

METODOLOGIA DE CARACTERIZACIÓN Y AGRUPAMIENTO DE LOS SISTEMAS GANADEROS DEL DEPARTAMENTO DE ARAUCA

INTRODUCCIÓN

El alto grado de heterogeneidad que existe entre los sistemas ganaderos que conforman una región dificulta el desarrollo y la validación de modelos alternativos de gestión tecnológica para finalmente tomar decisiones. En tal sentido al agrupar las explotaciones de acuerdo a sus principales diferencias y relaciones, se busca maximizar las semejanzas dentro de los grupos y las diferencias entre los mismos.

De esta forma la caracterización no es más que la descripción de las características principales y las múltiples interrelaciones de las organizaciones, mientras que el agrupamiento se refiere al establecimiento y construcción de grupos posibles basados en las características observadas en la realidad. La información obtenida en este análisis es considerada de gran utilidad a fin de proponer estrategias que permitan mejorar los aspectos de mayor incidencia en el desarrollo de las empresas ganaderas registradas.

Para ello se aplicaron técnicas de análisis multivariante como el análisis de componentes principales y el análisis de clúster. Además se interpretaron los resultados obtenidos por medio de la estadística descriptiva.

OBJETIVO

Determinar las principales características que inciden en el grado la heterogeneidad y homogeneidad existente entre las explotaciones ganaderas del departamento de Arauca; generando grupos representativos o subsistemas productivos, mediante la aplicación de técnicas de análisis multivariante.

METODOLOGIA Y RESULTADOS

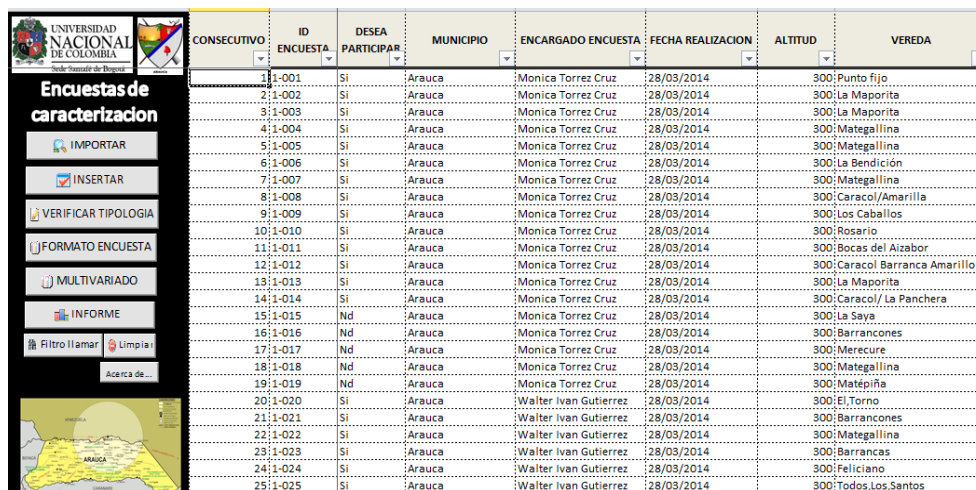
La metodología utilizada se estructura en las siguientes etapas:

1. Construcción del instrumento de recolección de la información y procesamiento de la información (elaboración de la base de datos, clasificación y descripción de las variables).
2. Filtro de registros con información completa.
3. Revisión y selección de las variables.
4. Aplicación de técnicas estadísticas multivariantes (ACP).
5. Determinación de tipos o subsistemas (clusters).
6. Descripción de los tipos o grupos.

1. Instrumento de recolección y procesamiento de información

Se construyó un aplicativo (EncuestSoft 1.0.19) para la digitación y el procesamiento de la información recolectada por medio de encuestas personalizadas (imagen 1).

Imagen 1. Interfaz del aplicativo



CONSECUTIVO	ID ENCUESTA	DESEA PARTICIPAR	MUNICIPIO	ENCARGADO ENCUESTA	FECHA REALIZACION	ALTITUD	VEREDA
1	1-001	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Punto fijo
2	1-002	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	La Maporita
3	1-003	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	La Maporita
4	1-004	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Mategallina
5	1-005	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Mategallina
6	1-006	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	La Bendición
7	1-007	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Mategallina
8	1-008	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Caracol/Amarilla
9	1-009	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Los Caballos
10	1-010	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Rosario
11	1-011	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Bocas del Aizabor
12	1-012	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Caracol Barranca Amarillo
13	1-013	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	La Maporita
14	1-014	SI	Arauca	Monica Torrez Cruz	28/03/2014	300	Caracol/ La Panchera
15	1-015	Nd	Arauca	Monica Torrez Cruz	28/03/2014	300	La Saye
16	1-016	Nd	Arauca	Monica Torrez Cruz	28/03/2014	300	Barrancones
17	1-017	Nd	Arauca	Monica Torrez Cruz	28/03/2014	300	Merecure
18	1-018	Nd	Arauca	Monica Torrez Cruz	28/03/2014	300	Mategallina
19	1-019	Nd	Arauca	Monica Torrez Cruz	28/03/2014	300	Matepiña
20	1-020	SI	Arauca	Walter Ivan Gutierrez	28/03/2014	300	El Torno
21	1-021	SI	Arauca	Walter Ivan Gutierrez	28/03/2014	300	Barrancones
22	1-022	SI	Arauca	Walter Ivan Gutierrez	28/03/2014	300	Mategallina
23	1-023	SI	Arauca	Walter Ivan Gutierrez	28/03/2014	300	Barrancas
24	1-024	SI	Arauca	Walter Ivan Gutierrez	28/03/2014	300	Feliciano
25	1-025	SI	Arauca	Walter Ivan Gutierrez	28/03/2014	300	Todos, Los Santos

La base de datos se depuró (estandarización información) y a partir de esta se generó un informe técnico con las características de los productores y la representación gráfica del comportamiento de todas las variables registradas.

2. Filtro de registros con información completa

Se realizó un filtro para identificar fincas con información básica completa (tabla 1) para generar diferentes indicadores que son usados en los análisis estadísticos y la generación de conglomerados, de esta manera se pasó de tener 566 a 241 registros.

Tabla 1. Variables básicas usadas para filtrar fincas que pueden conformar clusters.

Tiempo desplazamiento (min)	Área	Área pastoreo	Escolaridad	Rango edad	Rango trabajadores en la finca	Años de producción	Inventario total
-----------------------------	------	---------------	-------------	------------	--------------------------------	--------------------	------------------

Es importante destacar que este proceso se realizara a profundidad y para el total de fincas registradas en el momento que se valide en campo la información suministrada por el productor en la primera encuesta. En tal caso no se necesitaría realizar el filtro anterior ya que la información estaría completa.

3. Revisión y selección de las variables.

Se identificaron 35 variables algunas de carácter netamente productivo , sin embargo, debido a la falta de información completa, solo a 21 variables y/o indicadores se les realizó un análisis de correlación (tabla 2), con el fin de eliminar variables altamente correlacionadas ($P < 0.05$).

Tabla 2. Matriz de correlación

	UFS	UFS_HA	MU_AÑO	ESCOLARIDAD	AREA	AREA_PASTOS	ANIMALES	ASTORES	CARGA_ANIMAL	DIST_TIEMPO	EDAD	ADORES_PUBLICOS	INFRAESTRUCTURA	SUPLEMENTACION	SISTEMA_REGISTRO	FINANCIACIÓN	PRODUCCIÓN	MONTA	CARÁCTER_EMPRESA	PRODUCCIÓN
UFS	1,00																			
UFS_HA	-0,10	1,00																		
MU_AÑO	1,00	-0,10	1,00																	
ESCOLARIDAD	0,12	-0,21	0,12	1,00																
AREA	0,69	-0,36	0,69	0,12	1,00															
AREA_PASTOS	0,58	-0,34	0,58	0,07	0,96	1,00														
TOTAL_ANIMALES	1,00	-0,10	1,00	0,12	0,69	0,58	1,00													
PORC_PASTOREO	-0,08	0,10	-0,08	-0,10	-0,08	0,06	-0,08	1,00												
CARGA_ANIMAL	0,01	0,37	0,01	-0,08	-0,08	-0,10	0,01	-0,57	1,00											
DIST_TIEMPO	0,01	-0,14	0,01	-0,07	0,15	0,17	0,01	0,03	-0,03	1,00										
EDAD	0,06	0,07	0,06	-0,35	-0,02	-0,02	0,06	-0,02	0,06	-0,09	1,00									
TRABAJADORES_FINCA	0,32	0,02	0,32	-0,07	0,13	0,12	0,32	-0,02	0,00	0,07	0,19	1,00								
SERVICIOS_PUBLICOS	0,03	0,17	0,03	0,03	-0,06	-0,05	0,03	0,05	0,00	-0,05	0,00	0,05	1,00							
INFRAESTRUCTURA	0,17	0,07	0,17	0,15	-0,01	-0,04	0,17	-0,03	-0,03	-0,05	-0,07	0,14	0,52	1,00						
SUPLEMENTACION	0,02	0,15	0,02	0,13	-0,07	-0,09	0,02	-0,06	0,03	-0,05	-0,13	0,07	0,48	0,85	1,00					
SISTEMA_REGISTRO	0,16	0,04	0,16	0,07	0,12	0,13	0,16	-0,02	0,03	-0,08	-0,06	0,04	0,12	0,11	0,11	1,00				
FUENTE_FINANCIACIÓN	0,11	0,04	0,11	0,07	0,05	0,07	0,11	0,20	-0,14	0,01	0,06	0,10	0,20	0,10	0,04	-0,06	1,00			
AÑOS_PRODUCIÓN	0,16	0,14	0,16	-0,17	0,05	0,07	0,16	0,05	0,09	0,01	0,41	0,19	0,16	0,02	-0,06	0,05	0,17	1,00		
MONTA	0,09	0,12	0,09	0,05	-0,04	-0,02	0,09	0,01	0,08	-0,05	-0,02	0,15	0,15	0,19	0,12	0,22	0,09	0,00	1,00	
CARÁCTER_EMPRESA	0,13	0,08	0,13	0,07	-0,02	-0,02	0,13	-0,02	0,06	-0,07	-0,09	0,00	0,33	0,30	0,29	0,17	0,06	-0,06	0,21	1,00
HEMBRAS_PRODUCIÓN	0,63	0,04	0,63	0,02	0,33	0,30	0,63	-0,02	0,06	-0,10	0,17	0,32	0,16	0,26	0,12	0,14	0,12	0,23	0,33	0,08

De esta manera se escogieron 18 variables para el análisis de componentes principales (tabla 3).

Tabla 3. Variables utilizadas para el ACP (análisis de componentes principales).

1.UFS_HA	2.DIST_TIEMPO	3.SISTEMA_REGISTROS
4.ESCOLARIDAD	5.EDAD	6. FUENTE FINANCIACIÓN
7.AREA_PASTOS	8.TRABAJADORES_FINCA	9.AÑOS_PRODUCCION
10.TOTAL_ANIMALES	11.SERVICIOS_PUBLICOS	12.MONTA
13.PORC_PASTOREO	14.INFRAESTRUCTURA	15.CARÁCTER_EMPRESA
16.CARGA_ANIMAL	17.SUPLEMENTACION	18.HEMBRAS PRODUCCION

4. Aplicación de técnicas estadísticas multivariantes (ACP)

El Análisis de Componentes Principales (ACP) es una técnica estadística de síntesis de la información, o reducción de la dimensión (número de variables). Es decir, ante un banco de datos con muchas variables, el objetivo será reducirlas a un menor número perdiendo la menor cantidad de información posible.

Para este caso se usó el paquete estadístico INFOSTAT.

La tabla 4 presenta los 8 componentes elegidos, los cuales explicaron el 73% de la varianza total.

Tabla 4. Valor de cada componente con la varianza propia y acumulada (autovalores)

<u>Lambda</u>	<u>Valor</u>	<u>Proporción</u>	<u>Prop Acum</u>
1	2,96	0,16	0,16
2	2,29	0,13	0,29
3	2,05	0,11	0,41
4	1,69	0,09	0,50
5	1,21	0,07	0,57
6	1,08	0,06	0,63
7	0,95	0,05	0,68
8	0,87	0,05	0,73
9	0,82	0,05	0,77
10	0,76	0,04	0,82
11	0,71	0,04	0,86
12	0,59	0,03	0,89
13	0,54	0,03	0,92
14	0,44	0,02	0,94
15	0,43	0,02	0,97
16	0,28	0,02	0,98
17	0,20	0,01	0,99
18	0,11	0,01	1,00

La tabla 5 muestra el peso de cada una de las variables dentro de cada componente.

Tabla 5. Correlación de cada variable sobre los componentes elegidos (autovectores)

Variables	e1	e2	e3	e4	e5	e6	e7	e8
UFS_HA	0,10	-0,26	0,38	-0,04	0,18	-0,15	0,14	0,53
ESCOLARIDAD	0,08	-0,02	-0,39	-0,08	0,19	0,53	0,01	-0,14
AREA_PASTOS	0,12	0,47	-0,21	-0,03	-0,03	-0,04	-0,27	0,21
TOTAL_ANIMALES	0,29	0,45	-0,08	-0,13	-6,5E-05	0,06	-0,03	0,23
PORC_PASTOREO	-0,01	0,04	-0,05	0,63	0,17	-0,28	0,08	0,21
CARGA_ANIMAL	0,04	-0,11	0,28	-0,59	0,02	0,11	-0,05	0,28
DIST_TIEMPO	-0,07	0,19	-0,28	-0,13	-0,31	-0,42	0,12	0,28
EDAD	0,03	0,17	0,46	0,10	-0,25	-0,06	-0,12	-0,33
TRABAJADORES_FINCA	0,22	0,23	0,16	-3,3E-03	-0,11	-0,04	0,50	-0,22
SERVICIOS_PUBLICOS	0,37	-0,24	0,03	0,16	-0,09	0,05	-0,27	0,01
INFRAESTRUCTURA	0,45	-0,22	-0,15	0,03	-0,33	-0,04	0,09	-0,07
SUPLEMENTACION	0,39	-0,31	-0,16	-0,01	-0,34	-0,08	0,07	-0,02
SISTEMA_REGISTRO	0,18	0,05	-0,07	-0,12	0,40	-0,41	-0,40	-0,27
FUENTE_FINANCIACIÓN	0,16	0,06	0,06	0,37	0,12	0,43	0,01	0,32
AÑOS_PRODUCION	0,14	0,17	0,41	0,13	-0,13	0,11	-0,34	-0,08
MONTA	0,25	-0,02	0,05	-0,06	0,51	-0,14	0,38	-0,22
CARÁCTER_EMPRESA	0,27	-0,17	-0,10	-0,04	0,19	-0,13	-0,29	0,09
HEMBRAS_PRODUCION	0,36	0,32	0,10	-0,06	0,10	0,06	0,18	0,03

De esta manera se agruparon las correspondientes variables dentro de cada componente y se les asignó una explicación lógica e interpretable (tabla 6).

Tabla 6. Explicación de cada componente elegido y sus variables

COMPONENTE	VARIABLES	VARIANZA ACUMULADA
1-Poder adquisitivo	SERVICIOS_PUBLICOS	0,16
	INFRAESTRUCTURA	
	SUPLEMENTACION	
2-Tamaño finca	AREA_PASTOS	0,29
	TOTAL_ANIMALES	
	HEMBRAS_PRODUCION	
3-Experiencia	EDAD	0,41
	AÑOS_PRODUCION	
4-Uso del suelo	PORC_PASTOREO	0,5
	CARGA_ANIMAL	
5- Estado reprod	MONTA	0,57
6-Nivel especialización y ubicación estratégica	ESCOLARIDAD	0,63
	DIST_TIEMPO	
	SISTEMA_REGISTRO	
	FUENTE_FINANCIACIÓN	
7-Empresarizacion	TRABAJADORES_FINCA	0,68
	CARÁCTER_EMPRESA	
8- Capacidad de carga	UFS_HA	0,73

5. Determinación de tipos o subsistemas (clusters).

Una vez se seleccionaron los factores (componentes) se procedió al análisis multivariante de clúster, el cual es un método estadístico de clasificación de datos, que permite establecer grupos de explotaciones.

El método de agrupamiento más utilizado es el de Ward ya que maximiza la homogeneidad dentro de los grupos y plantea todas las posibles fusiones entre las variables, por esto se aplicó tal método y se implementó la distancia euclídea como medida de agrupamiento jerárquica.

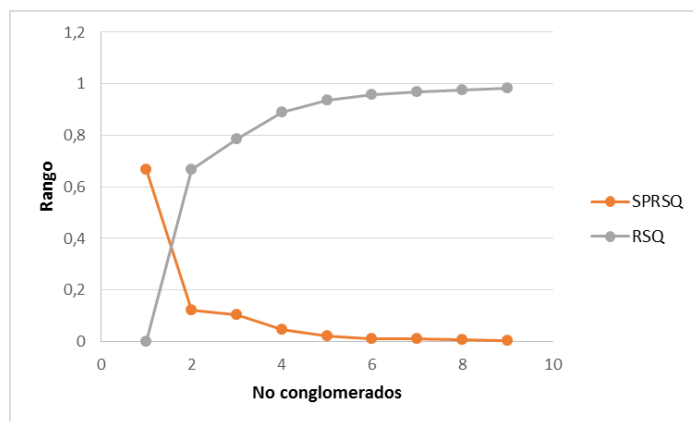
En la tabla 7, se observa la historia de conglomerado usando el paquete estadístico SAS. Es importante destacar que el alto valor del RSQ (0,936) y el bajo SPRSQ (0,0218) reducido proporcionalmente en el conglomerado 5 permite hacer una aproximación para detener el agrupamiento en tal punto y componer 5 clusters.

Tabla 7. Historia de conglomerado usando el programa estadístico SAS.

NCL	Conglomerados und		Frecuencia	SPRSQ	RSQ	PSF	PST2
	CL15	CL18	150633	0,0032	0,986	1839	271
9	CL24	CL14	325206	0,0034	0,983	1662	73,1
8	CL13	OB180	1070749	0,0054	0,977	1443	7
7	CL11	CL17	625343	0,009	0,968	1196	38
6	CL10	CL16	222439	0,0104	0,958	1072	321
5	CL9	CL12	476494	0,0218	0,936	866	191
4	CL7	CL8	1291648	0,0478	0,888	629	46,4
3	CL5	CL6	577225	0,1023	0,786	437	502
2	CL4	OB200	2105888	0,1206	0,666	476	35,8
1	CL3	CL2	1437148	0,6656	0	0	476

La grafica 1 permite hacer una aproximación más exacta al número de clusters ya que se observa una inflexión notable en el conglomerado 5.

Grafica 1. Representación gráfica del SPRS1 y RSQ en la historia de conglomeración.



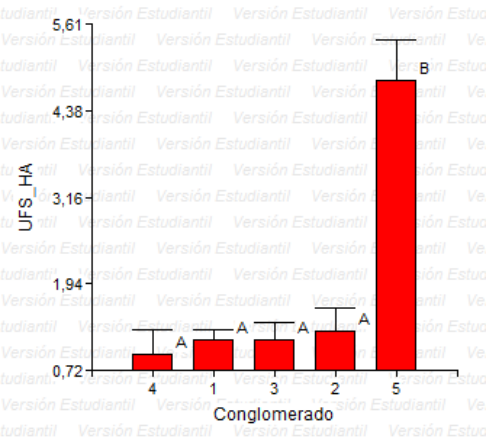
Grafica 2. Dendrograma y representación de los 5 clusters seleccionados.



En la tabla 9 y en la gráfica 3 se puede observar el comportamiento de la variable UF/ha por conglomerado.

CONGLOMERADO	Promedio UFS_HA	Desvest
1	1,15	0,67
2	1,28	0,52
3	1,16	0,60
4	0,94	1,15
5	4,81	8,59
Total general	1,27	1,75

Grafica 3. Representación de las UF/ha para cada conglomerado

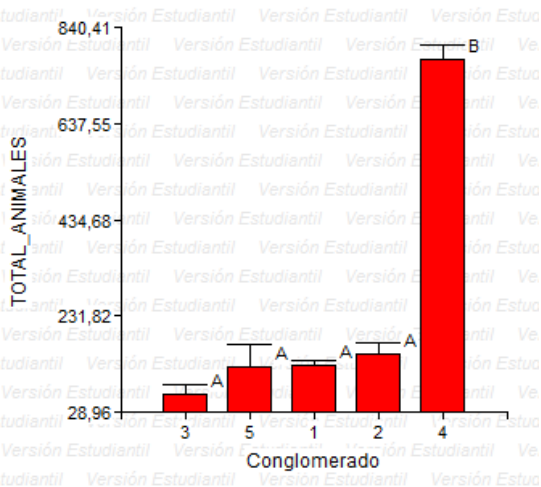


Las fincas del conglomerado 5 presentan una mayor carga animal que las de los demás clúster, así mismo el clúster 4 se caracteriza por presentar bajas UF/ha.

Tabla 9. Inventario total para cada conglomerado

CONGLOMERADO	Promedio TOTAL_ANIMALES	Desvest
1	127,07	97,95
2	149,56	145,80
3	65,84	63,13
4	772,95	355,20
5	122,63	104,24
Total general	173,02	233,51

Grafica 4. Representación del promedio de inventario por conglomerado

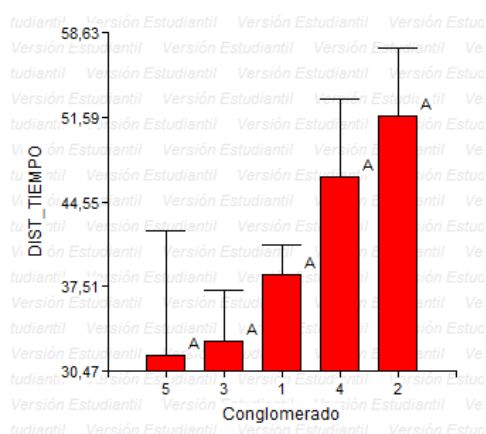


El conglomerado 4 se caracteriza por fincas con alto número de animales a diferencia de los demás conglomerados. Por otra parte el conglomerado 3 tiene una tendencia a manejar pocos animales.

Tabla 10. Promedio de distancia en minutos para cada conglomerado

CONGLOMERADO	Promedio DIST_TIEMPO	Desvest
1	38,45	30,23
2	51,67	34,78
3	32,98	21,87
4	46,67	33,25
5	31,75	30,36
Total general	39,29	29,84

Grafica 5. Representación del promedio de distancia en minutos por conglomerado

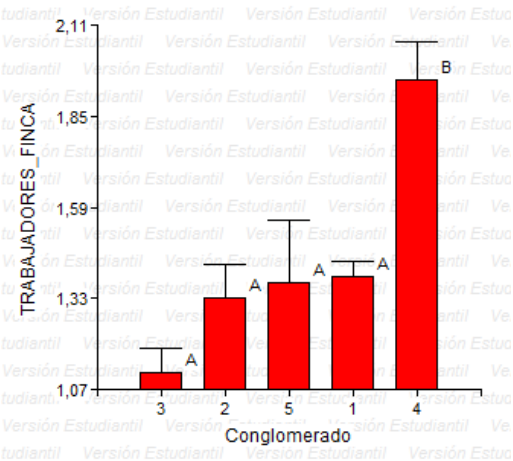


No existen diferencias significativas entre clusters, sin embargo, existen una tendencia marcada en las fincas del conglomerado 2 caracterizadas por estar más alejadas del centro urbano. El conglomerado 5 y 3 lo comprenden fincas más cercanas a los centros urbanos aunque se presenta una alta dispersión en los datos.

Tabla 11. Promedio trabajadores para cada conglomerado (1=1 a 3, 2= 3 a 7, 3= >7)

CONGLOMERADO	Promedio TRABAJADORES_FINCA	Desvest
1	1,39	0,52
2	1,33	0,55
3	1,12	0,33
4	1,95	0,67
5	1,38	0,52
Total general	1,38	0,54

Grafica 6. Representación del promedio de trabajadores por conglomerado

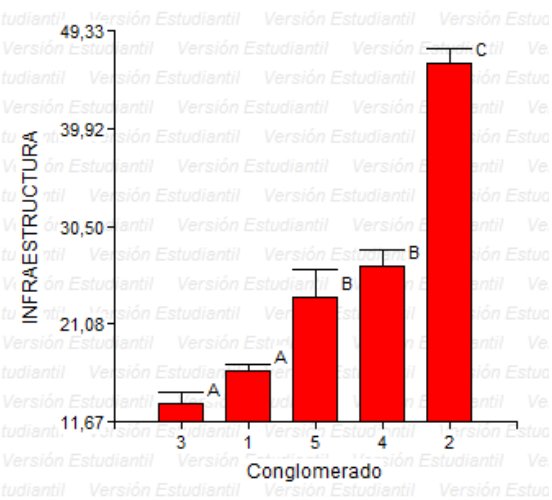


El conglomerado 4 presenta fincas con mayor disponibilidad de mano de obra contrariamente a los demás conglomerados. El conglomerado 3 tiene una tendencia a manejar poca mano de obra.

Tabla 12. Promedio del índice de infraestructura para cada conglomerado

CONGLOMERADO	Promedio INFRAESTRUCTURA	Desvest
1	16,48	6,62
2	46,19	4,23
3	13,38	5,75
4	26,57	10,74
5	23,63	20,16
Total general	20,28	12,33

Grafica 7. Representación del promedio del índice de infraestructura por conglomerado (>valor= mayor disponibilidad de infraestructura)

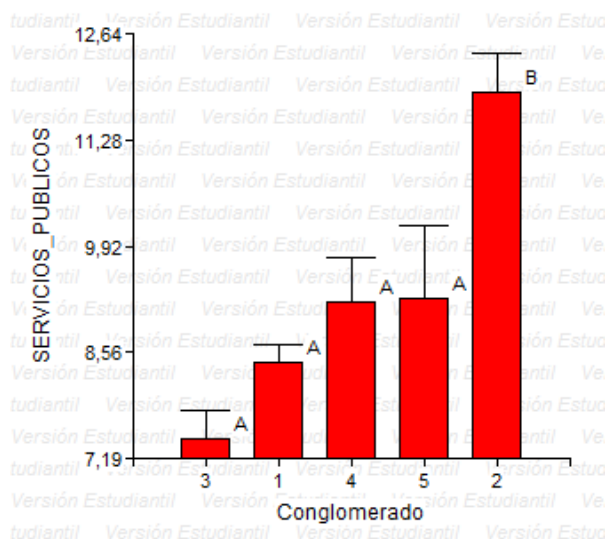


Las fincas del conglomerado 2 disponen de mayor infraestructura (corrales, bodegas, bebederos, casa, establos, etc.) que los demás conglomerados. Contrariamente sucede con el conglomerado 3 y 1 que lo componen fincas de menor disponibilidad de infraestructura.

Tabla 13. Promedio del índice de servicios públicos para cada conglomerado

CONGLOMERADO	Promedio SERVICIOS_PUBLICOS	Desvest
1	8,43	2,72
2	11,89	0,58
3	7,44	2,95
4	9,19	2,58
5	9,25	3,49
Total general	8,71	2,90

Grafica 8. Representación del promedio del índice de servicios públicos por conglomerado

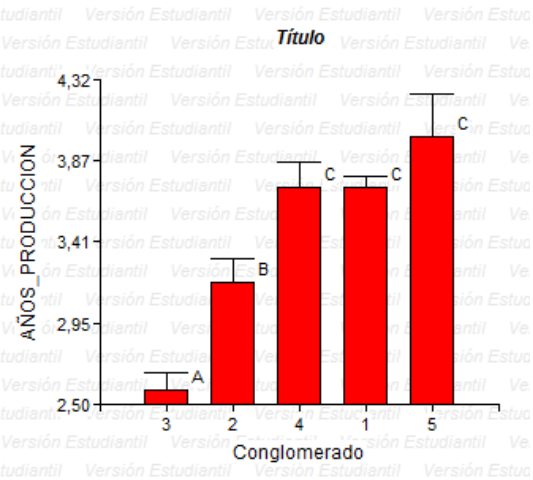


Las fincas del conglomerado 2 además de tener más infraestructura poseen mayor acceso a servicios públicos contrariamente a las fincas del conglomerado 3.

Tabla 14. Promedio de años de experiencia para cada conglomerado

CONGLOMERADO	Promedio AÑOS_PRODUCCION	Desvest
1	3,72	0,61
2	3,19	0,92
3	2,58	0,81
4	3,71	0,56
5	4,00	0,00
Total general	3,43	0,82

Grafica 9. Representación del promedio de años de experiencia por conglomerado

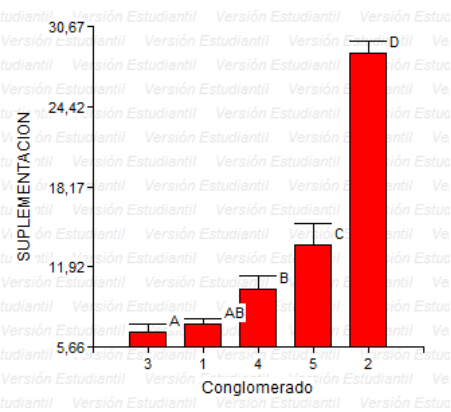


Las fincas del conglomerado 3 se caracterizan por ser productores jóvenes de poca experiencia seguidos por las fincas del conglomerado 2.

Tabla 15. Promedio del índice de suplementación para cada conglomerado

CONGLOMERADO	Promedio SUPLEMENTACION	Desvest
1	7,44	3,89
2	28,63	1,92
3	6,80	3,24
4	10,19	8,47
5	13,63	12,87
Total general	10,13	8,19

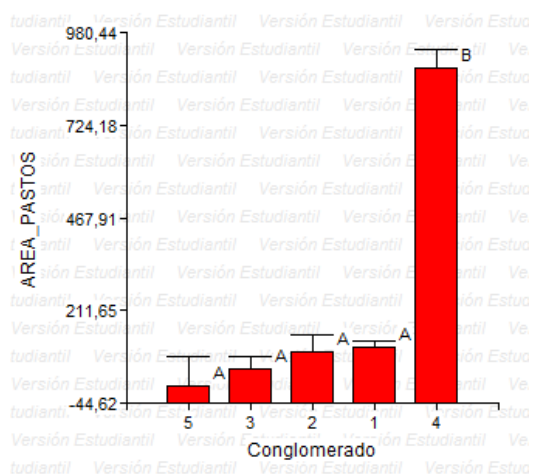
Grafica 10. Representación del promedio del índice de suplementación por conglomerado



Las fincas del conglomerado 2 además de tener más infraestructura poseen mayores recursos para suplementar a los animales contrariamente a las fincas del conglomerado 3.

Así mismo en la gráfica 9 se observan los valores promedios de la variable área de pastos relacionada con el tamaño de las fincas y la disponibilidad de forraje en cada uno de los conglomerados.

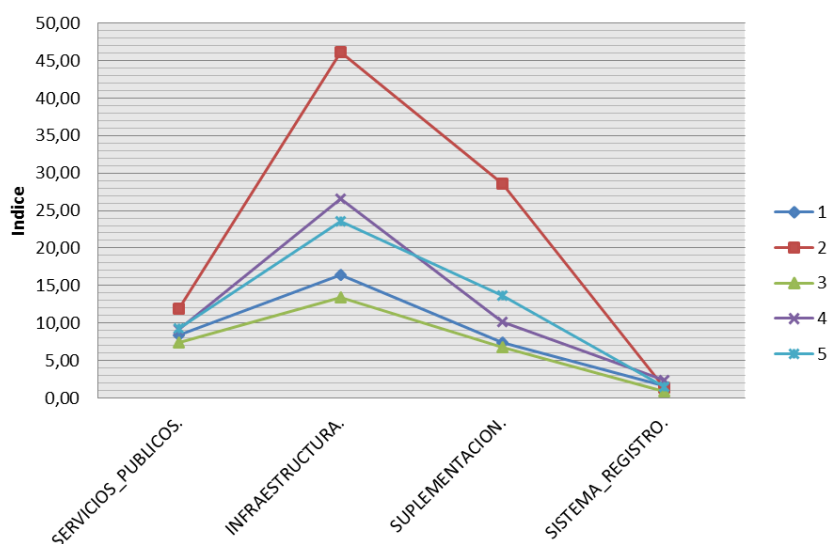
Gráfica 11. Representación del promedio del área de pasturas por conglomerado



De esta manera se observa que el conglomerado 4 abarca fincas de mayor área a diferencia principalmente del conglomerado 5.

Finalmente en la gráfica 10 se pueden ver los valores promedios de otras variables analizadas para cada uno de los conglomerados.

Gráfica 12. Resumen de conglomerados



El conglomerado 2 se destaca por comprender fincas con potencial de acceso a servicios públicos, infraestructura y suplementación. Por otra parte los conglomerados 3 y 1 lo comprenden fincas con menores capacidades de acceso a infraestructura y servicios públicos.

Finalmente se resumen los diferentes conglomerados

Conglomerado 1= Productores de alta experiencia, baja infraestructura, baja suplementación, fincas promedio.

Conglomerado 2= Productores de poca experiencia, alta infraestructura, alta suplementación, gran acceso a servicios públicos.

Conglomerado 3= Fincas de poca área, pocos animales, productores con poca experiencia, poco acceso a infraestructura, baja suplementación, limitada mano de obra, poco acceso a servicios públicos.

Conglomerado 4= Fincas con baja carga animal, altas áreas, altos inventarios, alta mano de obra.

Conglomerado 5= Fincas con alta carga animal, poca escolaridad, poca área, pocos animales, limitado uso del suelo (bajo pastoreo), cercanas al centro urbano.