

Ciencia, tecnología e innovación: una aplicación multivariante para el estado Falcón

Science, technology and innovation: a multivariate application at the Falcon state

Luis Piña¹

Domingo Maldonado²

Universidad Politécnica Territorial de Falcón Alonso Gamero (UPTFAG), Falcón, Venezuela¹

Universidad Nacional Experimental Francisco de Miranda (UNEFM), Falcón, Venezuela²

luispinalopez@gmail.com¹

dmaldona@gmail.com²

Fecha de recepción: 29/12/2020

Fecha de aceptación: 24/04/2021

Pág: 28–50

Resumen

Esta investigación tuvo como finalidad obtener un conjunto de datos multivariados sobre ciencia, tecnología e innovación en el estado Falcón, para posteriormente realizar un análisis de clúster con apoyo de programas informáticos y presentar los hallazgos más resaltantes. En este sentido, se empleó tanto el software estadístico “R” como el paquete de apoyo “FactoMineR” para generar una serie de corridas de computador con graficación que facilitaron el diseño de tres clústeres asociados a los Investigadores e Innovadores, respectivamente. Entre los resultados figuran, la posibilidad de establecer polos centrales de clústeres diferenciados con grupos de clasificación en áreas de concentración sobre ambos tipos de sujetos estudiados; el interés de los Investigadores por estar apoyados con recursos financieros y materiales, siendo esto contrastante con los Innovadores, quienes ven en la co-participación un objetivo de alto valor; y además, la existencia de algunos focos de atención: los municipios Miranda y Carirubana para los Investigadores, y la zona oriental del estado para los Innovadores; en ambos casos como clústeres geográficos principales.

Palabras clave: Análisis multivariante, ciencia, tecnología, innovación, estado Falcón.



Esta obra está bajo licencia CC BY-NC-SA 4.0.

Abstract

The purpose of this research was to obtain a multivariate data set on science, technology and innovation in the state of Falcon, to later carry out a cluster analysis with the support of computer programs and present the most outstanding findings. In this sense, both the statistical software R and the support package FactoMineR were used to generate a series of computer runs with graphing that facilitated the design of three clusters associated with researchers and innovators. The results include the possibility of establishing central poles of differentiated clusters with classification groups in concentration areas on both types of subjects studied, researchers' interest in being supported with financial and material resources, this being in contrast to innovators, who see co-participation as a high-value objective, and also, the existence of some attention centers: the Miranda and Carirubana municipalities for researchers, and the eastern part of the state for innovators; in both cases as main geographic clusters.

Key words: Multivariate analysis, science, technology, innovation, Falcon state.

Introducción

El análisis de datos multivariantes (ADM) es un área de la estadística que estudia datos resultantes de la observación con más de una variable sobre una muestra determinada. Tales variables han de ser homogéneas y correlacionadas, sin que alguna predomine sobre el resto. La información es multidimensional, por lo que la geometría, cálculo matricial y distribuciones multivariantes son fundamentales, tal como lo expresa [Cuadras, 2014]. En este sentido, interesa saber cómo están asociadas las variables porque se trata de grupos de datos multivariados, muy diferente a disponer de datos que sean univariados (o muy parecidos estructuralmente). La homogeneidad trata acerca del ámbito de estudio, así como su forma de correlacionarse, pues muchas veces los datos multivariados han de ser representados en cúmulos cuya caracterización es su fisonomía; en otras palabras, la asociación puede estar determinada según la disposición espacial de los datos en una nube.

Al mismo tiempo, la estadística como ciencia ayuda a comparar fuentes de varianza entre fenómenos para decidir sobre la aceptación o el rechazo de relaciones hipotéticas, en tanto que, facilita la elaboración de supuestos sobre observaciones empíricas [Kerlinger y Lee, 2002]. Esto permite el hallazgo de ciertos patrones de repetencia típicos en el enfoque epistemológico empírico-inductivo, que pueden explicar interdependencias entre clases e inferir probabilísticamente comportamientos futuros [Padrón, 2001]. Por otra parte, el describir una situación real como lo es caracterizar la población de un estado al asumir la ciencia, tecnología e innovación (CTI) como parte motivacional de su propio desarrollo, requiere tener en cuenta simultáneamente varias variables; por lo que el objeto de estudio estadístico pretendería resumir al conjunto de variables en otros nuevos construidos como

transformaciones de las variables originales con una mínima pérdida informativa, verificar la existencia de grupos en los datos, clasificar nuevas observaciones en grupos definidos, y finalmente relacionar los conjuntos de variables, lo que es referido por [Peña, 2002].

Al tratarse de Venezuela, el término CTI adquiere firmeza por cuanto existe una ley que lo promulga, la Ley Orgánica de Ciencia, Tecnología e Innovación (LOCTI), [MPPCTI, 2010]; una base constitucional, que es la Constitución de la República Bolivariana de Venezuela [Constitución, 2009], y planes nacionales e instituciones establecidos a partir de organismos de interacción. La información estadística que manejan es representativa de las unidades territoriales donde subyacen y del contexto donde se desenvuelven los innovadores e investigadores registrados. Sin embargo, únicamente obedecen a un grupo que ciertamente es significativo (o al menos representativo), de un universo “ideal” de productores del conocimiento, ya que habría que contrastar la información resultante con aquellos no registrados (o no actualizados), ante las estadísticas emitidas en el portal del Observatorio Nacional de Ciencia, Tecnología e Innovación [ONCTI, 2016].

Es importante señalar, que desde el año 2018 no se realizan arqueos a través del ONCTI, y en su lugar el Ministerio del Poder Popular para Ciencia y Tecnología (MPPCT) ha creado el Plan Nacional de Innovación Tecnológica (PNIT) a partir de 2019 para propiciar la labor investigativa entre los Investigaciones e Innovadores, independientemente de corresponder a alguna adscripción institucional en el país. Este cuenta con áreas definidas según prioridades nacionales, que agrupan a una población general estimada entre cincuenta (50) y setenta (70) personas, con solicitudes de financiamiento que esperan por su aprobación final y desembolso para la ejecución y evaluación transformacional.

Planteamiento

La problemática abordada se enfoca hacia los investigadores e innovadores en el estado Falcón, ya que no refieren a un mismo género o tipología, en otras palabras, los investigadores corresponden a un tipo de individuo que, generalmente, no es el mismo de los innovadores; con algunas excepciones. De por sí, ha sido una tradición que como investigadores se agrupen aquellos docentes que ejercen en Instituciones de Educación Universitaria (IEU), y al mismo tiempo cuenten con el aval mínimo de productos de investigación como para calificar ante el Registro Nacional de Innovadores e Investigadores (RNII) del ONCTI. Con respecto a los innovadores, se estima que la participación diversa entre artesanos, tecnólogos populares y demás cultores ocupe un espacio importante, mas sin embargo, siempre han representado una proporción que gira en torno al diez por ciento (10 %) del total de acreditados en el Programa de Estímulo a la Innovación e Investigación (PEII), según información de la Unidad Territorial Falcón de Fundacite.

Internamente, ocurren varias desigualdades que básicamente se presentan entre los

innovadores, como por ejemplo, docentes registrados como innovadores en lugar de haber sido clasificados como investigadores, y algunos estudiantes, cultores o demás actores de participación con incipientes requisitos, que no lejos de avizorar su proyección dentro del renglón que han ocupado, su continuidad en el programa rara vez pasa de los primeros períodos. A partir del año 2011, cuando empieza a regir el PEII, producto de la reestructuración del anterior Programa de Promoción al Investigador (PPI) en las Instituciones de Educación Universitaria (IEU), se han podido obtener las siguientes acreditaciones en Investigadores e Innovadores: 2011: ciento setenta y siete (177) y nueve (9); 2012: trescientos treinta y siete (337) y veinticinco (25); 2013: quinientos doce (512) y cincuenta (50); 2014: quinientos cuarenta (540) y treinta y tres (33); y 2015: cuatrocientos setenta (470) y treinta y cinco (35), respectivamente; así como cuatro (4) innovadores en 2019 (PNIT).

Como tal, el PEII ha sido muy discutido por cuanto abarca gran parte de actores, estén o no bajo el amparo de instituciones concebidas para la investigación o demás organismos oficiales, es decir, es de amplio espectro en el sentido del conjunto que agrupa; y a expensas de las clasificaciones que pudieren vislumbrarse, las estadísticas implican indicadores según los renglones de: recursos financieros, talento humano, productos de investigación y proyectos de investigación [ONCTI, 2016].

En estas, se colocan resultados a través de tablas y gráficos con breves comentarios que únicamente redefinen un comportamiento en alta, sostenimiento o baja porcentual con respecto a los períodos anteriores, con breves explicaciones que justifican las razones de su ocurrencia; esto pudo ser apreciado en el último documento conocido, emitido en el año 2016. En todo caso, aún falta más por evidenciar tales razones, lo que ha requerido de un estudio estadístico-multivariante que ayude a responder descriptivamente dichos renglones; o bien, aclaren el por qué tanto individuos como proyectos se circunscriben en el marco de CTI, lo relevante de sus contribuciones bajo el apoyo de varios organismos, la corresponsabilidad entre grupos, el reconocimiento de patrones, y en general, cómo algunas variables asociadas tanto a ellos como a CTI se manifiestan.

Por lo tanto, fue realizada una investigación que articulara múltiples datos partiendo de información heterogénea, para luego proceder a agruparlos por homogeneidad a los fines de lograr una apreciación de fácil interpretación. Para esto, se requirió de la utilización simultánea de todas las estructuras de grupos de datos, aunque sin suposiciones previas sobre las distribuciones de los sujetos o entidades; a expensas de que existan diferencias significativas entre sí. Así, el estudio multivariante permitiría la exploración exhaustiva de los datos en CTI regional, al tanto de conocer más sobre las observaciones, en especial si tanto investigadores como innovadores falconianos tienden a agruparse o no, y sobre la base de cuáles criterios manejar. Además, ya que se ha presumido del análisis clúster como técnica eminentemente descriptiva, ésta podría confirmar la validez definitiva del instrumento empleado, no obstante, con el uso de una prueba estadístico-inferencial se consiguió la conformación estructural de los

grupos creados y su determinación para un diseño de clúster.

En atención a lo expuesto, el objetivo fue el de estudiar la relación entre variables de dependencia e interdependencia asociadas al contexto de CTI en el estado Falcón mediante el análisis de conglomerados. Para lograrlo, se previó de: 1. Obtener un conjunto de datos multivariados relacionados con CTI en el contexto del estado Falcón, 2. Realizar un análisis por conglomerados a los datos multivariados encontrados con apoyo de programas informáticos, y finalmente, 3. Presentar los hallazgos resultantes de los clúster conformados en CTI falconiana con base en el análisis efectuado. A tales efectos, se requirió de suficiente versatilidad gráfica tridimensional dada la intención de atenuar la apreciación del diseño de clúster en 3D, de mejores características de representación frente a un diseño bidimensional (2D).

Precisiones teóricas

Existen algunos estudios previos relevantes, como el de Ferrer, 2014, quien trabajó con innovadores estimando una muestra de diecinueve (19) Consejos Comunales y veinticuatro (24) tecnólogos populares pertenecientes a diez (10) municipios falconianos, donde recalcó la importancia que desempeña la sociedad organizada dentro de la gestión tecnológica, apoyada en los Consejos Comunales como parte del poder popular. Asimismo, el ONCTI como agente emisor de estadísticas en boletines publicados por el MPPEUCT, generó varios recuentos clasificados traducidos en indicadores sobre CTI en el año 2016, mostrando determinados valores clasificados y totalizados.

Es conveniente destacar que, los grupos temáticos son de cobertura nacional, y el informe hace gala descriptiva de la cédula de indicadores, su definición, fórmula de cálculo, concepto de variables que le componen, metodología, forma de recolección de los datos, periodicidad y algunos comentarios. Si bien otros trabajos han propiciado ideas locales, el caso de Zamora et al., 2014, resulta totalmente cónsono. Allí precisan que luego de la convocatoria PEII para ese año, y posterior al proceso de evaluación final, renovación y reparos; Falcón alcanzó la cifra de seiscientos veintisiete (627) investigadores, y setenta y tres (73) innovadores acreditados. Comenta además, sobre el impulso y trascendencia de las Redes Socialistas de Innovación Productiva (RSIP) y de la Unidad Territorial Falcón de Fundacite en torno a los proyectos y diversidad de objetivos que administra. Como documentos ilustrativos de la región, figuran los catálogos del Instituto del Patrimonio Cultural, producto del Censo del Patrimonio Cultural Venezolano (el primero de estos, realizado en el año 2004); en especial, los referidos a todos los municipios de la entidad falconiana.

De manera crítica, Gutiérrez, 2016, elaboró la redacción editorial: “ La investigación e innovación hoy en Venezuela”, resaltando aspectos débiles del sistema nacional. Recalca la crisis que vive el país, reflejando que todos los indicadores y estadísticas revelan el retroceso a niveles alarmantes del desarrollo científico, tecnológico, al igual que en la generación de

innovaciones [Gutiérrez, 2016]. Se apoya en trabajos publicados, como el de Van Noorden, 2014, quien mostró estadísticas del declive venezolano en producción científica: una caída del veintinueve por ciento (29 %) entre los años 2009 y 2013; sumado esto al fenómeno migratorio de investigadores, visto como pérdida de talento [Requena y Caputo, 2016].

Ahora bien, ¿qué son los métodos multivariantes?. Dado que en estadística existen técnicas univariantes y bivariantes, el uso del prefijo “multi” denota una extensión mucho más general. Además, se centran en la investigación simultánea de tres o más características que parten de un conjunto determinado; es decir, varias variables que pueden condicionar a personas, entidades u objetos. Por tanto, estos métodos son un conjunto de técnicas particulares que tienen como fin el análisis paralelo y compatible de grupos de datos multivariados, dado que existen varias variables medidas para cada asunto en estudio [Closas et al., 2013].

La esencia de las técnicas multivariantes radica en la estadística asistida mediante software para modelaje matemático de múltiples variables relacionales. A saber, los computadores han resultado cada vez con mayores potencialidades en cuanto a capacidad, velocidad y disponibilidad que en la década de 1960, cuando surge un interés masivo por la estadística multivariante [Tinsley y Brown, 2000]. Por esto, existe la tendencia actual por emplear varios tipos de procedimientos, o lo que pudiere llamarse multi-técnicas para explorar casos con datos estadísticos multivariantes. En tal sentido, al conjunto de técnicas también puede catalogarse de “herramientas”, debido al uso de programas informáticos para su procesamiento incorporando capacidades para la emisión de ambientes gráficos versátiles (tridimensionales).

A todo esto, ¿cómo surgen o se derivan tales multi-técnicas o herramientas en el ADM?. La respuesta yace en el estudio de varianzas que al pretender relacionarle varias variables originó lo que se conoce como *MANOVA* (*multivariate analysis of variance*). Primeramente surge *ANOVA* (*analysis of variance*), que comprende un grupo de técnicas estadísticas para analizar cómo operan diversos factores simultáneamente en un mismo diseño factorial sobre una variable respuesta; según [Cuadras, 2014]. En seguida, la proporción de métodos generados hizo que se reorganizaran en dos áreas: dependencia e interdependencia, abarcando regresión lineal múltiple, análisis discriminante, regresión de variable dependiente limitada, correlación canónica, *MANOVA* y ecuaciones estructurales dentro de la primera área; mientras que análisis de componentes principales, análisis factorial, análisis de correspondencias, análisis de conglomerados y escalamiento multidimensional en la segunda área. Los métodos de interdependencia (descriptivos) se usan en la situación de que sea imposible distinguir entre variables dependientes e independientes. Su único interés, es el de determinar cómo y por qué se encuentran correlacionadas las variables entre sí.

En este sentido, la idea de emplear una técnica exploratoria inicial hace que el análisis de conglomerados sea atractivo, para luego aplicar otras técnicas de mayor rigor. De esta manera,

los planteamientos sobre análisis exploratorio multivariante de tipo clúster estimados por [Husson y Josse, 2010, Husson et al., 2011], resultan propicios por cuanto la heterogeneidad de una población representa la materia prima del análisis cuantitativo: 1. se utiliza información de una serie de variables sobre cada sujeto u objeto, 2. conforme a dichas variables se mide la similitud, 3. se crean grupos homogéneos internamente y distintos entre sí, y 4. la nueva disposición se aprovecha para facilitar un determinado análisis.

La intencionalidad por disponer de los datos juega un papel crucial al inicio de la exploración, por lo tanto, se debe desestimar el colocar a las “unidades de datos” en arreglos que conlleven a estudiarles aisladamente. Lo correcto es mostrarles por su consistencia en los términos característicos de sus atributos, clases u otro tipo de propiedades, ya que apreciadas colectivamente son los descriptores que representan a las variables del problema [Anderberg, 1973]. Proviene de registros documentales, así como de la aplicación de instrumentos para su recolección; en todo caso, tienen un origen empírico, con muestras poblacionales de cien (100) a doscientos (200), colocados en hojas de datos para su procesamiento informático: estadísticas resumidas y visualización gráfica.

La fase exploratoria multivariante es rica en aplicaciones con multi-técnicas, por esto al referirse sobre análisis clúster inmediatamente se sintetiza en aquellas de mayor trascendencia: matrices de similaridad y de distancia, algoritmos para clúster, jerarquización, partición, solapamiento u ordenación; mas sin embargo, al precisar qué técnica(s) deban emplearse en esta temática (ADM en CTI falcóniana), es justo apostar por las más asertivas en el manejo de los datos, como lo son: análisis de componentes principales, similaridad, partición, jerarquización, aglomeración de Ward u otras que [Husson et al., 2011], han establecido para “clusterización”. El proceso analítico presume iniciar con el estudio de variables que forman nubes de individuales a semejanza de las distancias euclídeas.

Ya que el manejo de la *data* es crucial, los criterios enmarcados por programas como R y FactoMineR serán determinantes en la interpretación, considerando que todo valor es “procesable” a expensas de algunas diferencias (tabla 1).

Con respecto al análisis clúster, es una técnica que se apoya en bases de datos multidimensionales cuyo procesamiento genera “objetos de representación” que suelen ser: sujetos, variables o categorías de variables categóricas, las cuales corresponden a filas y columnas de los datos tabulados. La forma típica de representar a un conjunto de objetos es la nube de puntos, donde cada punto es un objeto envuelto dentro de un espacio euclídeo (llevado al plano cartesiano). El término “euclídeo” hace referencia a la distancia entre puntos (con respecto a los ángulos formados entre las variables cuantitativas); a su vez, interpretados por la similaridad entre individuos o categorías (en cuanto a correlación, para variables cuantitativas). En la técnica clúster, el análisis de componentes principales y el análisis de correspondencias se encuentran vinculadas a las representaciones euclídeas [Husson et al., 2011].

Tabla 1: Ventajas y desventajas de SAS, SPSS, R y Python

	SAS	SPSS	R	Python
Ventajas	<ul style="list-style-type: none"> - Alta aceptación en la industria. Asistencia y soporte. - Manejo de grandes grupos de datos. -Flujos basados en interface pesada. - Enlaces SQL 	<ul style="list-style-type: none"> - Usado en muchas universidades. - Buena interface de usuario con gran documentación. - Funcionalidad click-play. - Fácil escritura de código (copy paste). - Asistencia y soporte 	<ul style="list-style-type: none"> - Comunidad amplia co-creadora de librerías. - Código fuente libre. - Rápida capacidad explicativa y de modelaje predictivo. - Fácil conexión a fuentes de datos: NoSQL, <i>webscraping</i> 	<ul style="list-style-type: none"> - Escalabilidad. - Lenguaje de propósito general. - Fácil de aprender. - Bueno en el aprendizaje mecánico. - Gran comunidad. - Código fuente libre
Desventajas	<ul style="list-style-type: none"> - Altos costos. - Deben escribirse códigos fuentes - Lenta adaptación a nuevas técnicas. - Diferentes paquetes para visualización o minería de datos 	<ul style="list-style-type: none"> - Altos costos. - Diferentes licencias para cada funcionalidad. - Sintaxis limitada - Lenta adaptación a nuevas técnicas. - Manejo lento de grandes grupos de datos 	<ul style="list-style-type: none"> - Puede ser lento en el manejo de grandes grupos de datos. - Lenta curva de aprendizaje. - Sin soporte oficial. - Sin interface de usuario 	<ul style="list-style-type: none"> - Débil capacidad para el modelado explicativo - Elección de versión aún no definida: ¿2.7 ó 3.5?. -Sin interface de usuario. - Sin soporte oficial

Fuente: Elaboración propia (2020)

Otra forma de representarles e ilustrarles, es por medio del árbol jerárquico, o más precisamente bajo el concepto de indexación jerárquica, la cual establece niveles donde los objetos son agrupados para interpretarles gráficamente en lo que se conoce como “dendograma”.

Metodología

Se parte de que esta investigación es regida bajo el enfoque epistemológico empírico-inductivo según Padrón, 2001, de tipo “kuhniano”, y catalogada según como otros denominan: metodología cuantitativa, positivista o medicionista. Representa una postura asumida donde coexisten las variaciones: “descriptivas”, por cuanto se focalizan registros de una problemática que estadísticamente ha sido centrada en datos de CTI para una región como la comprendida por el espacio geográfico falconiano. Como tal, se evidenciaron patrones de regularidad siguiendo las premisas de [Peña, 2002, Cuadras, 2014, Kerlinger y Lee, 2002, Closas et al., 2013, Husson et al., 2011], entre otros.

Asimismo, fueron revisados boletines del ONCTI hasta el 2016, información de Fundacite Falcón, algunas IEU y demás fuentes referenciales en análisis multivariante y uso de software para agilizar su procesamiento, específicamente R y FactoMineR. De esta forma, los ítems de cada instrumento de recogida de datos emergieron empíricamente con respecto a los hechos

observados desde inicios de la aproximación al objeto de estudio; por tanto, fueron creadas dos listas de descripción con escalas de clasificación: una aplicada a los investigadores y la otra a innovadores, ambas dentro del contexto multivariado en CTI falconiana. Los sujetos informantes investigadores fueron profesionales universitarios adscritos al PEII (no limitante), sin menoscabo del nivel al cual han sido clasificados e indistintamente de su condición laboral (ordinaria o no) en la institución donde prestan servicios; y como innovadores, se abordó aquellos personajes figurados en el PEII (tampoco restrictivo) o reconocidos por otros organismos oficiales, principalmente tecnólogos populares o artesanos.

La *data* fue cotejada por *MANOVA* y correlación canónica, así como su clusterización. Previamente, se calculó la confiabilidad mediante Alfa de Cronbach en escalamientos *Likert* sobre las sumatorias de varianzas y las varianzas totales, resultando valores aceptables.

Estructuración estadística

Durante la fase inicial, se obtuvo que para el ADM en CTI falconiana las variables consideradas para los investigadores e innovadores fueron: producción científica y proyectos adscritos, así como tecnología popular y artesanía, respectivamente. Con ellas se logró establecer la operacionalización de variables que se muestra a continuación (tabla 2):

Tabla 2: Operacionalización de variables

	Variable	Dimensión	Sub-dimensión	Indicadores	Ítem (A y B)
Investigadores (A)	Producción científica	Investigaciones	Trabajo científico informes largos	Investigaciones libres	1, 2, 3, 4, 5
			Síntesis arbitrable (informes cortos)	Publicaciones	6, 7, 8, 9, 10
	Proyectos adscritos	Planes de desarrollo	EIUF / Centros de investigación	Líneas de investigación	11, 12, 13, 14, 15
			Fuentes de Financiamiento	Apoyo financiero	16, 17, 18, 19, 20
Innovadores (B)	Tecnología Popular	Desarrollos tecnológicos	Trabajos individualizados	Prototipos, aparatos y desarrollos finales	1, 2, 3, 4, 5
			Co-participaciones	Apoyos interinstitucionales	6, 7, 8, 9, 10
	Artesanía	Rubros	Recursos materiales	Aprovisionamiento	11, 12, 13, 14, 15
			Productividad y calidad	Productos finales	16, 17, 18, 19, 20

Fuente: Elaboración propia (2020)

Como parte fundamental del ADM, el tamaño muestral suele ser igual o mayor a cien (100) objetos, según [Closas et al., 2013]; aunque no es un hábito restrictivo entendiéndose que existe la valoración al dejar entrever la existencia de información importante en cada uno de los datos multivariados.

En este sentido, el volumen de investigadores e innovadores que se consideró manejar fue de una data general de cien (100) sujetos: sesenta y cinco (65), y treinta y cinco (35) para cada caso; con un cúmulo variado de profesores-investigadores de IEU ubicados en el eje de los municipios Miranda-Carirubana-Falcón-Democracia-Urumaco-Tocópero-Píritu; por concentrar buena parte del conglomerado de interés, y de la misma forma para tecnólogos populares y artesanos.

Anticipando el proceso de análisis, se reclasificaron los ítems instrumentales por áreas de incidencia temática (tabla 3), respetando la operacionalización de variables antes presentada. Además, se efectuaron pruebas pilotos antes de proceder con las aplicaciones de rigor, siendo los valores conseguidos de α : A con trece (13) investigadores, y B con siete (7) artesanos o tecnólogos populares, de setenta y siete por ciento (0,77) para el instrumento A y de ochenta y cuatro por ciento (0,84) para el B; es decir, aceptables.

Tabla 3: Incidencia temática por áreas

Investigadores (Instrumento A)		Innovadores (Instrumento B)	
Áreas temáticas	Ítems asociados	Áreas temáticas	Ítems asociados
Intencionalidad investigativa, perspectivas de divulgación	1, 2, 3, 4, 5, 6, 7, 8, 9, 10	Construcción de piezas o artefactos por iniciativa propia o participativa	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Participación en proyectos de investigación financiados para el desarrollo	11, 12, 13, 14, 15, 16, 17, 18, 19, 20	Productividad artesanal en base al aprovisionamiento y emprendimiento	11, 12, 13, 14, 15, 16, 17, 18, 19, 20

Nota: Existencia de relaciones inter-ítems e inter-temáticas

Fuente: Elaboración propia (2020)

Se continuó con la aplicación instrumental hasta completar 100 sujetos, al igual que su tabulación multivariada donde además de los valores ítems y la ubicación municipal fueron incluidos: Instituciones de Educación Universitaria (IEU) e Instituciones No Universitarias (INU), en la manera de IEU/INU, Producción, Zona; Oficio y Aval. Resumen en las tablas 4 y 5

Tabla 4: Resumen tabular: Investigadores falconianos

Suj.	Ítems (K_i)											Municipio	Zona	IEU/INU	Producción
	1	2	3	4	5	6	7	8	9	...	20				
1	1	1	2	1	3	2	3	3	1	..	1	Miranda	Central	UPTFAG	No
2	3	1	2	2	2	1	2	1	2	...	1	Miranda	Central	UPTFAG	Si
3	1	1	1	3	1	1	2	2	3	...	2	Miranda	Central	UPTFAG	Si
4	1	1	2	4	1	4	1	1	1	...	2	Colina	Central	UNEFM	Si
...
65	2	2	1	1	2	1	1	3	3	...	2	Miranda	Central	INIA	Si

Fuente: Elaboración a partir de datos recabados (2020)

Tabla 5: Resumen tabular: Investigadores falconianos

Suj.	Ítems (Ki)											Municipio	Zona	IEU/INU	Producción
	1	2	3	4	5	6	7	8	9	...	20				
1	2	1	2	1	1	1	1	1	1	..	1	Colina	Central	Artesano	Si
2	1	1	2	1	1	1	3	1	1	...	1	Miranda	Central	Tec. Pop.	Si
3	1	1	1	1	1	1	1	1	1	...	1	Falcón	Paraguaná	Artesano	Si
4	2	1	1	1	1	1	1	2	1	...	1	Colina	Central	Tec. Pop.	Si
...
35	1	1	1	1	1	1	1	1	1	...	1	Colina	Central	Artesano	No

Fuente: Elaboración a partir de datos recabados (2020)

De manera resumida, la data recabada por aplicación del instrumento A en docentes-investigadores tomó las siguientes locaciones: UBV-Paraguaná (1), UNEFM-El Sabino (6), UNEFA-Paraguaná (4), UPTFAG-Punto Fijo (2), LUZ-PF (1), UPTFAG-Coro (23), UNEFM-Los Perozo (10), UNA-Coro (1), UNESR-Coro (1), UNEFA-Coro (1), UNERMB-Coro (3), INIA-Coro (1), UPTFAG-Dabajuro (3), UNEFM-Churuguara (3), UNEFM-El Hatillo (2) y UNEFM-Tucacas y Morón, estado Carabobo (3). Asimismo, los datos que corresponden por aplicación del instrumento B se valieron de artesanos y tecnólogos populares ubicados en los municipios: Falcón (3), Miranda (9), Urumaco (1), Dabajuro (2), Colina (9), Zamora (5), Jacura (2) y Petit (4).

Dado que las tablas fueron transcritas con extensión “.xlsx”, se debió reconvertirlas en archivos “CSV” para compatibilidad en R y FactoMineR. De esta manera, los datos contentivos en los ficheros “DATAInv_IL_4” y “DATAInn_IL_4”, se trataron con lenguaje de programación R a través de los siguientes comandos:

```
pfmrdatainv4 = read.table("C:/Users/personal/Desktop/Luis Piña/
IL N° 4/PROCESAM FactoMineR/DATAInv_IL_4.csv", header=TRUE,
sep=";", dec=".", row.names=1)
```

```
summary(pfmrdatainv4)
```

En el primer comando, la instrucción “read.table()” efectuó la captura/lectura de data, y el segundo comando: “summary()”, permitió desplegar un resumen estadístico de lo encontrado en la variable “pfmrdatainv4”. Inmediatamente, fue cargada la librería FactoMineR empleando el comando:

```
library(FactoMineR)
```

Por medio del análisis de componentes principales, fue posible generar dos gráficos: el mapa de variables (en este caso, los veinte ítems del instrumento A), y el mapa de individuos (la opinión de los sesenta y cinco sujetos sobre tales ítems); de la forma:

```
res<- PCA(pfmrdatainv4[,1:20])
```

Ahora bien, tales mapas se muestran seguidamente en las figuras 1 y 2, y luego se tomó un procedimiento similar para los innovadores con las mismas líneas de comando, pero para la variable “pfmrdatainn4”, resultando otros mapas (figuras 3 y 4).

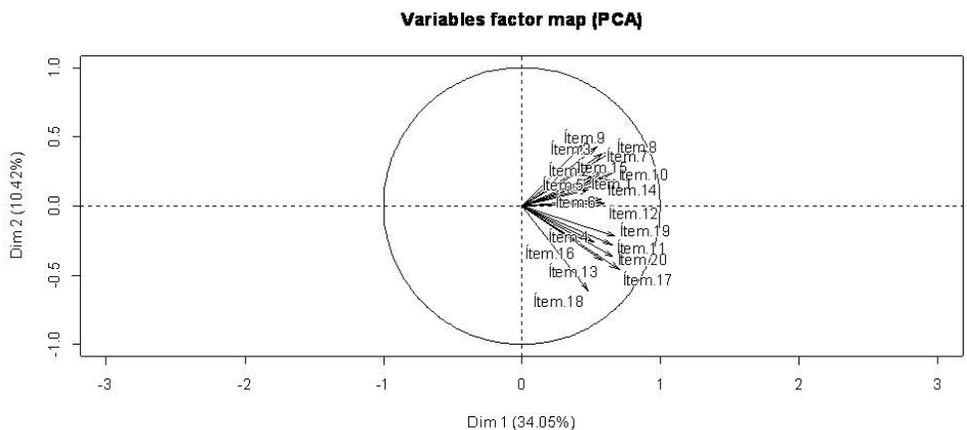


Figura 1: Mapa de variables PCA de “pfmrdatainv4”
Fuente: Elaboración obtenida a través de R Console (2020)

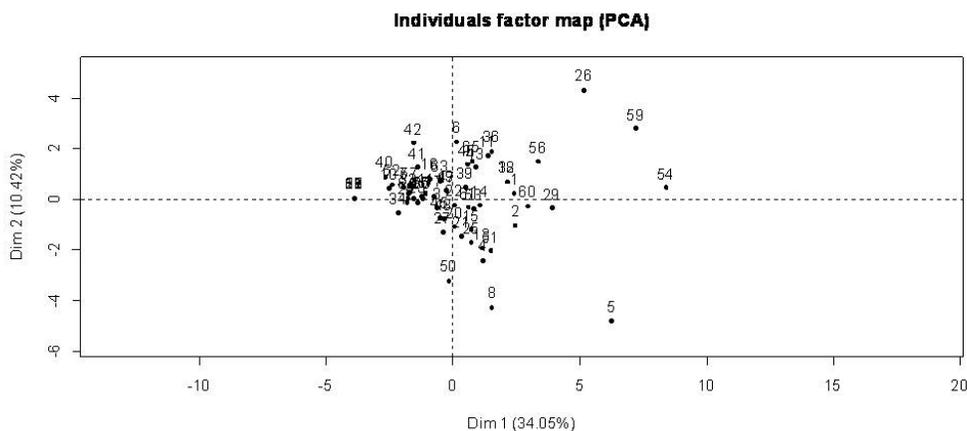


Figura 2: Mapa de individuos PCA de “pfmrdatainv4”
Fuente: Elaboración obtenida a través de R Console (2020)

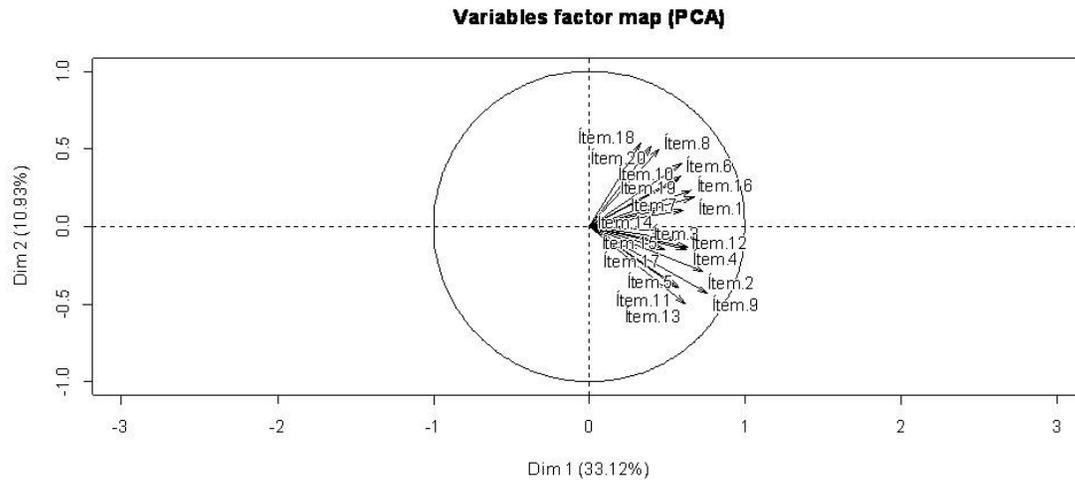


Figura 3: Mapa de variables PCA de “pfmrdainn4”
Fuente: Elaboración obtenida a través de R Console (2020)

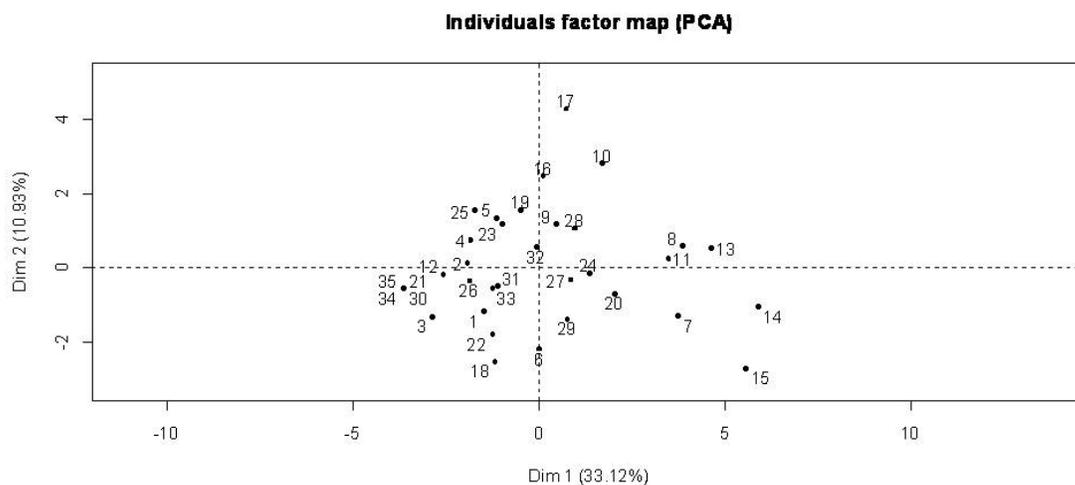


Figura 4: Mapa de individuos PCA de “pfmrdainn4”
Fuente: Elaboración obtenida a través de R Console (2020)

En vistas del reconocimiento de datos multivariado, se pudo obtener resultados de tipo ANOVA, entre estas, la evidencia de valores en “áreas positivas” del mapa de variables en Investigadores para las dimensiones 1 y 2 (34,05 % y 10,42 %), específicamente en los ítems 3, 5, 6, 7, 8, 9, 10 y 14 (figura 1); así como la forma de la nube que describen los sujetos en el mapa de individuos (figura 2). Por su parte, en innovadores las “áreas positivas” destacan los

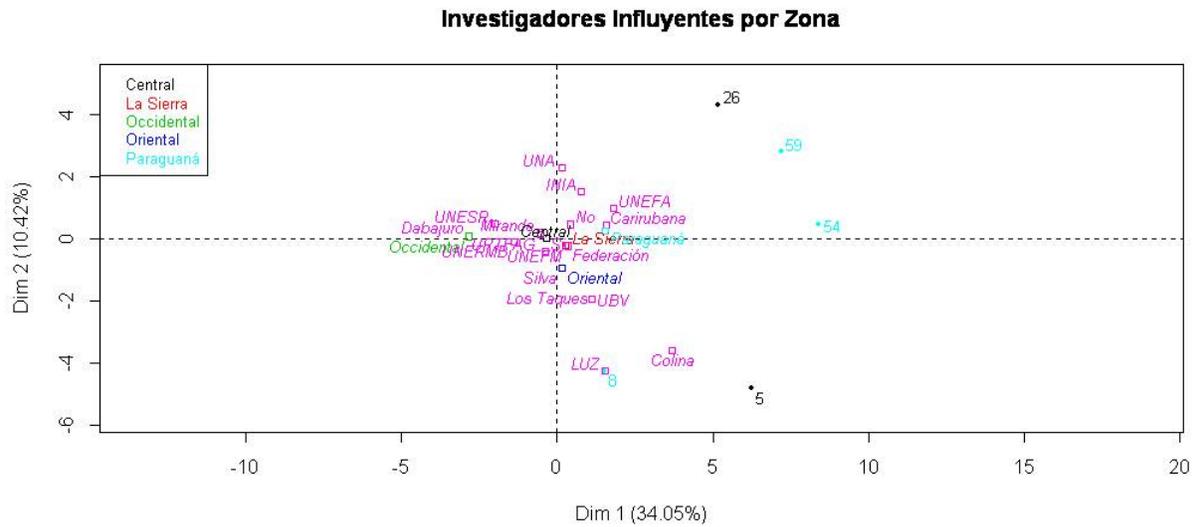


Figura 6: Mapa de investigadores influyentes → variable: Zona
Fuente: Elaboración obtenida a través de R Console (2020)

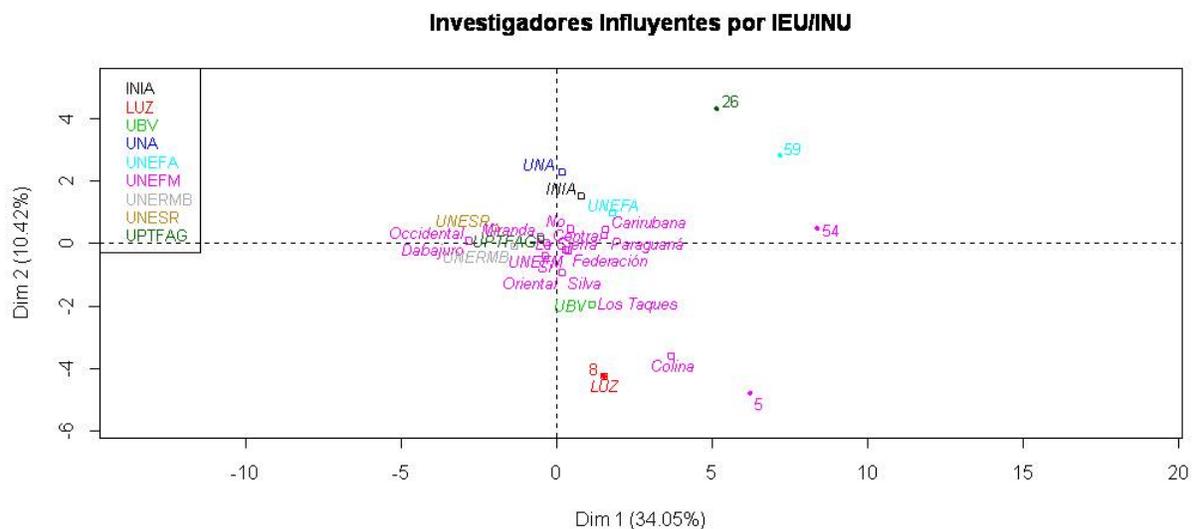


Figura 7: Mapa de investigadores influyentes → variable: IEU/INU
Fuente: Elaboración obtenida a través de R Console (2020)

Investigadores Influyentes por Producción Científica

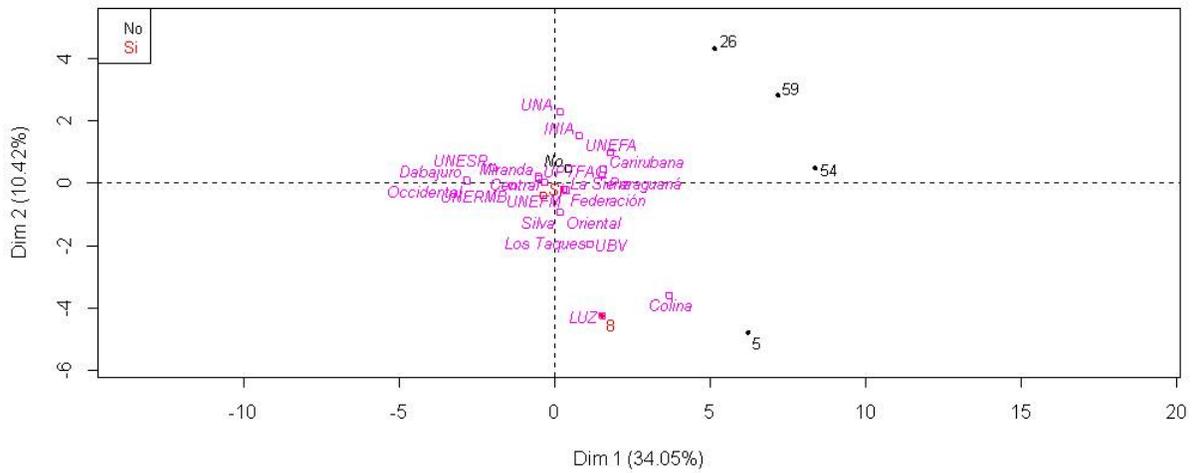


Figura 8: Mapa de investigadores influyentes → variable: Producción
 Fuente: Elaboración obtenida a través de R Console (2020)

Innovadores Influyentes por Municipio

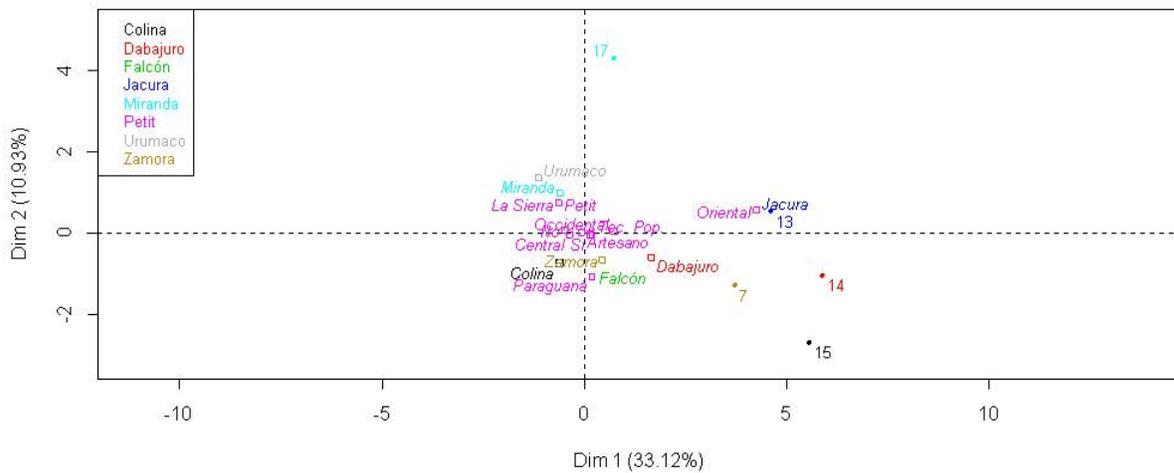


Figura 9: Mapa de innovadores influyentes → variable: Municipio
 Fuente: Elaboración obtenida a través de R Console (2020)

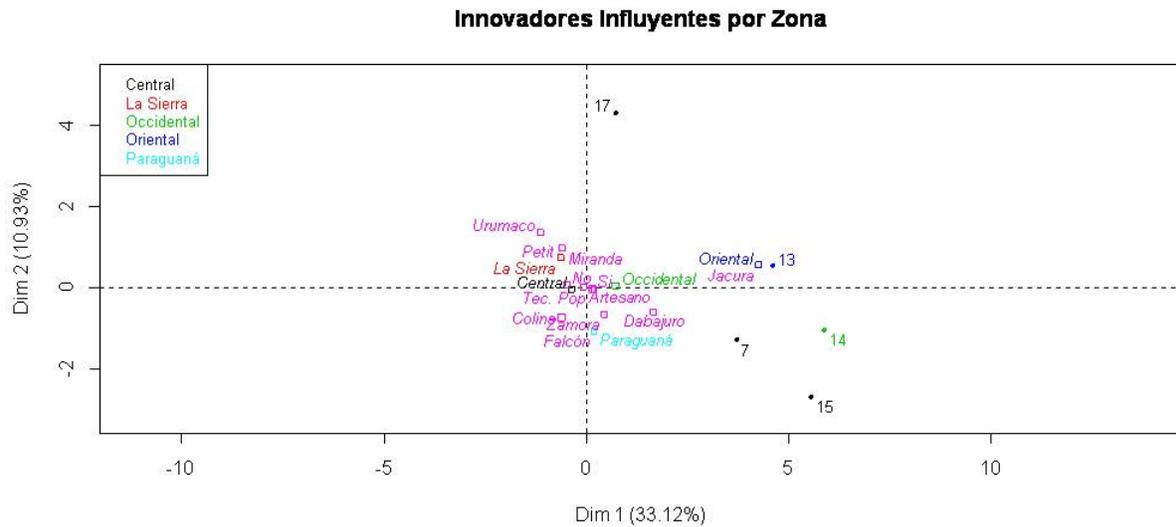


Figura 10: Mapa de innovadores influyentes → variable: Zona
 Fuente: Elaboración obtenida a través de R Console (2020)

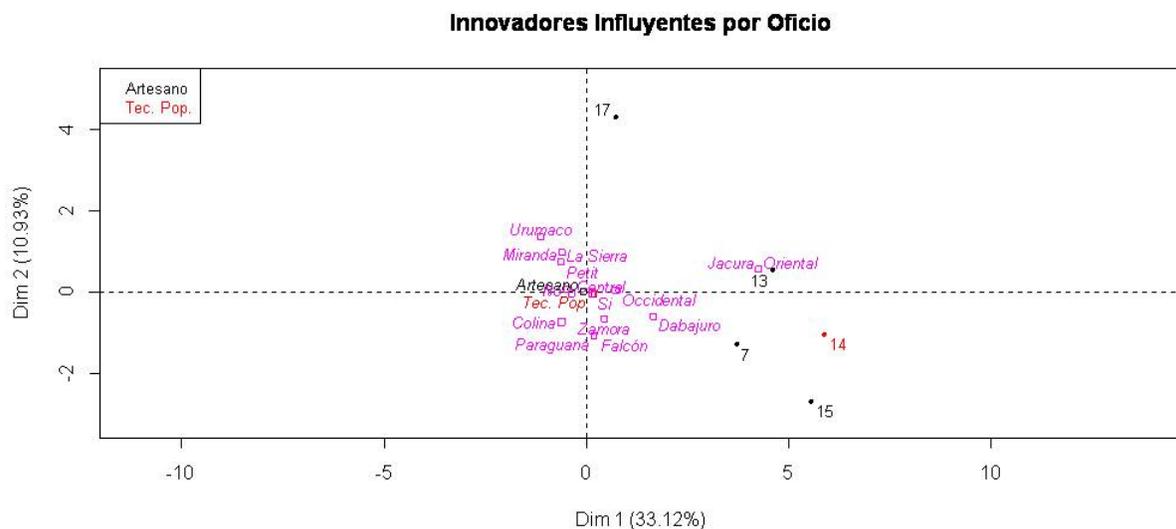


Figura 11: Mapa de innovadores influyentes → variable: Oficio
 Fuente: Elaboración obtenida a través de R Console (2020)

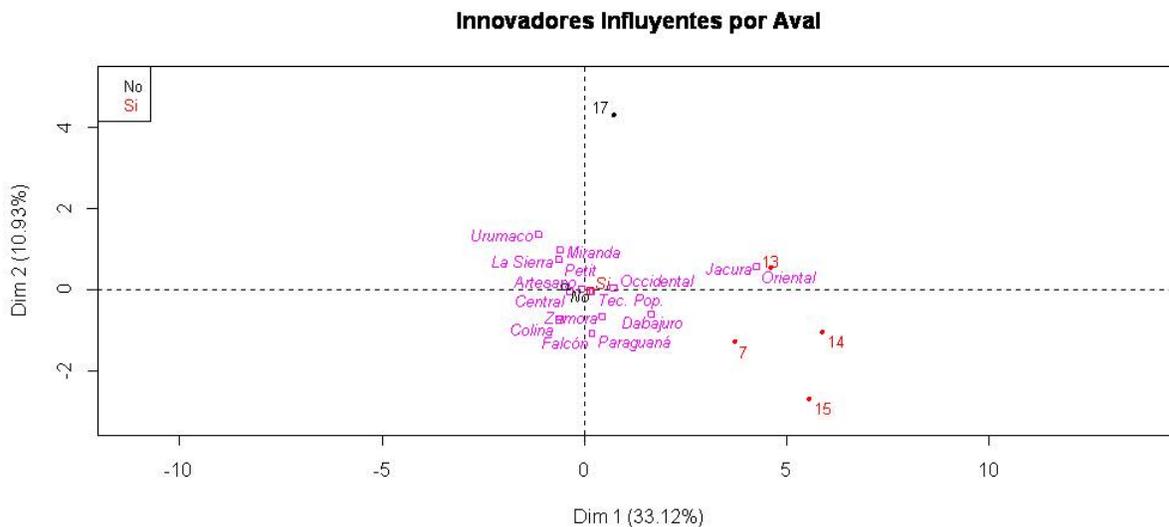


Figura 12: Mapa de innovadores influyentes → variable: Aval
 Fuente: Elaboración obtenida a través de R Console (2020)

Interpretación analítica multivariante

Hay más proporción de Investigadores en las zonas Central, con cuarenta y dos (42), y Paraguaná, con catorce (14), debido a que en los municipios Miranda, con cuarenta (40) y Carirubana, con trece (13), se concentran la mayor parte. El peso lo tienen la UPTFAG con veintiocho (28) y la UNEFM con veinticuatro (24). Al revisar la data de producción, un ligero margen se inclina al “Si”, con treinta y cuatro (34) frente al “No”, con treinta y uno (31); o sea, tienen discreta participación en actividades científicas. En Innovadores, resultó mayor la participación de artesanos sobre tecnólogos populares, de veinticinco a diez (25 a 10), con proporciones más cónsonas en Colina y Miranda, con nueve (9), y Zamora y Petit, con cinco y cuatro (5 y 4, respectivamente). La zona Central fue la que obtuvo mayor registro con veintitrés (23), con registro aval del “Si” de veintiocho (28) sobre el “No”, con siete (7). Los artesanos centrales reconocidos son mayoría en Falcón frente a los tecnólogos populares (tablas 6 y 7, figura 13). Anteriormente, las primeras dos figuras (1 y 2) abordaron las asociaciones entre variables e individuos de los investigadores. Luego, las dos siguientes (figuras 3 y 4), las asociaciones relativas a los innovadores, sosteniendo pocas dispersiones sobre los individuos, asunto reflejado en los mapas de influencia por variable para ambos tipos de actores (figuras 5 a la 12).

Tabla 6: Descriptores de clústeres (Investigadores)

Conglomerado	Nro. de Sujetos	Características categorial	
Clúster 1	29	→	^a : Cla/Mod = 57.50
Clúster 2	30		^a : Mod/Cla = 76.67
Clúster 3	6		^b : Cla/Mod = 21.43
Total de Sujetos	65		^b : Cla/Mod = 10.00
% Acumulado de varianza (componentes 1 y 2)		34,05206 + 44,47298 = 78,52504	
Dimensiones (mapa de factores)		Dim 1: 34,05 %	Dim 2: 10,42 %
Clases de interés	Municipio= Miranda ^a a	Zona= Paraguaná ^b	

Fuente: Elaboración obtenida a través de R Console (2020)

Tabla 7: Descriptores de clústeres (Innovadores)

Conglomerado	Nro. de Sujetos	Características categorial	
Clúster 1	16	→	^a : Cla/Mod = 100.00
Clúster 2	12		^a : Mod/Cla = 28.57
Clúster 3	7		^b : Cla/Mod = 100.00
Total de Sujetos	35	^b : Cla/Mod = 28.57	
% Acumulado de varianza (componentes 1 y 2)		33,12065 + 44,05364 = 77,17429	
Dimensiones (mapa de factores)		Dim 1: 33,12 %	Dim 2: 10,93 %
Clases de interés	Municipio= Jacura ^a a	Zona= Oriental ^b	

Fuente: Elaboración obtenida a través de R Console (2020)

Al respecto, las tablas 6 y 7 presentados denotan un resumen de las caracterizaciones estadísticas arrojadas por el software, donde se especifican conglomerados de interés al igual que clases de interés, en especial del clúster 1 para Investigadores (como polo potencial de desarrollo), junto al municipio Miranda y consecuentemente la zona de Paraguaná. De igual forma, del clúster 3 para Innovadores con especial foco en el municipio Jacura, subsecuentemente y la zona oriental del estado Falcón.

Clústeres jerárquicos sobre mapas de factores

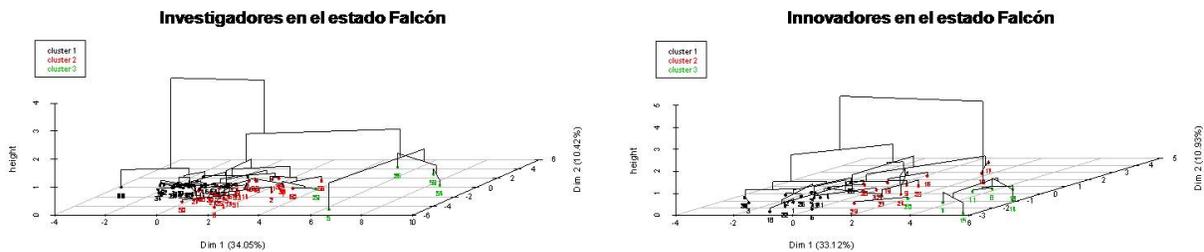


Figura 13: Clústeres de Investigadores e Innovadores en Falcón

Fuente: Elaboración obtenida a través de R Console (2020)

Conclusiones

El desarrollo de esta investigación ha permitido erigir las siguientes aseveraciones en relación a una clusterización sobre Investigadores e Innovadores en el estado Falcón, las cuales se enuncian así:

- La posibilidad de establecer polos centrales en los clústeres diseñados representa una alternativa totalmente viable, debido a que pueden disponerse de al menos tres conglomerados en cada clase de sujetos analizados, no solo por los valores de ganancia obtenidos mediante software, sino por la disposición geográfica del espacio falconiano; esto es, resultaría práctico definir grupos de clasificación en áreas de concentración bajo los términos de desarrollo territorial que atiendan objetivos con particularidades comunes en CTI.
- De acuerdo a lo anterior, la apreciación que los Investigadores pudieren tener de los clústeres guardaría especial interés con el apoyo ofrecido a través de recursos financieros y materiales para sus trabajos de investigación; no así, para los Innovadores quienes ven en la co-participación un objetivo de alto valor, jugando un papel determinante dentro de sus obligaciones sociales.
- Si bien, la data sobre artesanos y tecnólogos populares se mostró prácticamente cohesionada, es posible diferenciarlos mediante una variable su ubicación tabular; pero con respecto a la clusterización, habría que redefinir espacios geográficos lo suficientemente diferenciados si se requiere una participación de cada ámbito, rubro u oficio particular; esto sería, colocar clústeres para artesanos y otros para tecnólogos populares. De todas formas, la participación conjunta tiene sentido en cuanto a una hipotética visión de desarrollo que involucre a los Innovadores en el contexto de su radio de acción.
- Por otra parte, los focos de atención por clúster se concentran entre los municipios Miranda y Carirubana para los Investigadores, lo cual tiene sentido desde el punto de vista del número

de instituciones y de la población residente en dicho eje; no siendo así en los Innovadores ya que los sujetos estudiados se encontraron más explayados geográficamente, resultando favorecida la zona oriental de la geografía falconiana.

Finalmente, la consideración de clúster ha sido interesante para abordar decisiones estratégicas en políticas públicas, tratándose de capacidades tecnológicas de desarrollo, como lo expresa [Pérez et al., 2017]. La realización de este estudio ha servido de test relevante para la planificación de ejes directrices, a expensas de las limitaciones encontradas de todo análisis multivariante [Blázquez y García, 2009], pues se precisa en intentos por contextualizarle analíticamente para conformar nuevos mapas estratégicos en CTI. Además, ante nuevos escenarios (como el del PNIT), los resultados sirven para potenciar las capacidades de innovación, tal como lo manifiesta [García et al., 2012].

Agradecimientos

Al Comité Académico del Programa Doctoral en Planificación y Gestión del Desarrollo Regional de la Universidad del Zulia, Núcleo Punto Fijo, por la aprobación de la Investigación Libre: “Análisis de datos multivariantes en ciencia, tecnología e innovación del estado Falcón”.

Bibliografía

- [Anderberg, 1973] Anderberg, M. (1973). Cluster analysis for applications. *Academic Press*, Inc. New York.
- [Blázquez y García, 2009] Blázquez, M. y García, M. (2009). *Clusters de innovación tecnológica en Latinoamérica*. GCG Georgetown University. Vol. 3, Nro. 3. pp. 16-33. DOI:10.3232/GCG.2009.V3.N3.01. <https://dialnet.unirioja.es/servlet/articulo?codigo=3116622>
- [Closas et al., 2013] Closas, A., Arriola, E., Kuc, C., Amarilla, M. y Jovanovich, E. (2013). Análisis multivariante, conceptos y aplicaciones en psicología educativa y psicometría. *Enfoques*. Vol. XXV, Nro. 1. Otoño. pp. 65-92. <https://www.redalyc.org/articulo.oa?id=25930006005>
- [Cuadras, 2014] Cuadras, C. M. (2014). *Nuevos métodos de análisis multivariante*. Ediciones CMC (CMC éditions Sàrl). Barcelona.
- [Constitución, 2009] CRBV. (2009). *Constitución de la República Bolivariana de Venezuela*. Enmienda Nro. 1: 15/02/2009. G.O. Nro. 5.908 Ext. Asamblea Nacional. (19/02/2009). Caracas.
- [García et al., 2012] García, M., Blázquez, M. y López, J. (2012). Uso y aplicación de la técnica de análisis estadístico multivariante de clúster sobre la capacidad

- de innovación tecnológica en Latinoamérica y España. *Revista Innovar*. Vol. 22, Nro. 44. Abril-Junio. pp. 21-40. <https://revistas.unal.edu.co/index.php/innovar/article/view/35564/36015>
- [Gutiérrez, 2016] Gutiérrez, A. (2016). La investigación e innovación hoy en Venezuela. *Revista Venezolana de Endocrinología y Metabolismo*. Vol. 14, Nro. 3. Octubre. Editorial. pp. 166-167. <http://www.saber.ula.ve/handle/123456789/42685?locale-attribute=fr>
- [Husson y Josse, 2010] Husson, F. y Josse, J. (2010). *Multivariate data analysis: special focus on clustering and multiway methods. Tutorial*. Applied Mathematics Department. Agrocampus Rennes. <https://www.r-project.org/conferences/user-2010/tutorials/Husson+Josse.pdf>
- [Husson et al., 2011] Husson, F., Lê, S. y Pagès, J. (2011). *Exploratory multivariate analysis by example using R*. CRC Press. Taylor & Francis. Boca Raton, FL, USA.
- [Kerlinger y Lee, 2002] Kerlinger, F. y Lee, H. (2002). *Investigación del comportamiento: métodos de investigación en ciencias sociales*. McGraw-Hill. 4ta Edición. México.
- [MPPCTII, 2010] MPPCTII. (2010). *Ley orgánica de ciencia, tecnología e innovación (LOCTI)*. FONACIT. Caracas.
- [ONCTI, 2016] ONCTI. (2016). Indicadores venezolanos de ciencia, tecnología e innovación. *Boletín Año 2016*. Ediciones ONCTI. Caracas. <http://otrasvoceseneducacion.org/archivos/251967>
- [Padrón, 2001] Padrón, J. (2001). *La estructura de los procesos de investigación. Educación y Ciencias Humanas*. Año XI, Nro. 17. UNESR. Jul-Dic. pp. 33-54. Caracas. http://padron.entretemas.com.ve/Estr_Proc_Inv.htm
- [Peña, 2002] Peña, D. (2002). *Análisis de datos multivariantes*. Editorial McGraw-Hill / Interamericana de España, SAU. Madrid.
- [Pérez et al., 2017] Pérez, C., Lara, G. y Gómez, D. (2017). Evolución de la capacidad tecnológica en México. Aplicación del análisis estadístico multivariante clúster. *Contaduría y Administración*. Nro. 62. pp. 505-527. <http://dx.doi.org/10.1016/j.cya.2017.01.002>. <http://www.cya.unam.mx/index.php/cya/article/view/966/898>
- [Requena y Caputo, 2016] Requena, J. y Caputo, C. (2016). Pérdida de talento en Venezuela: migración de sus investigadores. *Interciencia*. Vol. 41, Nro. 7. Julio. pp. 444-453. <https://dialnet.unirioja.es/servlet/articulo?codigo=5601420>
- [Tinsley y Brown, 2000] Tinsley, H. y Brown, S. (2000). *Handbook of applied multivariate statistics and mathematical modeling*. Editores. Academic Press. San Diego.

- [Van Noorden, 2014] Van Noorden, R. (2014). The impact gap: South America by the numbers. *Nature - International Weekly Journal of Science*. Volume 510, Issue 7504. 11/06/2014. Nature. 12/06/2014. Vol. 510. pp. 202-203. <https://wcrif.org/2015-news/51-the-impact-gap-south-america-by-the-numbers>.
- [Zamora et al., 2014] Zamora, F., Pérez, R. y Núñez, M. (2014). Falcón cuenta con 627 investigadores y 73 innovadores acreditados. *Ciencia y Tecnología para el Pueblo*. Fundacite Falcón. Edición impresa. Año 3. Nro. 5. Coro.