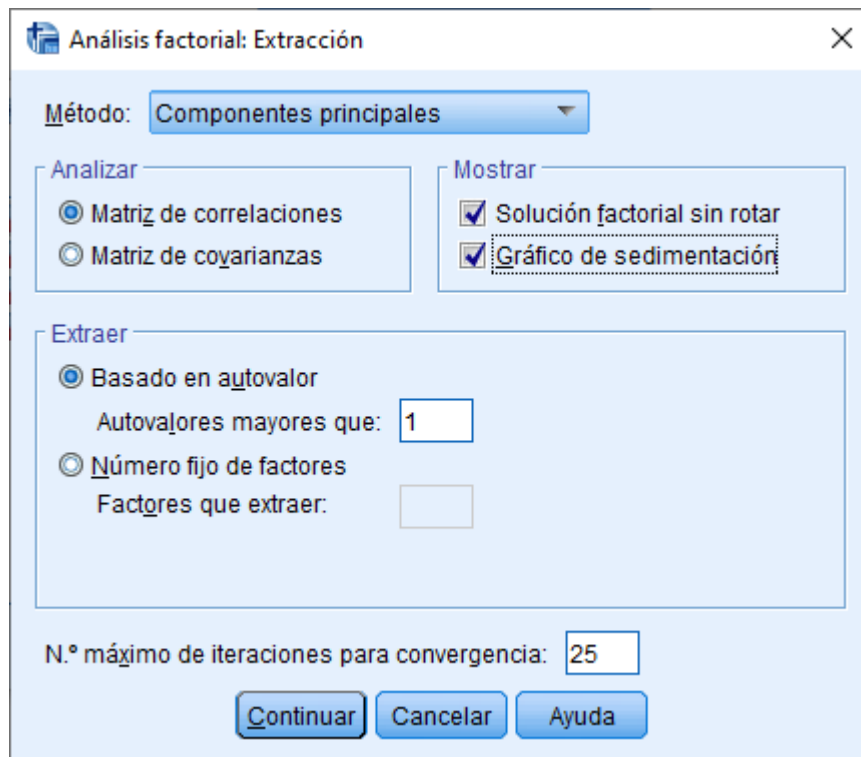


Figura N° 8

Ventana de selección de análisis factorial: Extracción

Una vez aceptada las selecciones, el SPSS reportará un gráfico (ver N° 5) y unos cuadros, que se detallan a continuación:

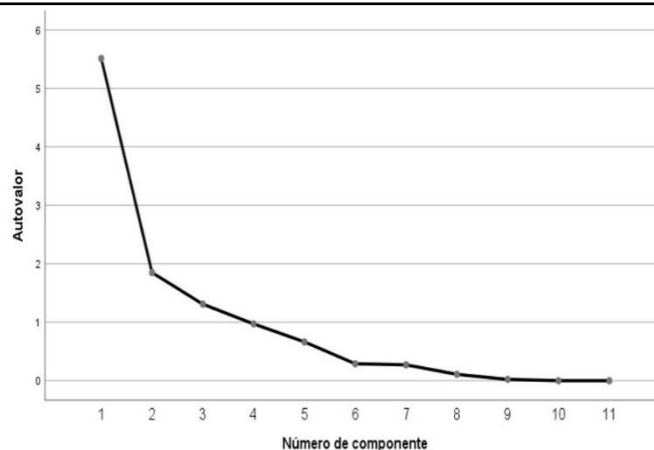


Gráfico N° 5
Sedimentación

Este gráfico N° 5 nos explica la variable en estudio ya que sus valores están por encima de 1, además muestra los componentes y su aporte a la descripción de lo que se desea explicar.

En el siguiente cuadro N° 3 se observa la misma presentación de los cuadros 2A hasta 2E, por lo tanto, la explicación sería la misma.

Cuadro N° 3
Análisis de extracción de los componentes principales

Matriz de componente ^a			
	Componente		
	1	2	
Mat	,875	,222	
Fis	,874	-,051	
Qui	,309	,903	
Fil	-,742	-,410	
Ing	-,885	,281	
Bio	-,882	,260	
Leng	,759	-,336	
Método de extracción: análisis de componentes principales.			
a. 2 componentes extraídos.			
Comunalidades		Extracción	
Mat		,815	
Fis		,767	
Qui		,911	
Fil		,719	
Ing		,862	
Bio		,846	
Leng		,689	
Método de extracción: análisis de componentes principales.			
Varianza total explicada			
Componente	Sumas de cargas al cuadrado de la extracción		
	Total	% de varianza	% acumulado
1	4,313	61,620	61,620
2	1,295	18,495	80,115
Método de extracción: análisis de componentes principales.			

Por último, en la ventana de análisis factorial identificada como la figura N° 5, se encuentra otro sobre reducción de dimensiones por “*Factorial de Rotación*” que corresponde a procedimientos que tratan de obtener factores más interpretables mediante la transformación de la solución inicial. Tiene como objetivo la búsqueda de soluciones factoriales en las que cada factor tenga correlaciones altas con un grupo de variables y baja con el resto.

Existen dos tipos de procedimiento para rotar los factores: 1.- la ortogonal extrae factores no correlacionados entre sí, y 2.- la oblicua se basa en el supuesto de intercorrelación entre los factores. Este análisis Factorial de Rotación dispone de variados métodos, ver cuadro N° 4

Cuadro N° 4

Descripción de los métodos de rotación de factores

Rotaciones ortogonales: Factores perpendiculares, no correlacionados	
Varimax	Minimiza el número de variables que tienen cargas altas en cada factor. Simplifica la interpretación de los factores
Equamax	Minimiza el número de factores necesarios para explicar cada variable. Simplifica la interpretación de las variables observadas.
Quartimax	Combina el método varimax, que simplifica los factores, y el método quartimax, que simplifica las variables. Se minimiza tanto el número de variables que saturan alto en un factor, como el número de factores necesarios para explicar una variable.
Rotaciones oblicuas: Factores no necesariamente perpendiculares, correlacionados	
Oblimin	Es el método más utilizado para el cálculo de las rotaciones oblicuas. Calcula el grado de oblicuidad de los factores en función del parámetro delta, que permite ponderar la maximización de la matriz por filas o por columnas. En función del objetivo a perseguir se puede modificar este parámetro, que en los programas estadísticos al uso suele ser por defecto 1.
Promax	Calcula los factores a partir de una matriz construida analíticamente partiendo de una solución ortogonal hasta crear una solución factorial lo más cercana posible a la estructura ideal.

Tomado de: López-Aguado y Gutiérrez-Provecho (2019)

Luego de ser vista en forma general la teorización del análisis factorial por rotación, podemos centrarnos en la Figura N° 9 cuyo cuadro permitirá visualizar la selección del método de rotación que necesita el investigador para su estudio.

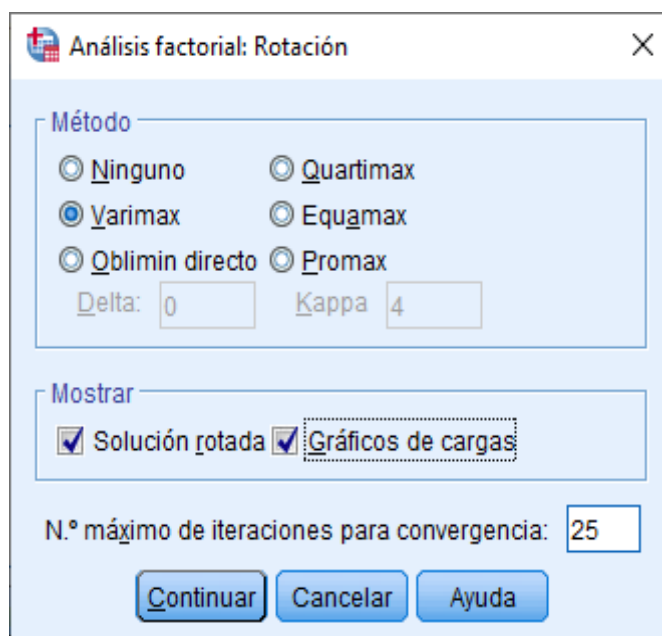


Figura N° 9

Ventana en el SPSS sobre Análisis Factorial: Rotación

En esta ventana encontramos diferentes métodos, pero como estamos centrados en el análisis de componentes principales, seleccionamos el método Varimax, que tiene como finalidad calcular las varianzas máximas con el menor número de componentes.

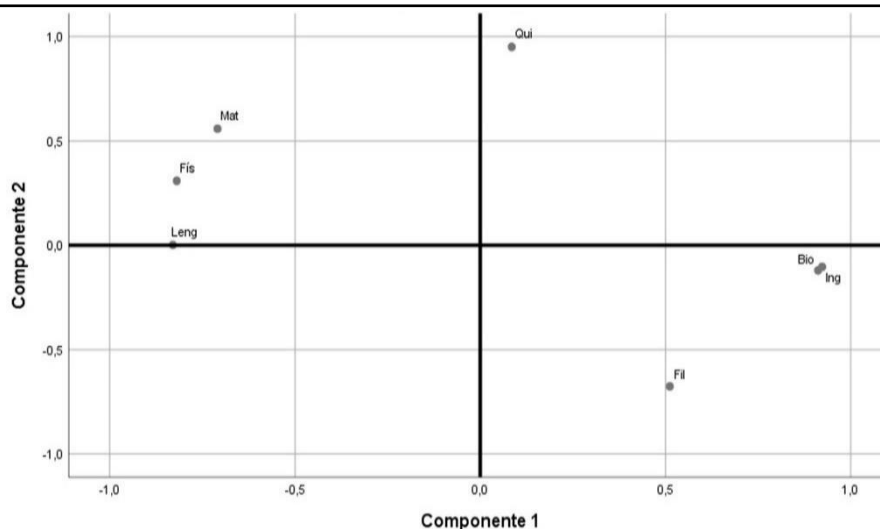


Gráfico N° 6

Componente en espacio de rotación

El gráfico N° 6 presenta el aporte de cada variable, indicando si son aporte positivo o negativos según en caso. Por ejemplo, Química (Qui) para el componente 1 tiene un coeficiente cercano a 0,0, mientras que para el componente 2 el coeficiente de aporte es próximo al 1. Al observa Matemática (Mat) para la componente 1 su coeficiente es negativo y a la izquierda de -0,5; para la componente 2 es superior a 0,5.

Nota

En todas las ventanas de diálogo el botón Continuar permite pasar a la ventana de diálogo Análisis factorial para seguir fijando especificaciones. Una vez elegidas se pulsa Aceptar para obtener los resultados del Análisis Factorial.

CONCLUSIONES

Con la pauta se establece un puente entre los docentes de aula y la estadística aplicada a la investigación desarrollada en sus aulas para desarrollar teorías que fortalezcan los procesos de enseñanza aprendizaje.

Crear conciencización de la investigación en el aula como un proceso más que se puede abordar como algo cotidiano y se puede mejorar en el tiempo evaluando los procesos desarrollados como un sistema en constante evolución.

REFERENCIAS

- Calvo, D. (2020). *Diferencias entre análisis factorial y análisis de componentes principales*. [Documento en línea] Disponible en: <https://www.diegocalvo.es/diferencias-entre-analisis-factorial-y-analisis-de-componentes-principales-pca/> [Consultado: 2020, mayo 30]
- Escobar, M. (2013). *El análisis multivariable*. [Documento en línea] Disponible en: <https://sociocav.usal.es/blog/modesto-escobar/files/2013/01/Escobar2005a.pdf> [Consultado: 2020, junio 13]
- Guardiola J., Plácido (2014). *Análisis Multivariantes*. [Blog Universidad de Murcia]. Disponible en: https://www.um.es/docencia/pguardio/documentos/Tec_seg.pdf [Consultado: 2020, julio 14]
- López-Aguado, M. y Gutiérrez-Provecho, L. (2019). Cómo realizar e interpretar un análisis factorial exploratorio utilizando SPSS. [Artículo en línea] *Revista d'Innovació i Recerca en Educació*. 12(2). [Documento en línea] Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=7057076>. [Consultado: 2020, julio 30]
- Pérez López, C. (2004). *Técnicas de análisis multivariante de datos*. Madrid: Pearson Educación, S.A.