

HJ-BIPLLOT COMO METODOLOGÍA EXPLORATORIA PARA EL ANÁLISIS MULTIDIMENSIONAL DE LOS OBJETIVOS DE DESARROLLO SOSTENIBLE (ODS) A NIVEL MUNICIPIOS EN BOLIVIA

HJ-BIPLLOT AS AN EXPLORATORY METHODOLOGY FOR THE MULTIDIMENSIONAL ANALYSIS OF THE SUSTAINABLE DEVELOPMENT GOALS (SDG) AT THE MUNICIPALITY LEVEL IN BOLIVIA

Carlos Fernando Silva Viamonte¹
Universidad de Salamanca, Salamanca, España,
✉ fersilvania@usal.es

Artículo recibido: 02/01/2024
Artículo aceptado: 28/03/2024

RESUMEN

La Agenda 2030 y los Objetivos de Desarrollo Sostenible (ODS) se presentan como oportunidades clave para reconsiderar prácticas y abordar los desafíos de desarrollo, medir y analizar el cumplimiento de estos objetivos es de vital importancia para una planificación objetiva. Por tanto, se destaca la importancia de un análisis estadístico multidimensional descriptivo, tales como, las técnicas de análisis Biplot, metodologías usadas por ciencia de datos, inteligencia artificial y machine learning, como un nuevo paradigma para comprender la información de manera más profunda y como base para el diseño de políticas con alto impacto social. La medición continua y evaluación de indicadores ODS se consideran esenciales para ajustar y mejorar las políticas públicas a lo largo del tiempo. En este sentido, el artículo presenta el HJ-Biplot como una técnica de análisis multivariante, una herramienta analítica avanzada para interpretar grandes volúmenes de información, en este caso, el cumplimiento a los ODS a nivel municipio, logrando clústeres a nivel departamental y nacional de estos de acuerdo a la similitud y disimilitud como insumo para el diseño de políticas públicas objetivas y de alto impacto.

Palabras clave: Estadística multivariante, métodos Biplot, HJ-Biplot, ODS, Clústeres.

ABSTRACT

The 2030 Agenda and the Sustainable Development Goals (SDGs) are presented as key opportunities to reconsider practices and address development challenges. Measuring and analyzing compliance with these goals is of vital importance for objective planning. Therefore, the importance of a descriptive multidimensional statistical analysis is highlighted, such as Biplot analysis techniques, methodologies used by data science, artificial intelligence and machine learning, as a new paradigm to understand information in a deeper way and as a basis for the design of policies with high social impact. The continuous measurement and evaluation of SDG indicators are considered essential to adjust and improve public policies over time. In this sense, the article presents the HJ-Biplot as a multivariate analysis technique, an advanced analytical tool to interpret large volumes of information, in this case, compliance with the SDGs at the departmental and municipal level, achieving clusters of these according to their similarity and dissimilarity.

Key words: Multivariate statistics, Biplot methods, HJ-Biplot, SDGs (Sustainable Development Goals), Clusters.

¹ Coordinador de Gerencia General del Banco de Desarrollo Productivo BDP-SAM, Licenciado en Administración de Empresas, Máster en Estadística Aplicada, Máster en Economía, Doctor en Investigación. Estudiante de Doctorado en Estadística Multivariante Aplicada. ORCID: [0000-0001-5117-6565](https://orcid.org/0000-0001-5117-6565).

1. INTRODUCCIÓN

Todos los días nos enfrentamos a un mundo cada vez más interconectado, complejo, inestable y sobre todo rápidamente cambiante, los nuevos retos se añaden a los ya conocidos. Los niveles de pobreza y desigualdad en nuestro país, el cambio climático; la violencia e inseguridad social; la digitalización y la robotización de nuestras vidas muestran que nos desenvolvemos en un contexto claramente complejo y es oportuno e importante entenderlo, aprender y desarrollar instrumentos que nos ayuden a tomar las mejores decisiones para enfrentarlo y aprovechar de sus características, de sus cualidades y de las oportunidades que nos brinda.

Bolivia tiene ventajas comparativas y capacidades tanto en el sector público como en el privado, para articular políticas públicas y acciones colectivas en el marco de la nueva agenda global de desarrollo sostenible. En ese sentido, la Agenda 2030 está ya en el centro de todas las mesas de discusión del mundo, Bolivia y todos los gobiernos subnacionales no deberían ser la excepción.

Lograr que el mundo sea más sostenible y equitativo en 2030 solo se alcanzará si se tiene un entendimiento real de la situación en la que vivimos, para ello, debemos definir nuevas herramientas y metodologías de análisis objetivas, que se caractericen por la rigurosidad matemática y estadística necesaria dejando de lado los métodos de trabajo convencionales. En ese entendido, trabajar con firmeza hacia una evaluación técnica científica y a partir de ello diseñar nuevas estrategias y políticas más eficientes para consolidar un consumo y producción sostenibles, una distribución igualitaria de los recursos y un enfoque transversal por los derechos humanos y por ende la mejora de la calidad de vida de los habitantes.

Un análisis estadístico descriptivo desde una perspectiva multidimensional, en términos de cantidad y variabilidad hoy se constituye en un nuevo paradigma en el que se justifica la ciencia de datos, la inteligencia artificial y el *Machine Learning*, técnicas estadísticas que nos ayudarán a analizar e interpretar la información con un enfoque completamente distinto a las técnicas convencionales de tal forma que nos ayudarán a diseñar políticas objetivas y de alto impacto social minimizando el uso de recursos escasos necesarios.

2. METODOLOGÍA

Los antecedentes básicos para el presente trabajo se constituyen en todo aquello que se desarrolló en los últimos años en el marco de dos temáticas independientes y al mismo tiempo complementarios.

Inicialmente, se tiene todo lo referido a los Objetivos de Desarrollo Sostenible ODS en cuanto a su diseño, su socialización e implementación en el mundo y por otro lado, está el desarrollo de la ciencia de datos, específicamente la inteligencia artificial, el *Big Data*, el *Machine Learning* y todas las técnicas de análisis de datos en términos matemático estadísticos que se vienen desarrollando e implementando como una herramienta fundamental a la hora de darle un valor agregado a la información, a través de diferentes algoritmos que se programan en diferentes softwares, y que en estos últimos años han dado la posibilidad de trabajar con grandes volúmenes de información e identificar patrones que antes la estadística clásica no podía identificar.

Dada la agenda ambiciosa que tienen los países, Bolivia no es la excepción, con los 17 Objetivos de Desarrollo Sostenible (ODS) que componen la Agenda 2030, objetivos universales determinados y ratificados, en

2015, por los 193 países que conforman la Asamblea General de la Organización de las Naciones Unidas (ONU). Que básicamente apunta a un mundo sin pobreza, con gente saludable, educada y productiva, viviendo en comunidades prósperas con todas sus necesidades básicas satisfechas. Con el propósito de ir mejorando es necesario establecer evaluaciones periódicas acerca de la situación en la que se encuentran estos objetivos con sus respectivas metas a través del uso de herramientas que ayuden a interpretar y explorar la información de manera multidimensional.

Los datos usados en este artículo provienen del Atlas Municipal de los Objetivos de Desarrollo Sostenible en Bolivia 2020, este dossier presenta una ficha de diagnóstico multidimensional por municipio. El territorio de Bolivia se divide política y administrativamente en departamentos, provincias, municipios y Territorios Indígena Originario Campesino (TIOCs). El Atlas nos da la oportunidad de contar con datos agregados a nivel país, datos correspondientes a los nueve departamentos y datos que corresponden a cada uno de los 339 municipios.

2.1 Objetivos del estudio.

El objetivo fundamental es explorar a través de la técnica HJ-Biplot los ODS's municipales desde una perspectiva multidimensional, es decir, la identificación de los municipios de acuerdo al cumplimiento de sus ODS's y de esa forma, crear clústeres de acuerdo a la similaridad de los mismos a nivel nacional y departamental.

HJBplot es una técnica poderosa que se puede utilizar para visualizar y analizar la relación entre variables en un conjunto de datos multivariante con alta calidad de representación a nivel municipio y

cumplimiento de ODS's (medidos de 0 a 1, donde valores mayores, cercanos a 1 siempre son mejores, y un valor de 0,5 significa que el municipio está en la mitad del camino hasta el valor óptimo, mientras que un valor de 0,75 significa que está en tres cuartos del camino).

2.2 Métodos Biplot.

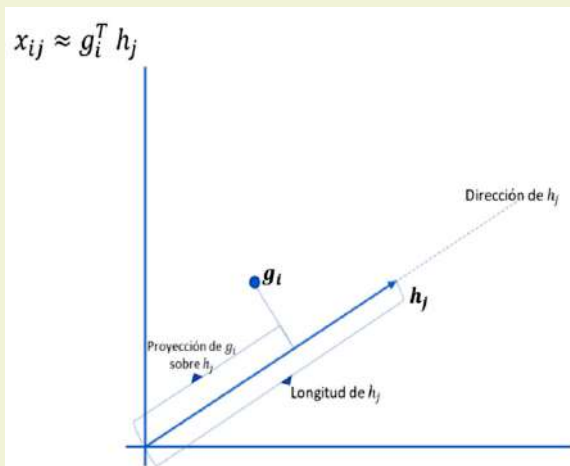
Los métodos Biplot fueron propuestos inicialmente por Gabriel (1971) como representaciones gráficas de datos multivariantes, es decir, al igual que un diagrama de dispersión muestra la distribución conjunta de dos variables, un Biplot representa tres o más variables (Gabriel y Odoroff, 1990); son pues, técnicas multivariantes. Usualmente, las filas de la matriz son representadas mediante puntos (marcadores fila) y las columnas con vectores (marcadores columna), siguiendo la terminología introducida por el autor.

Formalmente se definen de la siguiente manera: un Biplot para una matriz de datos $X_{n \times p}$ (arreglo rectangular con n filas y p columnas) es una representación gráfica mediante marcadores $g_1, g_2, g_3, \dots, g_n$ para las filas de la matriz de datos X y $h_1, h_2, h_3, \dots, h_p$ para las columnas de X , de forma que el producto escalar $g_i^T h_j$ aproxime el elemento x_{ij} de la matriz de partida, tan bien como sea posible (Gabriel, 1971). El producto escalar, en el que se basa, es un concepto matemático que en un principio podría suponer una barrera para el usuario, pero su traducción geométrica es sencilla. En este documento, por ejemplo, los datos están contenidos en una matriz $X_{339 \times 15}$ que tiene en filas los 339 municipios del Estado Plurinacional de Bolivia y en columnas, los 15 ODS. Así, para cada fila i (cada municipio en nuestro caso) y cada columna j (indicadores de cumplimiento de los ODS) aparece en la matriz de datos un valor x_{ij} que es el valor de ese marcador j para

el municipio i . Un Biplot permite representar la fila i de la matriz de datos (un municipio) con el marcador g_i y la columna j con el vector h_j , de forma que al proyectar el punto g_i sobre el vector h_j , esa proyección coincide con el valor que ese centro ha tenido para ese indicador.

El interés práctico reside en que el orden de las proyecciones de cada marcador fila sobre un marcador columna reproduce el orden de la matriz de partida, de forma que analizando la posición de cada unidad (municipio) sobre cada variable (indicador de cumplimiento del ODS), es posible ordenar las unidades en función del valor que toman en ese indicador, y eso puede hacerse para todos y cada uno de los indicadores (ver Figura 1).

Figura 1 Representación geométrica del producto escalar

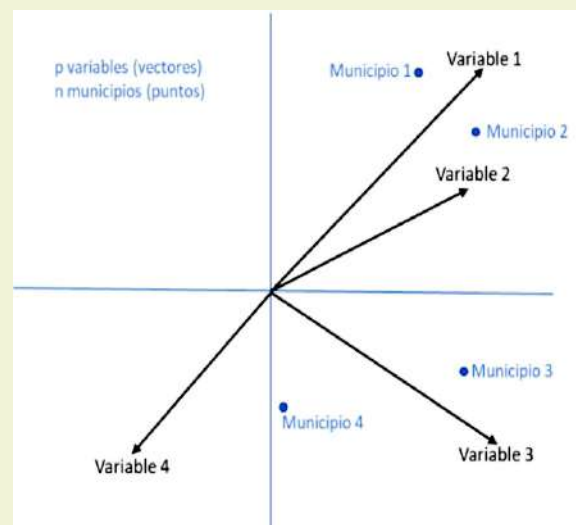


Hay infinitas formas de representar un Biplot pero solo algunas tienen propiedades interesantes en el análisis de datos. Los Biplots propuestos originalmente por Gabriel (1971) fueron dos: JKBiplot, en el cual las filas son representadas con la máxima calidad de representación (fiabilidad de las posiciones de los puntos que representan a los centros) y GH-Biplot en el cual las columnas son representadas con la máxima calidad, pero no las filas.

Galindo (1986) propone el HJ-Biplot como una representación gráfica multivariante de los datos de una matriz $X_{n \times p}$, mediante marcadores $j_1, j_2, j_3, \dots, j_n$ para las filas y $h_1, h_2, h_3, \dots, h_p$ para las columnas, elegidos de forma que ambos marcadores puedan ser superpuestos en un mismo sistema de referencia con máxima calidad de representación. Al presentar filas y columnas idéntica bondad de ajuste es posible interpretar no sólo la posición de las filas y de las columnas, sino también las relaciones fila-columna. Los ejes que conforman el sistema de referencia son las Componentes Principales del espacio de los indicadores.

Las reglas para la interpretación del HJ-Biplot son una combinación de las reglas empleadas en otras técnicas como el Escalamiento Multidimensional, el Análisis de Correspondencias, el Análisis Factorial y los Biplots clásicos (Galindo y Cuadras, 1986). En la Figura 2 se muestra un ejemplo con cuatro variables y cuatro individuos (municipios).

Figura 2 Interpretación HJ-Biplot.



El análisis se ha llevado a cabo a través del programa MultiBiplot desarrollado por

HJ-Biplot como metodología exploratoria para el análisis multidimensional de los objetivos de desarrollo sostenible (ODS) a nivel municipios en Bolivia

Vicente-Villardón (2010) en el entorno de programación orientado a matrices MATLAB.

Para la selección del tipo de Cluster se han aplicado, con fines exploratorios, Cluster jerárquicos y se ha afinado la solución mediante los métodos no jerárquicos, en concreto, se usó el método K-means y como medida la distancia euclídea.

3. ANÁLISIS DE RESULTADOS.

Para la primera parte se han retenido tres ejes factoriales pues, de esa forma, se consigue una inercia acumulada muy elevada de 83,6%, más que suficiente para caracterizar, con garantías, los municipios en relación a las variables de alcance de los objetivos de desarrollo.

Tabla 1 Valores propios y varianza explicada

Ejes	Eigenvalor	Var. Expl.	Inercia
			Acumulada
Eje 1	1666889.97	65.651	65.651
Eje 2	326143.74	12.845	78.496
Eje 3	128796.25	5.073	83.569
Eje 4	97398.83	3.836	87.405
Eje 5	79182.02	3.119	90.523

Fuente: Elaboración propia.

Atendiendo a las contribuciones del factor al elemento para las columnas (ver Tabla 2), se observa que todas las variables han de interpretarse en el plano factorial 1-2, a excepción del ODS15, que queda mejor recogido en el plano 1-3.

En cuanto a las filas, de los 339 municipios todos han quedado bien recogidos en los tres primeros ejes.

Tabla 2 Calidad de representación para columnas

Columna	Ejes		
	Eje 1	Eje 2	Eje 3
ODS1	488	458	1
ODS2	546	1	60
ODS3	833	89	12
ODS4	17	303	5
ODS5	777	67	11
ODS6	631	158	0
ODS7	13	818	3
ODS8	297	81	12
ODS9	641	177	1
ODS10	482	125	17
ODS11	700	150	6
ODS13	674	74	76
ODS15	35	2	775
ODS16	829	30	10
ODS17	940	3	0

Elaboración: Propia.

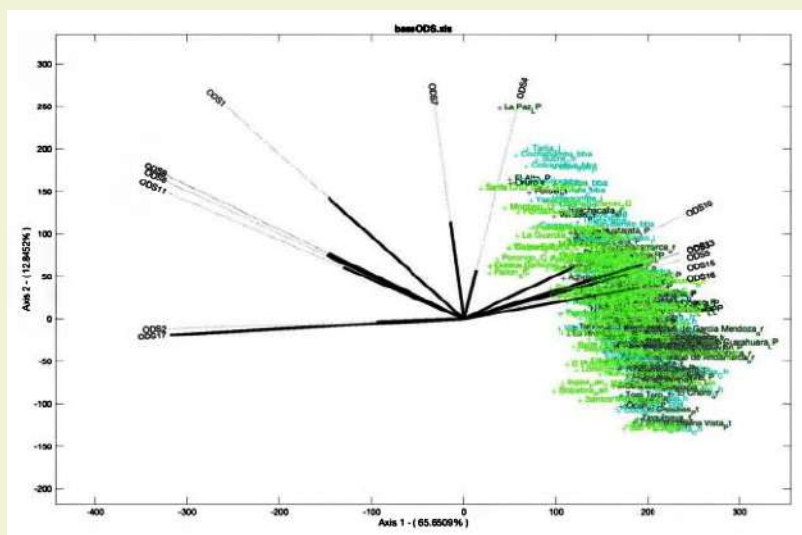
3.1. Análisis del impacto y la colaboración: plano 1-2

En el Biplot se muestra el gráfico factorial del plano 1-2 (Gráfico 1), donde la inercia acumulada asciende al 78,5%. Los cumplimientos de los objetivos de desarrollo sostenible (es decir, el cumplimiento de la agenda de los Objetivos de Desarrollo Sostenible) se representan mediante vectores, mientras que los municipios se identifican mediante puntos, cuya etiqueta incluye el nombre y la abreviación al departamento al que pertenecen.

Estos vectores indican la relación existente o inexistente y si la misma es directa y fuerte o no, como se observa en el Gráfico 1, podríamos dividir a las variables (cumplimiento de los objetivos) en tres grupos.

El grupo 1, con los objetivos ODS5, ODS8, ODS10, ODS13, ODS15 y ODS16 con una relación muy fuerte por el ángulo que forman sus vectores, de la misma forma, el grupo 2, se tienen los objetivos ODS1, ODS9, ODS6 y ODS11, y el grupo 3 con los objetivos los ODS2 y ODS17.

Gráfico 1 Representación factorial resultante del HJ-Biplot, plano 1-2



Existen otras relaciones, aunque son más débiles. Adicionalmente se puede observar que entre el grupo 1 y grupo 2 existe una relación inversa debido a que el ángulo entre los vectores, se aproximan a 180° . Finalmente, entre el grupo 1 y grupo 2 respecto al grupo 3, se puede observar una relación casi nula por ser perpendiculares entre si.

Con respecto a las filas (los municipios) se puede ver claramente que los mismos se ubican prácticamente en una misma área

lo que indica que presentan las mismas características o similares en lo que respecta al grupo 1 (Igualdad de Género, Trabajo Decente y Crecimiento Económico, Reducción de Las Desigualdades, Vida de Ecosistemas Terrestres y Paz Justicia e Instituciones Sólidas). También se encuentran cerca de los vectores ODS4 y ODS7 (Educación de Calidad y Energía Asequible y No Contaminante), principalmente el municipio de La Paz respecto al objetivo Educación de Calidad con 83,2% de cumplimiento.

Gráfico 2 Correlación entre vectores (cumplimiento de los objetivos)

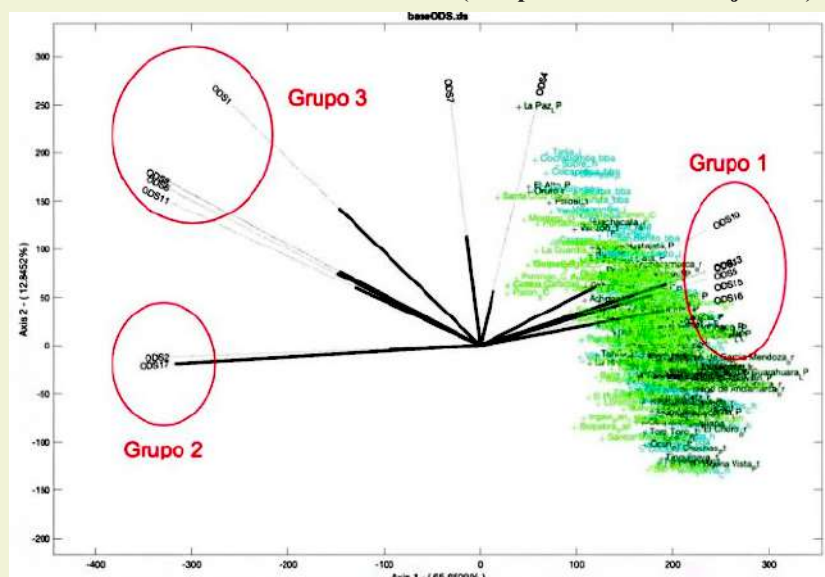
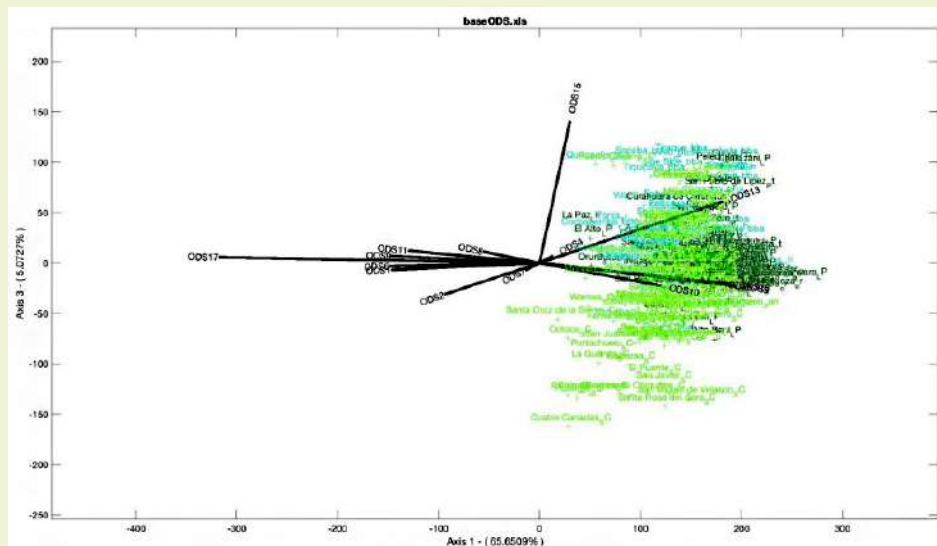


Gráfico 3 Representación factorial resultante del HJ- Biplot, Plano 1-3



En el caso del plano 1-3, el eje 3 tiene una absorción de la inercia parcial que alcanza 5.03%, ambos ejes alcanzan 70% como se puede apreciar en el Gráfico 3, solo el ODS15 (vida de ecosistemas terrestres) parece ser independiente de los otros objetivos de desarrollo sostenible.

3.2. Clúster según tipo de cumplimiento a los ODS

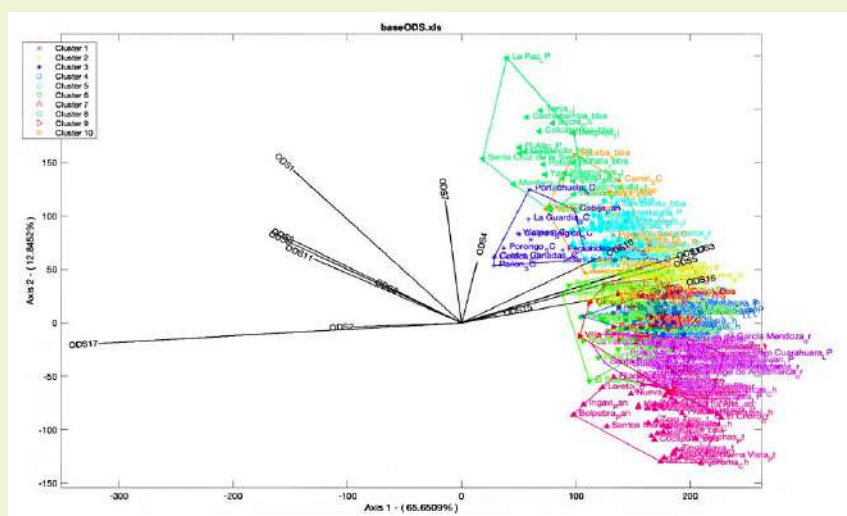
A través de las coordenadas Biplot se han calculado los Clusters (método K-means, distancia euclídea), en este caso se solicitaron el cálculo de 10 clusters. Se observa en el gráfico 4 la representación factorial donde los centros forman conglomerados en función de su comportamiento en las variables (objetivos) ODS3, ODS5, ODS10, ODS13 y ODS16 (Salud y Bienestar, Igualdad y Género, Reducción de Las Desigualdades, Acción por El Clima y Vida de Ecosistemas Terrestres, Respectivamente). En el [Anexo 1](#) Tabla Municipios y sus clústeres para método k-means clústeres para biplot scorese, se pueden observar la composición de cada uno de los clústeres, se puede destacar inmediatamente que para

las ciudades más grandes en población se encuentran en el clúster 10 agrupadas en función al comportamiento de las variables ODS4 (Educación de Calidad), debido a sus características generales en términos de similitud del cumplimiento de los objetivos de desarrollo sostenible. Se puede destacar el clúster 7 con 47 municipios y la agrupación principalmente depende de la variable ODS10 (Reducción de las Desigualdades). En general se puede observar que los clústeres tienen un bajo cumplimiento de los ODS1, ODS2, ODS6, ODS8, ODS9, ODS11 y el ODS17 (Fin de La Pobreza, Hambre Cero, Agua Limpia y Saneamiento, Trabajo Decente y Crecimiento Económico, Ciudades y Comunicaciones Sostenibles y Alianzas para Lograr Objetivos, respectivamente). La mayoría de los grupos se caracteriza por un buen cumplimiento en la variable ODS16, debido a que en los últimos años se han consolidado las formas de articulación de las instituciones y organizaciones sociales principalmente territoriales en los municipios más rurales. También se destaca el aporte de la variable ODS5 (Equidad de Género) en las agrupaciones, esto debido a las diferentes políticas enfocadas a la búsqueda de igual y

equidad de género como a Ley Integral para Garantizar a las Mujeres una Vida Libre de Violencia (Ley 348), la participación de la mujer en cargos de liderazgo en las estructuras de poder político principalmente, los movimientos sociales y organizaciones

de mujeres que han desempeñado un papel muy importante en la pasada década y otros factores donde las mujeres se destacaron cambiando y sensibilizando a la sociedad en general.

Gráfico 4 Representación factorial del HJ Biplot por clúster, plano 1-2



Las calidades de representación para cada conglomerado en el plano 1-2 se exponen en la tabla 3 (medida que proporciona información sobre la precisión con la que las variables y las observaciones están representadas en el biplot), en este caso es alta en la mayoría de los clústeres para el plano 1-2, con calidad media para los clústeres 4, 5 y 9. Estas medidas se calculan como parte del proceso

de construcción del biplot en el análisis HJ-Biplot, después de calcular los ejes principales y las coordenadas de las variables y observaciones en el espacio biplot, se pueden calcular las medidas de calidad de representación para evaluar la relevancia y la contribución de cada variable y observación al análisis.

Tabla 3 Conglomerados y calidades de información de representación plano 1-2

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Clúster 1	35	177,66	-33.929	90.015	32.832	93.298
Clúster 2	21	165,97	-100,75	70.639	26,03	96.668
Clúster 3	37	176,76	-65,27	86.436	11.786	98.222
Clúster 4	22	126,47	78.826	68.223	0,26502	68.488
Clúster 5	30	157,44	38.693	78.988	0,047707	79.036
Clúster 6	62	156,63	27.857	96.125	30.406	99.166
Clúster 7	47	129,55	79.455	71.935	27,06	98.994
Clúster 8	46	168,6	-13,99	97.262	0,66971	97.931
Clúster 9	16	68.085	81,59	22.381	32.141	54.522
Clúster 10	23	74.839	152,04	18.996	78.405	97.401

El clúster 6 es el que tiene a mayor cantidad de municipios (62), también destaca el clúster 9 con 16 municipios (el clúster más pequeño) que se caracteriza por municipios de tamaño medio, la mayoría ubicados en el departamento de Santa Cruz y la ciudad de Cobija, capital del departamento de Pando.

En segundo lugar, se encuentran los clústeres 2 y 4 con 21 y 22 municipios respectivamente, en este sentido, si se quiere diseñar políticas a nivel nacional de acuerdo al cumplimiento de los objetivos de desarrollo, los clústeres del 1 al 10 muestran aquellos que tienen menos disimilaridad de acuerdo a un análisis multidimensional, tomando en cuenta la reducción del espacio de 15 dimensiones a 2, perdiendo muy poca información.

3.3. Resultados a nivel departamental

El objetivo del presente documento no solo es mostrar una herramienta de análisis estadístico multidimensional que ayude a identificar municipios que tengan una similaridad en términos de cumplimiento de los 15 objetivos medidos en el Atlas, en ese sentido, además se puede observar adicionalmente la relación entre cada una de las columnas o variables entre sí y los municipios, pero con la idea de brindar adicionalmente una herramienta que nos ayude a plantear políticas de desarrollo local focalizadas, podemos estratificar los resultados a nivel departamental, ya que la información así nos lo permite.

De esta manera, se presenta el HJ-Biplot a nivel departamental y su clusterización. Una de las principales ventajas de este método consiste en determinar cuáles son las variables (ODS y nivel de cumplimiento) responsables de la configuración de los diferentes clusteres. De esa forma, se quiere determinar las diferentes configuraciones que los clústeres tienen en los diferentes planos factoriales, analizando sus contribuciones relativas del

factor al elemento de las variables (Ver [Anexo 2](#) Clusteres por departamento).

En este sentido se tiene en los gráficos a continuación la aplicación de la metodología HJ-Biplot y el proceso de clusterización por departamento. Se destaca en todos los casos la existencia de diferencias importantes en el cumplimiento de los diferentes ODS's de acuerdo al tamaño de los municipios en términos de población, aquellos municipios grandes como las ciudades capitales y ciudades intermedias grandes siempre están agrupadas en un clúster, a continuación municipios intermedios en otro clúster y finalmente un grupo con los municipios más alejados, pequeños y pobres. Para el caso del departamento de La Paz, se tiene los municipios de La Paz y El Alto, como las ciudades más pobladas con cumplimiento principalmente de los ODS4 (Educación de Calidad), ODS7 (Energía Asequible y No Contaminante) y ODS8 (Trabajo Decente y Crecimiento Económico) entre otras.

Para todos los departamentos la retención al segundo eje mínimo es de 77.958% de inercia acumulada, más que suficiente para caracterizar, con garantías, los alcances a los objetivos de desarrollo sostenible y sus respectivos municipios, (las tablas de los valores propios y varianza explicada se pueden ver en el [Anexo 3](#)).

Las calidades de representación para cada conglomerado en el plano 1-2 se exponen en el [Anexo 4](#), para cada departamento, donde se destaca la calidad en todos los clústeres por encima del 95%, para el departamento de La Paz por encima del 92.456, Cochabamba tiene más del 83.172%, el departamento de Oruro una calidad de representación mínima de 94.432% para el clúster 4, Potosí para su segundo clúster una calidad mínima de 96.186, Tarija una calidad de representación mínima de 81.855, el departamento de Santa

Cruz en su cuarto clúster tiene una calidad de presentación de 88.0471 siendo esta la mínima de los clústeres contruidos, el Beni con una calidad mínima de representación de 89.254 y el departamento de Pando con calidad mínima de 74.467, para su primer clúster siendo este el menor de todos los departamentos.

4. DISCUSIÓN

Los resultados de la revisión bibliográfica realizada ponen de manifiesto que los métodos Biplot han logrado un creciente desarrollo desde su formulación en 1971, en el que Ruben Gabriel introduce el GH y el JK Biplot; el MANOVA Biplot de una vía en 1972, el MANOVA Biplot para tablas de 2 vías, en 2004; HJ Biplot en 1986, Biplot Canónico en 1992, Modelos AMMI GGE Biplot en 2001, Biplot para datos composicionales, en 2002, MANOVA Biplot para tablas de 3 vías en 2004, Biplot Logístico externo en 2008, el HJ Biplot inferencial, en 2015, el Biplot Logístico Nominal en 2016, Biplot Logístico Ordinal en 2016, y el HJ Biplot Composicional en 2016.

Este artículo se ha analizado la distribución de los municipios en función a la implementación y alcance de los Objetivos de Desarrollo ODS Sostenible medidos y expuestos en el Atlas de los ODS 2020 mediante la técnica HJ-Biplot que permite realizar una representación gráfica de datos multivariantes en el que filas y columnas pueden ser superpuestas en un mismo sistema de referencia con máxima calidad de representación.

Así, el estudio permite observar que las características analizadas son, en parte, dependientes de cada departamento y/o de cada municipio ya que estos tienden a posicionarse en la misma zona de los gráficos factoriales resultantes. Sin embargo, existe también cierta heterogeneidad intra-

municipio, de forma que los municipios grandes que tienen ciudades capitales caracterizados por grandes poblaciones respecto al departamento muestran la mayor homogeneidad, mientras que los municipios pequeños presentan mayor variedad de comportamientos entre sí mismos de un departamento a otro.

El análisis de los datos a nivel multidimensional ha puesto de manifiesto que los municipios, se pueden agrupar de manera que se tomen en cuenta sus variables más débiles, así como las más fuertes para identificar disimilaridades entre ellas y así establecer políticas públicas que generen una especie de economías de escala, que reduzcan costos eliminando la duplicidad de actividades y así ser más eficientes.

El artículo permite analizar el comportamiento específico de cada municipio, y su situación en el contexto de su cumplimiento a los ODS a nivel nacional y departamental. Así, es posible identificar municipios que se salen del patrón general de departamento, y determinar qué faceta de su cumplimiento les hace singulares.

5. CONCLUSIONES

En conclusión, el HJ-Biplot se ha revelado como una herramienta de análisis multivariante sumamente útil en el análisis de datos económico sociales, en este caso cumplimientos a los ODS en la etapa descriptiva. La técnica, ha resultado muy buena al ser aplicable a cualquier matriz de datos y la posibilidad de detectar qué indicadores sociales son los responsables de las agrupaciones de los municipios.

La aplicación del HJ-Biplot al estudio de los ODS nos ha permitido caracterizar la actividad de los municipios en cuanto a su comportamiento sobresaliente o singular, que

los diferencia o no del resto de municipios en términos de cumplimiento de sus ODS's a nivel nacional y departamental.

Finalmente, es importante destacar, que a partir, de la implementación de la técnica HJ-Biplot se obtuvieron un conjunto de conglomerados a nivel nacional (Ver [Anexo 1](#)) y departamental (Ver [Anexo 2](#)), destacan a nivel nacional el cumplimiento distinto que tienen los municipios más grandes respecto de los pequeños, en el caso de los valores más altos están las ciudades como La Paz, El Alto, Cochabamba, Santa Cruz, Montero, etc. se encuentran en un clúster (Clúster 10), por otro lado, están los municipios con valores bajos de cumplimiento, en otro clúster (Clúster 2) como Poroma en Chuquisaca, San Pedro de Buena Vista en Pando, Tinguipaya

en Potosí, entre otros, que concuerda con los municipios de más bajo Índice Municipal de Desarrollo Sostenible.

Toda el análisis multidimensional puede servir de base para la implementación de políticas que nos ayuden a mejorar y/o alcanzar objetivos de desarrollo con el fin de dar cumplimiento a nuestros compromisos internacionales y así alcanzar metas que nos lleven a tener la oportunidad de vivir una vida digna en función a nuestros propios esfuerzos.

6. CONFLICTO DE INTERESES

El autor declara que no hay conflicto de intereses con respecto a la publicación de este documento.

REFERENCIA BIBLIOGRÁFICA

- Aitchison, J. (1982). The Statistical Analysis of Compositional Data. Journal of the Royal Statistical Society: Series B (Methodological), 44(2), 139–160.
<https://doi.org/10.1111/j.2517-6161.1982.tb01195.x>
- Aitchison, J. (1997). The one-hour course in compositional data analysis or compositional data analysis is simple. In: V. Pawlowsky-Glahn (Ed.), Proceedings of IAMG'97 – The III Annual Conference of the International Association for Mathematical Geology. (Vols. I, II and addendum, p. 3-35). Barcelona: International Center for Numerical Methods in Engineering (CIMNE).
- Aitchison, J., & Greenacre, M. (2002). Biplots of compositional data. Journal of the Royal Statistical Society: Series C (Applied Statistics), 51(4), 375–392.
<https://doi.org/10.1111/1467-9876.00275>
- Aitchison, J., Barceló-Vidal, C., Egozcue, J. J., & Pawlowsky-Glahn, V. (2002). A concise guide for the algebraic-geometric structure of the simplex, the sample space for compositional data analysis. In Proceedings of Eighth Annual Conference of the International Association for Mathematical Geology. (Vol. 2, pp. 387-392).
- Alcaraz Lamana, A.; Alonso Torres, P. (2019). La contribución de las universidades a la Agenda 2030. Unitat de Cooperació, Servei de Relacions Internacionals i Cooperació Universitat de València
- Amaratunga, D., & Cabrera, J. (2016). High-dimensional data. Journal of the National Science Foundation of Sri Lanka, 44(1).
- Amaro, I. R. (2001). Manova biplot para diseños con varios factores, basado en modelos lineales generales multivariantes. [Tesis doctoral]. Universidad de Salamanca, España.

- Amaro, I. R., Vicente-Villardón, J. L. y, & Galindo-Villardón, P. (2004). MANOVA BIPLLOT para arreglos de tratamientos con dos factores basado en modelos lineales generales multivariantes. *Interciencia*, 29(1), 26–32.
- Baccalá, N. (2004). Contribuciones al Análisis de Matrices de Datos Multivía: tipología de las variables. [Tesis doctoral]. Universidad de Salamanca, España, Spain.
- Becker, H. C. (1981). Correlations among some statistical measures of phenotypic stability. *Euphytica*, 30(3), 835–840.
- Benzécri, J. P. (1973). *L'analyse des données* (Vol. 2). Paris: Dunod.
- Berman, J. J. (2013). Principles of big data: preparing, sharing, and analyzing complex information. Newnes.
- Billheimer, D., Guttorp, P., & Fagan, W. F. (2001). Statistical Interpretation of Species Composition. *Journal of the American Statistical Association*, 96(456), 1205–1214.
<https://doi.org/10.1198/016214501753381850>
- Blázquez, A. (1998). Análisis biplot basado en modelos lineales generalizados. [Tesis doctoral]. Universidad de Salamanca, España.
- Bodor, A., Csabai, I., Mahoney, M. W., & Solymosi, N. (2012). rCUR: an R package for CUR matrix decomposition. *BMC Bioinformatics*, 13(1), 103. <https://doi.org/10.1186/1471-2105-13-103>
- Braak, C. J. Ter, & Looman, C. W. (1994). Biplots in reduced-rank regression. *Biometrical Journal*, 36(8), 983–1003.
- Bradu, D., & Gabriel, K. R. (1974). Simultaneous statistical inference on interactions in two-way analysis of variance. *Journal of the American Statistical Association*, 69(346), 428–436.
- Bradu, D., & Gabriel, K. R. (1978). The biplot as a diagnostic tool for models of two-way tables. *Technometrics*, 20(1), 47–68.
- Breiman, L. (1995). Better Subset Regression Using the Nonnegative Garrote. *Technometrics*, 37(4), 373–384.
<https://doi.org/10.1080/00401706.1995.10484371>
- Cadima, J., & Jolliffe, I. T. (1995). Loading and correlations in the interpretation of principle compenents. *Journal of Applied Statistics*, 22(2), 203–214.
<https://doi.org/10.1080/757584614>
- Cárdenas, O. C., & Galindo, M. P. (2004). Biplot con información externa basado en modelos bilineales generalizados.
- Cárdenas, O., Noguera, C., Galindo, M. P., & Vicente-Villardón, J. L. (2003). El uso de información externa en aproximaciones Biplot. *Revista Venezolana de Análisis de Coyuntura*, 9(2), 257–276.
- Cárdenas, O., Noguera, C., Galindo, P., & Vicente-Villardón, J. L. (2006). An alternative to principal components regression based on Regression Biplot. *INTERCIENCIA*, 31(3), 160–167.
- Carlier, A., & Kroonenberg, P. M. (1996). Decompositions and biplots in threeway correspondence analysis. *Psychometrika*, 61(2), 355–373.
- Chessel, D., Dufour, A., & Thioulouse, J. (2004). The ade4 package-I-One-table methods. Pdfs.Semanticscholar.Org.

- Chessel, D., Dufour, A.B., Dray, S., Jombart, T., Lobry, J.R., Ollier, S. y, & Thioulouse, J. (2013). ade4. R package version 1.5-2: analysis of ecological data: exploratory and Euclidean methods in environmental sciences.cran.rproject.org/package=ade4, 2013.
- Choulakian, V. (1996). Generalized bilinear models. *Psychometrika*, 61(2), 271– 283.
- Cortés-Rodríguez, M., & Sánchez-Barba, M. (2013). Biplot de datos composicionales: una herramienta útil en el estudio de test psicológicos. Universidad de Salamanca, España.
- Crossa, J., Cornelius, P. L., & Yan, W. (2002). Biplots of linear-bilinear models for studying crossover genotype× environment interaction. *Crop Science*, 42(2), 619– 633.
- Crossa, J., Gauch, H. G., & Zobel, R. W. (1990). Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. *Crop Science*, 30, 493–500.
- Cubilla-Montilla, M., Nieto-Librero, A.-B., Galindo-Villardón, M. P., Vicente Galindo, M. P., & Garcia-Sanchez, I.-M. (2019). Are cultural values sufficient to improve stakeholder engagement human and labour rights issues? *Corporate Social Responsibility and Environmental Management*, 26(4), 938–955. <https://doi.org/10.1002/csr.1733>
- D'Aspremont, A., El Ghaoui, L., Jordan, M. I., & Lanckriet, G. R. G. (2007). A Direct Formulation for Sparse PCA Using Semidefinite Programming. *SIAM Review*, 49(3), 434–448. <https://doi.org/10.1137/050645506>
- De Falguerolles, A. (1996). Generalized bilinear models and generalized biplots: some examples. Publications Du Laboratoire de Statistique et Probabilités.
- Demey, J. R. (2008). Diversidad genética en bancos de Germoplasma: un enfoque Biplot.
- Denis, J. B. (1991). Ajustements de modèles linéaires et bilinéaires sous contraintes linéaires avec données manquantes. *Revue de Statistique Appliquée*, 39(2), 5–24.
- Donoho, D. L., & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3), 425–455. <https://doi.org/10.1093/biomet/81.3.425>
- Dray, S., Dufour, A.B., Chessel, D. (2007). The ade4 package-II: Two-table and Ktable methods. *R Journal*, 7(2), 47–52.
- Drineas, P., Kannan, R., & Mahoney, M. (2006). Fast Monte Carlo algorithms for matrices III: Computing a compressed approximate matrix decomposition. *SIAM, Journal on Computing*, 36(1), 184–206.
- Drineas, P., Mahoney, M. W., & Muthukrishnan, S. (2008). Relative-Error CUR Matrix Decompositions. *SIAM Journal on Matrix Analysis and Applications*, 30(2), 844–881. <https://doi.org/10.1137/07070471X>
- Eckart, C., & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3), 211–218.
- Eckart, C., & Young, G. (1939). A principal axis transformation for non-Hermitian matrices. *Bulletin of the American Mathematical Society*, 45(2), 118–121.
- Efron, B., & Tibshirani, R. (1994). An introduction to the bootstrap.

- Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician*, 37(1), 36–48.
- Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, 32(2), 407–499.173
- Efron, Bradley. (1979). Computers and the Theory of Statistics: Thinking the Unthinkable. *SIAM Review*, 21(4), 460–480. <https://doi.org/10.1137/1021092>
- Egido, J. (2014). dynBiplotGUI. R package versión 1.0.1. cran.rproject.org/web/packages/dynBiplotGUI.
- Erichson, N., Zheng, P., Manohar, K., Brunton, S., Kuetz, J., & Aravkin, A. (2018). Sparse Principal Component Analysis via Variable Projection. *ArXiv Preprint ArXiv:1804.00341*.
- Fan, J., & Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the American Statistical Association*, 96(456), 1348–1360. <https://doi.org/10.1198/016214501753382273>
- Farcomeni, A. (2009). An exact approach to sparse principal component analysis. *Computational Statistics*, 24(4), 583–604. <https://doi.org/10.1007/s00180-008-0147-3>
- Faria, J. C., & Demetrio, C. G. B. (2012). Biplot of multivariate data based on principal components analysis. R package version 1.02. URL <http://cran.rproject.org/package=bpca>.
- Fernández-Gómez, M. J. (1995). Contribuciones al análisis multivariante directo del gradiente mediante estudio combinado de configuraciones espaciales. [Tesis doctoral]. Universidad de Salamanca, España.
- Fisher, R. A. (1936). The use of mutiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2), 179–188. <https://doi.org/10.1111/j.1469-1809.1936.tb02137>
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1.
- Frieze, A., Kannan, R., & Vempala, S. (2004). Fast Monte-Carlo algorithms for finding low-rank approximations. *Journal of the ACM (JACM)*, 51(6), 1025–1041.
- Frutos, E., & Galindo, M. P. (2013). GGEbiplotGUI. R package version 1.0-6: interactive GGE biplots in R. cran.r-project.org/package=GGEbiplotGUI,.
- Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58(3), 453–467.
- Gabriel, K. R. (1972). Analysis of Meteorological Data by Means of Canonical Decomposition and Biplots. *Journal of Applied Meteorology*, 11(7), 1071–1077. [https://doi.org/10.1175/1520-0450\(1972\)011<1071:AOMDBM>2.0.CO;2](https://doi.org/10.1175/1520-0450(1972)011<1071:AOMDBM>2.0.CO;2)
- Gabriel, K. R. (1995). “MANOVA biplots for two-way contingency tables.” In: W.J. Krzanowski (Ed.), *Recent Advances in Descriptive Multivariate Analysis*, Oxford: Clarendon Press.
- Gabriel, K. R. (1998). Generalised bilinear regression. *Biometrika*, 85(3), 689–700.

- Gabriel, K. R., & Odoroff, C. L. (1990). Biplots in biomedical research. *Statistics in Medicine*, 9(5), 469–485. <https://doi.org/10.1002/sim.4780090502174>
- Galindo, M. P. (1986). Una alternativa de representacion simultanea: HJ-Biplot. *Qüestiió: Quaderns d'estadística i Investigación Operativa*, 10(1), 13–23.
- Galindo, M. P., & Cuadras, C. M. (1986). Una extensión del método Biplot y su relación con otras técnicas. *Publicaciones de Bioestadística y Biomatemática*, 17.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144.
- Gauch, H. G. (1988). Model selection and validation for yield trials with interaction. *Biometrics*, 705–715.
- Gauch, H. G., & Zobel, R. W. (1989). Accuracy and selection success in yield trial analyses. *Theoretical and Applied Genetics*, 77(4), 473–481. <https://doi.org/10.1007/BF00274266>
- Gollob, H. F. (1968). A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika*, 33(1), 73–115.
- Goreinov, S., Tyrtysnikov, E. E., & Zamarashkin, N. L. (1997). A theory of pseudoskeleton approximations. *Linear Algebra and Its Applications*, 261(1–3), 1–21.
- Goreinov, S. A., & Tyrtysnikov, E. E. (2001). The maximal-volume concept in approximation by low-rank matrices. *Contemporary Mathematics*, 280, 47–52.
- Gower, J. C. (1992). Generalized biplots. *Biometrika*, 79(3), 475–493.
- Gower, J. C., & Hand, D. J. (1995). *Biplots* (Vol. 54). CRC Press.
- Gower, J. C., & Hand, D. J. (1996). *Biplots*. Chapman&Hall, London UK.
- Gower, J. C., & Harding, S. A. (1988). Nonlinear biplots. *Biometrika*, 75(3), 445–455.
- Graffelman, J. (2012). Calibrate: calibration of scatterplot and biplot axes. R package version 1(1). URL cran.r-project.org/package=calibrate, 2012.
- Greenacre, M. J. (1984). *Correspondence analysis*. London: Academic Press.
- Greenacre, M. J. (1993). Biplots in correspondence analysis. *Journal of Applied Statistics*, 20(2), 251–269. <https://doi.org/10.1080/026647693000000021>
- Greenacre, M., & Nenadic, O. (2012). The ca R package version 0.53: simple, multiple and joint correspondence analysis. cran.r-project.org/package=ca.
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47, 98–115.
- Hausman, R. E. (1982). Constrained multivariate analysis. In: Zanakis SH, Rustagi JS (eds). *Optimisation in Statistics*. North-Holland, Amsterdam, 137–151.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Selección de la muestra en Metodología de la Investigación*. México: McGrawHill. 175
- Hernández Sánchez, J. C. (2016). Biplot logístico para datos nominales y

- ordinales. [Tesis Doctoral]. Universidad de Salamanca, España.
- Hernández Suárez, M., Molina Pérez, D., Rodríguez-Rodríguez, E., Díaz Romero, C., Espinosa Borreguero, F., & Galindo-Villardón, P. (2016). The Compositional HJ-Biplot—A New Approach to Identifying the Links among Bioactive Compounds of Tomatoes. *International Journal of Molecular Sciences*, 17(11), 1828.
- Hernández, J. C., & Vicente-Villardón, J. L. (2013a). Nominal Logistic Biplot. R package version 0.1: biplot representations of categorical data. cran.rproject.org/web/packages/NominalLogisticBiplot/index.html.
- Hernández, J. C., & Vicente-Villardón, J. L. (2013b). Ordinal Logistic Biplot. R package version 0.2: ordinal logistic biplots. cran.rproject.org/web/packages/OrdinalLogisticBiplot/index.html.
- Hernández, S. (2005). Biplots robustos. [Tesis doctoral]. Universidad de Salamanca, España.
- Hernández, S., & Galindo-Villardón, M. P. (2006). BIPROB: UN MÉTODO PARA OBTENER UN BILOT ROBUSTO. *Investigación Operacional*, 27(3), 287–299.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1), 55–67. <https://doi.org/10.1080/00401706.1970.10488634>
- Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Readings in Psychology and Culture*, 2(1), 8.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417.
- Jolliffe, I. T. (1995). Rotation of principal components: choice of normalization constraints. *Journal of Applied Statistics*, 22(1), 29–35. <https://doi.org/10.1080/757584395>
- Jolliffe, I. T., & Uddin, M. (2000). The Simplified Component Technique: An Alternative to Rotated Principal Components. *Journal of Computational and Graphical Statistics*, 9(4), 689–710. <https://doi.org/10.1080/10618600.2000.10474908>
- Jolliffe, I. T., Trendafilov, N. T., & Uddin, M. (2003). A Modified Principal Component Technique Based on the LASSO. *Journal of Computational and Graphical Statistics*, 12(3), 531–547. <https://doi.org/10.1198/1061860032148>
- Kempton, R. A. (1984). The use of biplots in interpreting variety by environment interactions. *The Journal of Agricultural Science*, 103(1), 123–135.
- La Grange, A. M., Le Roux, N. J., Rousseeuw, I. R., & Tukey, J. W. (2009). BiplotGUI: interactive biplots in R. R package version 0.0-7. cran.rproject.org/package=BiplotGUI.
- Lafortune, G., Fuller, G., Moreno, J., Schmidt-Traub, G., Kroll, C. (2018). SDG Index and Dashboards: Detailed Methodological paper. Paris: Sustainable Development Solutions Network (SDSN).
- Mahoney, M. W., & Drineas, P. (2009). CUR matrix decompositions for improved data analysis. *Proceedings of the National Academy of Sciences of the United States of America*, 106(3), 697–702. <https://doi.org/10.1073/pnas.0805349106>

[org/10.1073/pnas.0803205106](https://doi.org/10.1073/pnas.0803205106)

- Mahoney, M. W., Maggioni, M., & Drineas, P. (2008). Tensor-CUR Decompositions for Tensor-Based Data. *SIAM Journal on Matrix Analysis and Applications*, 30(3), 957–987. <https://doi.org/10.1137/060665336>
- Mandel, J. (1971). A New Analysis of Variance Model for Non-additive Data. *Technometrics*, 13(1), 1–18. <https://doi.org/10.1080/00401706.1971.10488751>
- Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate Analysis*. Press Inc. London.
- Markos, A. (2012). Package caGUI. The GUI ca R package version 0.1-4: a Tcl/Tk GUI for the functions. cran.r-project.org/package=caGUI, 2012.
- Martín-Rodríguez, J. (1996). Contribuciones a la integración de subespacios desde una perspectiva biplot. [Tesis doctoral]. Universidad de Salamanca, España.
- McCabe, G. P. (1984). Principal Variables. *Technometrics*, 26(2), 137–144. <https://doi.org/10.1080/00401706.1984.10487939>
- Moghaddam, B., Weiss, Y., & Avidan, S. (2006). Spectral bounds for sparse PCA: Exact and greedy algorithms. *Advances in Neural Information Processing Systems*.
- Nenadic, O., & Greenacre, M. (2007). Correspondence analysis in R, with two- and three-dimensional graphics: The ca package. *Journal of Statistical Software*, 20(3), 1–13.
- Nieto-Librero, A. B. (2015). Versión inferencial de los métodos Biplot basada en remuestreo Bootstrap y su aplicación a tablas de tres vías. [Tesis doctoral]. Universidad de Salamanca, España.
- Nieto-Librero, A. B., & Galindo-Villardón, P. (2015). biplotbootGUI. R package version 1.0: Bootstrap on Classical Biplots and Clustering Disjoint Biplot. [cran.rproject.org/web/packages/biplotbootGUI/](https://cran.r-project.org/web/packages/biplotbootGUI/).
- Nieto-Librero, A. B., Baccalá, N., & Galindo, M. P. (2012). MultibiplotGUI. R package version 0.0-1: Multibiplot Analysis. [cran.rproject.org/package=multibiplotGUI](https://cran.r-project.org/package=multibiplotGUI).
- Nieto-Librero, A. B., Sierra, C., Vicente-Galindo, M. P., Ruíz-Barzola, O., & Galindo-Villardón, M. P. (2017). Clustering Disjoint HJ-Biplot: A new tool for identifying pollution patterns in geochemical studies. *Chemosphere*, 176, 389–396.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Minchin, P. R., ... Wagner, H. (2013). Package vegan. Community Ecology Package, Version 2.9. R package version 2.0-8. cran.r-project.org/package=vegan.177
- Pawlowsky-Glahn, V., & Egozcue, J. J.-S. (2001). Geometric approach to statistical analysis on the simplex. *Stochastic Environmental Research and Risk Assessment*, 15(5), 384–398.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559–572.
- Qi, X., Luo, R., & Zhao, H. (2013). Sparse principal component analysis by choice of norm. *Journal of Multivariate Analysis*, 114:127-150.

- <https://doi.org/10.1016/j.jmva.2012.07.004>
- Quenouille, M. H. (1950). An application of least squares to family diet surveys. *Econometrica: Journal of the Econometric Society*, 27–44.
- Resolución 70/1 [Asamblea General de las Naciones Unidas]. Por la cual se aprobó la Agenda 2030 para el Desarrollo Sostenible. 25 de septiembre de 2015.
- Rodríguez-Mazahua, L., Rodríguez-Enríquez, C.-A., Sánchez-Cervantes, J. L., Cervantes, J., García-Alcaraz, J. L., & Alor-Hernández, G. (2016). A general perspective of Big Data: applications, tools, challenges and trends. *The Journal of Supercomputing*, 72(8), 3073–3113.
- R-TEAM. (2014). A language and environment for statistical computing. R foundation for statistical computing. Vienna, Austria, 2014. Retrieved from www.r-project.org.
- Shen, H., & Huang, J. (2008). Sparse principal component analysis via regularized low rank matrix approximation. *Journal of Multivariate Analysis*, 99(6), 1015–1034.
- Stewart, G. W. (1999). Four algorithms for the efficient computation of truncated pivoted QR approximations to a sparse matrix. *Numerische Mathematik*, 83(2), 313–323. <https://doi.org/10.1007/s002110050451>
- Talia, D. (2013). Clouds for scalable big data analytics. *Computer*, 46(5), 98–101.
- Ter Braak, C. J. (1986). Canonical Correspondence Analysis: a new eigenvector technique for Multivariate Direct Gradient Analysis. *Ecology*, 67(5), 1167–1179.
- Ter Braak, C. J. (1990). Interpreting canonical correlation analysis through biplots of structure correlations and weights. *Psychometrika*, 55(3), 519–531.
- Thioulouse, J., & Dray, S. (2007). Interactive multivariate data analysis in R with the ade4 and ade4TkGUI packages. *Journal of Statistical Software*, 22(5), 1–14.
- Thioulouse, J., & Dray, S. (2012). ade4TkGUI. R package version 0.2-6: ade4 Tcl/Tk graphical user interface. cran.r-project.org/package=ade4TkGUI, 2012.
- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(3), 273–282.178
- Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, 17(4), 401–419. <https://doi.org/10.1007/BF02288916>
- Trendafilov, N. T. (2014). From simple structure to sparse components: A review. *Computational Statistics*, 29(3–4), 431–454.
- Tukey, J. (1958). Bias and confidence in not quite large samples. *Ann. Math. Statis*, 29, 614.
- Vairinhos, V. M. (2003). Desarrollo de un sistema para minería de datos basado en los métodos Biplot.
- Van Eeuwijk, F. A. (1995). Multiplicative interaction in generalized linear models. *Biometrics*, 1017–1032.

- Vicente-Tavera, S. (1992). Las técnicas de representación de datos multidimensionales en el estudio del Índice de Producción Industrial en la CEE (Tesis Doctoral). Universidad de Salamanca, España.
- Vicente-Villardón, J. L. (1992). Una alternativa a las técnicas factoriales clásicas basada en una generalización de los métodos BILOT. [Tesis doctoral]. Universidad de Salamanca, España.
- Vicente-Villardón, J. L. (2001). Biplot for binary data based on logistic response surfaces. In Salamanca Statistics Seminar IV: Advances in Multivariate Analysis. Salamanca, Spain.
- Vicente-Villardón, J. L. (2010). MULTBILOT: package for multivariate análisis using biplots. Departamento de Estadística. Universidad de Salamanca. Retrieved from <http://biplot.usal.es/multibiplot/introduction.html>
- Vicente-Villardón, J. L. (2017). MultiBiplotR: Multivariate Analysis using Biplot. R package version 0.1.0. <http://biplot.usal.es/classicalbiplot/multibiplot-in-r/>.
- Vicente-Villardón, J. L., Galindo-Villardón, M. P., & Blázquez-Zavallos, A. (2004). Constrained Logistic Biplots. In SALAMANCA STATISTICS SEMINAR V. Advances in Descriptive Multivariate Analysis. Universidad de Salamanca, España.
- Vicente-Villardón, J. L., Galindo-Villardón, M. P., & Blázquez-Zavallos, A. (2006). Logistic biplots. Multiple Correspondence Analysis and Related Methods. London: Chapman & Hall. 503-521.
- Vichi, M., & Saporta, G. (2009). Clustering and disjoint principal component analysis. Computational Statistics & Data Analysis, 53(8), 3194–3208.
- Vines, S. K. (2000). Simple principal components. Journal of the Royal Statistical Society: Series C (Applied Statistics), 49(4), 441–451. <https://doi.org/10.1111/1467-9876.00204>
- Witten, D. M., Tibshirani, R., & Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. Biostatistics, 10(3), 515–534.
- Yan, W. (2001). GGEbiplot—a Windows application for graphical analysis of multi-environment trial data and other types of two-way data. Agronomy, 93(5), 1111–1118.
- Yan, W. (2002). Singular-value partitioning in biplot analysis of multi-environment trial data. Agronomy Journal, 94(5), 990–996.
- Yan, W., & Hunt, L. A. (2002). Biplot analysis of diallel data. Crop Science, 42(1), 21–30.
- Yan, W., & Kang, M. S. (2002). GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC Press.
- Yan, W., & Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: Principles and applications. Canadian Journal of Plant Science, 86(3), 623–645.
- Yan, W., Cornelius, P. L., Crossa, J., & Hunt, L. A. (2001). Two types of GGE biplots for analyzing multi-environment trial data. Crop Science, 41, 656–663.
- Yan, W., Hunt, L. A., Sheng, Q., & Szlavnick, Z. (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Science, 40, 597–605.

Yang, J., Rübel, O., Prabhat, Mahoney, M. W., & Bowen, B. P. (2015). Identifying Important Ions and Positions in Mass Spectrometry Imaging Data Using CUR Matrix Decompositions. *Analytical Chemistry*, 87(9), 4658–4666.

<https://doi.org/10.1021/ac5040264>

Young, G., & Householder, A. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika*, 3(1), 19–22.

Zhang, Z., Xu, Y., Yang, J., Li, X., & Zhang, D. (2015). A survey of sparse representation: algorithms and applications. *IEEE*, 490–530.

Zobel, R. W., Wright, M. J., & Gauch, H. G. (1988). Statistical analysis of a yield trial. *Agronomy Journal*, 80, 388–393.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*., 67(2), 301–320.

Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse Principal Component Analysis. *Journal of Computational and Graphical Statistics*, 15(2), 265–286. <https://doi.org/10.1198/106186006X113430>.

HJ-Biplot como metodología exploratoria para el análisis multidimensional de los objetivos de desarrollo sostenible (ODS) a nivel municipios en Bolivia

Anexo 1

Tabla Municipios y sus clústeres para método k-means clústers para biplot scores

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10
Belén de Andamarca	Alalay	Acasio	Alto Beni	Aiquile	Antequera	Achacachi	Ancoarimes	Cobija	Bernejo
Calacoto	Arque	Anzaldo	Ascension de Guarayos	Apolo	Arbieto	Achocalla	Ayo Ayo	Colpa Belgica	Carapari
Caquilaviri	Azurduy	Arampampa	Bella Flor	Charazani	Batalias	Arani	Belanzos	Cotoca	Cilza
Charaña	Bolpebra	Aucapata	Cabezas	Comarapa	Baures	Atocha	Cairama	Cuatro Cañadas	Cochabamba
Chipaya	Ckochas	Ayeta	Caranavi	Cuevo	Caiza D	Boyulibe	Chaqui	El Tono	Coleapirhua
Coro Coro	EL Sena	Bolivar	Coloni	Curahuara de Carangas	Cajuta	Buena Vista	Charagua	Fernández Alonso	El Alto
Corque	Ingavi	Callapa	El Puente	Curva	Calamarca	Camiri	Chayanta	Guayaramerín	Huachacalla
Cruz de Machacamarca	Loreto	Carangas	Machareti	Entre Rios	Camargo	Capinota	Chuma	La Guardia	La Paz
Culpina	Nueva Esperanza	Caripuyo	Porvenir	Eucaliptus	Caracollo	Chulumani	Coipasa	Mineros	Montero
El Villar	Oauri	Chacarilla	San Andrés	Ixiamas	Carmen Rivero Torrez	Copacabana	Colquiri	Okinawa Uno	Oruro
Escara	Poroma	Choque Cola	San Antonio de Lomerio	Magdalena	Catacora	Coripata	Comanche	Pailón	Potosí
Exaltación	Ravelo	Chuquihuta	San Javier	Moro Moro	Challapata	Corico	Combaya	Porongo	Puerto Quijarro
Guany	S.P. De Buena Vista	Cocapata	San Juan de Yapacani	Morochata	Chimoré	Desaguadero	Cotagaita	Portachuelo	Punata
Huacaraje	San Lorenzo	Colquechaca	San Julián	Muyupampa	Chua Cocani	El Puente	Cuchumuela	San José de Chiquitos	Quillacollo
Huacareta	San Pedro	El Choro	San Miguel de Velasco	Padilla	Colcha K	Entre Rios	Escoma	Warnes	Sacaba
Huacaya	Santos Mercado	Filadelfia	San Rafael	Pasorapa	Collana	Gral. Saavedra	Idia	Yacuiba	Santa Cruz de la Sierra
Huayllamarca	Tacobamba	Gutiérrez	San Ramón	Pelechuco	Colquencha	Huanuni	Llica	Sucre	
Inquisivi	Tarvita	Humanata	Santa Rosa del Sara	Pojo	Concepción	Huatajata	Malla	Tarija	
La Asunta	Tingupipaya	Ichoca	Teoponte	Pucara	Esmeralda	Llallagua	Mocomoco	Tiquipaya	
La Rivera	Vila Vila	Inchhuasi	Yapacani	Puerto Rico	Guaqui	Machacamarca	Mojocoya	Trinidad	
Pampa Aullagas	Villa Nueva - Loma Alta	Independencia		Reyes	Huari	Mairana	Paica	Vallegrande	
Puerto Gonzales Moreno		Mizque		San Matías	Huarina	Mecapaca	Papel Pampa	Villamontes	
Puerto Siles		Mojinete		San Pablo de Lipez	Irupana	Monteagudo	Pto. Carabuco	Villazón	
Sabaya		Pocoata		Santa Rosa	Jesús de Machaca	Padcaya	Pucarani		
Salinas de García Mendoza		Presto		Shinashota	Lagunillas	Pampa Grande	Puerto Acosta		
San Andrés de Machaca		Sacabamba		Tiraque	Laja	Palacamayaya	Puerto Pérez		
San Antonio de Esmoruco		San Lucas		Totorá	Las Carreras	Porco	Quiabaya		
San Pedro Cuarahua		Santa Rosa		Unibichá	Luribay	Puerto Suarez	San Agustín		
San Pedro de Totora		Sicaya		Villa Tunari	Mupiri	Roboré	San Borja		
Santiago de Andamarca		Sopachuy		Yunchará	Nazacara de Pacajes	Rurrenabaque	San Ignacio		
Tahua		Tapacaya			Omoreque	Salpina	Sapahaqui		
Todos Santos		Tapacarí			Palos Blancos	Samatpata	Soracachi		
Toledo		Tarabuco			Pazña	San Benito	Sorata		
Waldo Ballivian		Toro Toro			Pocona	San Carlos	Tacacoma		

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 9	Cluster 10
Yunguyo de Litoral		Umiri Villa Charcas Villa de Sacaca			Popó Postrer Valle Puerto Villaruel Puna Quillacas Quime Quirusillas Riberalta San Buenaventura San Ignacio de Velasco San Joaquín San Pedro de Quemes San Pedro de Tiquina Santa Ana de Yacuma Santiago de Huata Santiago de Machaca Santivañez Sica Sica Tacachi Taraco Tiahuanacu Tipuani Toko Uncía Villa Libertad Licoma Villa Rivero Villa Serrano Yocalla	San Pedro Sipe Sipe Tarata Tolata Trigal Tupiza Uriondo Uyuni Viacha Villa San Lorenzo Vinto Yanacachi Yotala	Tito Yupanqui Tomave Tomina Turco Urnala Vacas Villa Abecia Villa Alcalá Vifichi Yaco Yamparáez Zudafiez		

Anexo 2

Cochamba

La Paz

Chuquisaca

Tarija

Potosí

Oruro

Pando

Beni

Santa Cruz

Anexo 3

Valores propios y varianza explicada por Departamento

Chuquisaca

Axis	Eigenvalue	Inertia	
		Expl. Var.	Cummulative
Axis 1	185137.595	79.182	79.182
Axis 2	25788.823	11.03	90.212
Axis 3	5993.331	2.563	92.775
Axis 4	5325.475	2.278	95.053
Axis 5	2987.124	1.278	96.33

La Paz

Axis	Eigenvalue	Inertia	
		Expl. Var.	Cummulative
Axis 1	520420.521	78.583	78.583
Axis 2	51861.419	7.831	86.414
Axis 3	17672.898	2.669	89.082
Axis 4	16212.455	2.448	91.53
Axis 5	11903.42	1.797	93.328

Cochabamba

Axis	Eigenvalue	Inertia	
		Expl. Var.	Cummulative
Axis 1	224701.125	64.439	64.439
Axis 2	72237.96	20.716	85.155
Axis 3	23433.066	6.72	91.875
Axis 4	6865.459	1.969	93.844
Axis 5	5659.469	1.623	95.467

Oruro

Axis	Eigenvalue	Inertia	
		Expl. Var.	Cummulative
Axis 1	239748.278	80.181	80.181
Axis 2	23740.475	7.94	88.121
Axis 3	10004.003	3.346	91.467
Axis 4	5177.716	1.732	93.198
Axis 5	3988.51	1.334	94.532

Potosí

Axis	Eigenvalue	Inertia	
		Expl. Var.	Cummulative
Axis 1	219010.612	71.715	71.715
Axis 2	48402.875	15.85	87.565
Axis 3	13744.027	4.501	92.066
Axis 4	6121.438	2.004	94.07
Axis 5	5173.192	1.694	95.764

Tarija

Axis	Eigenvalue	Inertia	
		Expl. Var.	Cummulative
Axis 1	40343.617	70.745	70.745
Axis 2	9400.769	16.485	87.229
Axis 3	3349.97	5.874	93.104
Axis 4	1523.437	2.671	95.775
Axis 5	899.038	1.577	97.352

***HJ-Biplot* como metodología exploratoria para el análisis multidimensional de los objetivos de desarrollo sostenible (ODS) a nivel municipios en Bolivia**

Santa Cruz

Axis	Eigenvalue	Inertia Expl. Var.	Cummulative
Axis 1	210714.253	60.592	60.592
Axis 2	60392.787	17.366	77.958
Axis 3	28991.746	8.337	86.295
Axis 4	10722.097	3.083	89.378
Axis 5	9660.287	2.778	92.156

Beni

Axis	Eigenvalue	Inertia Expl. Var.	Cummulative
Axis 1	93729.284	72.551	72.551
Axis 2	12187.838	9.434	81.985
Axis 3	8914.587	6.9	88.886
Axis 4	4015.731	3.108	91.994
Axis 5	3244.435	2.511	94.505

Pando

Axis	Eigenvalue	Inertia Expl. Var.	Cummulative
Axis 1	82426.292	74.467	74.467
Axis 2	12861.201	11.619	86.086
Axis 3	6960.962	6.289	92.375
Axis 4	3135.552	2.833	95.207
Axis 5	1986.531	1.795	97.002

Anexo 4

Conglomerados y calidades de información de representación plano 1-2

Chuquisaca

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	9	955093	-18.219	95,48	0,03	95,5104
Cluster 2	12	1000555	-179340	94,90	3,05	97,9528
Cluster 3	8	786593	473099	72,88	26,36	99,2471

La Paz

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	20	121,6291	- 21,1521	93,452	2,826	96,279
Cluster 2	26	128,8094	- 20,5785	94,356	2,408	96,764
Cluster 3	22	122,7761	13,8238	97,558	1,237	98,795
Cluster 4	2	38,6991	148,3008	5,894	86,562	92,456
Cluster 5	17	100,7894	4,6493	81,450	17,332	98,781

Cochabamba

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	10	62,8338	85,9660	34,456	64,495	98,951
Cluster 2	11	105,4151	- 51,0837	79,884	18,759	98,644
Cluster 3	7	88,6169	- 62,3260	90,614	0,448	91,063
Cluster 4	8	103,2657	- 16,7116	81,049	2,123	83,172
Cluster 5	11	88,6041	29,5023	85,208	9,447	94,655

Oruro

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	10	91,0251	- 22,7754	91,726	5,743	97,468
Cluster 2	11	117,8279	13,7718	96,775	1,322	98,097
Cluster 3	4	56,0644	- 74,2944	35,055	61,558	96,612
Cluster 4	10	105,8824	15,4916	92,453	1,979	94,432

Potosí

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	14	109,828	40,952	87,028	12,100	99,127
Cluster 2	7	61,216	-75,777	37,983	58,203	96,186
Cluster 3	19	91,180	-20,038	93,264	4,504	97,768

Tarija

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	1	-53,095	24,997	67,004	14,851	81,855
Cluster 2	4	-58,290	-32,394	76,130	23,512	99,642
Cluster 3	6	-54,131	18,635	89,093	10,558	99,651

Santa Cruz

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	16	-84,1485	22,3513	83,9015	5,9195	89,821
Cluster 2	16	-91,4559	-28,583	85,5614	8,3574	93,9188
Cluster 3	12	-91,4236	49,6442	74,3296	21,9171	96,2467
Cluster 4	12	-64,4662	-53,8073	51,8944	36,1527	88,0471

Beni

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	6	80,942	-6,722	88,643	0,611	89,254
Cluster 2	6	67,094	-28,648	79,332	14,463	93,795
Cluster 3	7	72,965	27,480	84,292	11,956	96,249

Pando

CLUSTER	n	Coord-1	Coord-2	QLR - 1	QLR - 2	QLRA 1- 2
Cluster 1	2	41,687	-69,812	25,702	72,082	97,784
Cluster 2	3	71,356	-4,963	75,770	0,367	76,136
Cluster 3	10	78,478	7,771	97,186	0,953	98,139