

## Análisis de la Superficie Cultivada y Producción de Papa en el Altiplano Paceño con Escalamiento Multidimensional y Análisis Cluster

*Autor: Lic. Fernando Rivero Sugiura*

### 1. Introducción

El escalado multidimensional tiene el objetivo de identificar una representación gráfica perceptual de un grupo de objetos, mediante las distancias de similitud entre ellos. El método multivariado cluster, reagrupa a los objetos estudiados de acuerdo a su similitud empleando determinadas variables cuantitativas.

El presente tiene el objetivo de aplicar ambos métodos de clasificación con base algunos datos estadísticos de la superficie promedio cultivada en hectáreas y la producción promedio del producto

agrícola “papa” en quintales por municipio. Esta información es obtenida de la Encuesta Sociodemográfica Altiplano Sur y Norte<sup>1</sup> de los 32 municipios estudiados. La producción de papa es lo más común en todos estos municipios del Altiplano de La Paz, por lo cual todos los municipios cuentan con información.

### 2. Metodología

Los datos de superficie cultivada en Ha, y producción en qq en el periodo de un año por municipio, se presenta en el siguiente cuadro:

Cuadro 1

Región	Nro.	Municipio	Superficie Cultivada (Ha)	Producción Cosechada (qq)
Altiplano Norte	1	Puerto Acosta	200,8	20.620,5
	2	Puerto Carabuco	234,9	22.536,2
	3	Escoma	227,0	20.990,5
	4	Achacachi	882,4	92.816,5
	5	Ancoraimes	12,8	1.237,0
	6	Santiago de Huata	55,3	6.464,2
	7	Chua Cocani	23,9	1.896,5
	8	Huarina	362,0	31.631,0
	9	Huatajata	57,3	5.985,4
	10	Tiquina	3,5	288,0
	11	Tito Yupanqui	9,8	917,9
	12	Copacabana	255,9	25.103,9
	13	Pucarani	859,4	85.392,2
	14	Batallas	568,7	61.551,7
	15	Puerto Pérez	73,9	6.996,1
	16	Tiahuanacu	538,2	55.618,0
	17	Taraco	90,9	8.770,8
	18	Guaqui	156,4	15.560,6
	19	Desaguadero	55,3	5.717,7
Altiplano Sur	20	Coro Coro	809,5	51.248,8
	21	Caquiaviri	604,3	49.571,7
	22	Calacoto	452,2	37.772,2
	23	Comanche	100,5	6.577,2
	24	Nazacara de Pacajes	11,5	980,2
	25	Santiago de Callapa	503,0	42.606,7
	26	Jesús de Machaca	256,3	28.207,0
	27	Sica Sica	1.135,8	90.730,2
	28	Umala	1.101,9	95.653,9
	29	Calamarca	872,3	69.980,8
	30	Patacamaya	476,5	32.275,2
	31	Colquencha	539,8	44.366,8
	32	Collana	111,8	8.956,6

Fuente: Encuesta Sociodemográfica - IETA

1 - Información que procesa el IETA

Los detalles metodológicos del análisis de escalamiento multidimensional se encuentran en la Revista Nro. 8, donde el autor realizó un ejercicio de distancias entre las ciudades capitales de Bolivia. En este artículo se presentan los resultados y análisis de una aplicación agrícola descrita anteriormente.

La siguiente es la matriz de distancias euclídeas  $D$  de dimensión 32, es decir:

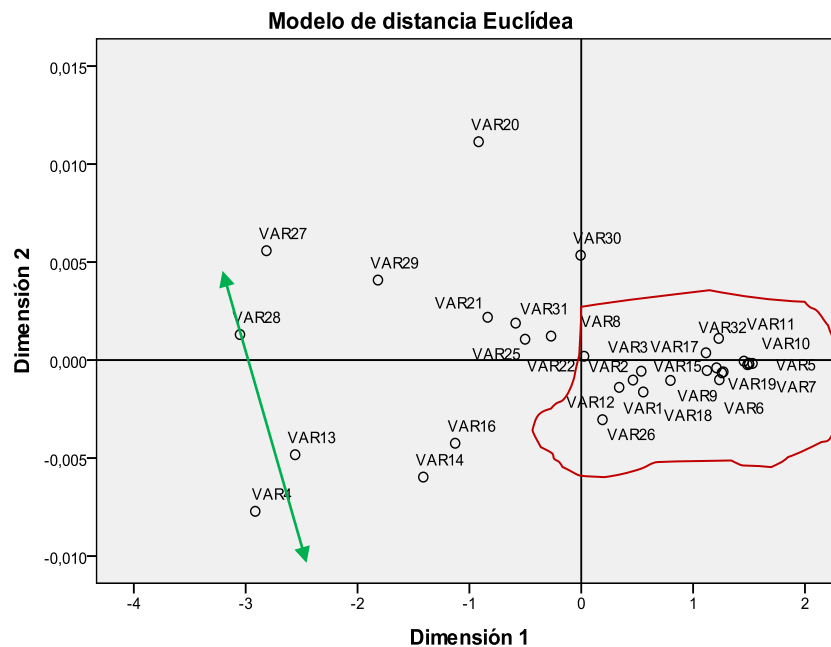
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32				
1	0																																			
2	0,09	0																																		
3	0,02	0,07	0																																	
4	3,47	3,38	3,45	0																																
5	0,93	1,02	0,95	4,40	0																															
6	0,68	0,77	0,70	4,15	0,25	0																														
7	0,90	0,99	0,92	4,37	0,03	0,22	0																													
8	0,53	0,44	0,51	2,94	1,46	1,21	1,43	0																												
9	0,70	0,80	0,72	4,17	0,23	0,02	0,20	1,23	0																											
10	0,98	1,07	1,00	4,45	0,05	0,30	0,08	1,51	0,27	0																										
11	0,95	1,04	0,96	4,42	0,02	0,27	0,05	1,48	0,24	0,03	0																									
12	0,22	0,12	0,20	3,25	1,15	0,90	1,12	0,31	0,92	1,19	1,16	0																								
13	3,11	3,02	3,09	0,36	4,04	3,79	4,01	2,58	3,82	4,09	4,06	2,90	0																							
14	1,97	1,88	1,95	1,50	2,90	2,65	2,87	1,44	2,67	2,94	2,91	1,75	1,15	0																						
15	0,66	0,75	0,67	4,12	0,28	0,03	0,25	1,18	0,05	0,32	0,29	0,87	3,77	2,62	0																					
16	1,68	1,59	1,66	1,79	2,61	2,36	2,58	1,15	2,39	2,66	2,63	1,47	1,43	0,29	2,34	0																				
17	0,57	0,66	0,59	4,04	0,36	0,11	0,33	1,10	0,13	0,41	0,38	0,79	3,68	2,54	0,09	2,25	0																			
18	0,24	0,34	0,26	3,71	0,69	0,44	0,66	0,77	0,46	0,73	0,70	0,46	3,36	2,21	0,41	1,93	0,33	0																		
19	0,72	0,81	0,73	4,19	0,22	0,04	0,18	1,25	0,01	0,26	0,23	0,93	3,83	2,68	0,06	2,40	0,15	0,47	0																	
20	1,47	1,38	1,45	2,00	2,40	2,15	2,37	0,94	2,18	2,45	2,42	1,26	1,64	0,50	2,13	0,21	2,04	1,72	2,19	0																
21	1,39	1,30	1,37	2,08	2,32	2,07	2,29	0,86	2,09	2,37	2,34	1,18	1,72	0,58	2,05	0,29	1,96	1,63	2,11	0,08	0															
22	0,82	0,73	0,81	2,65	1,76	1,50	1,72	0,30	1,53	1,80	1,77	0,61	2,29	1,14	1,48	0,86	1,39	1,07	1,54	0,65	0,57	0														
23	0,68	0,77	0,69	4,14	0,26	0,01	0,23	1,20	0,03	0,30	0,27	0,89	3,79	2,64	0,02	2,36	0,11	0,43	0,04	2,15	2,07	1,50	0													
24	0,94	1,04	0,96	4,41	0,01	0,26	0,04	1,47	0,24	0,03	0,00	1,16	4,06	2,91	0,29	2,63	0,37	0,70	0,23	2,42	2,34	1,77	0,27	0												
25	1,06	0,96	1,04	2,41	1,99	1,74	1,96	0,53	1,76	2,03	2,00	0,84	2,06	0,91	1,71	0,63	1,63	1,30	1,77	0,42	0,34	0,23	1,73	2,00	0											
26	0,37	0,27	0,35	3,10	1,30	1,05	1,26	0,17	1,07	1,34	1,31	0,15	2,75	1,60	1,02	1,32	0,93	0,61	1,08	1,11	1,03	0,46	1,04	1,31	0,69	0										
27	3,37	3,28	3,35	0,10	4,30	4,05	4,27	2,84	4,07	4,35	4,32	3,15	0,26	1,40	4,02	1,69	3,94	3,61	4,09	1,90	1,98	2,55	4,04	4,31	2,31	3,00	0									
28	3,61	3,51	3,59	0,14	4,54	4,29	4,51	3,08	4,31	4,58	4,55	3,39	0,49	1,64	4,26	1,92	4,18	3,85	4,32	2,13	2,21	2,78	4,28	4,55	2,55	3,24	0,24	0								
29	2,37	2,28	2,35	1,10	3,30	3,05	3,27	1,84	3,08	3,35	3,32	2,16	0,74	0,41	3,03	0,69	2,94	2,62	3,09	0,98	1,55	3,05	3,32	1,32	2,01	1,00	1,23	0								
30	0,56	0,47	0,54	2,91	1,49	1,24	1,46	0,03	1,26	1,54	1,51	0,35	2,55	1,41	1,22	1,12	1,13	0,80	1,28	0,91	0,83	0,26	1,24	1,50	0,50	0,20	2,81	3,05	1,81	0						
31	1,14	1,05	1,12	2,33	2,07	1,82	2,04	0,61	1,84	2,12	2,09	0,93	1,97	0,83	1,80	0,54	1,71	1,38	1,86	0,33	0,25	0,32	1,82	2,09	0,09	0,78	2,23	2,46	1,23	0,58	0					
32	0,56	0,65	0,58	4,03	0,37	0,12	0,34	1,09	0,14	0,42	0,39	0,78	3,67	2,53	0,09	2,24	0,01	0,32	0,16	2,03	1,95	1,39	0,11	0,38	1,62	0,93	3,93	4,17	2,93	1,12	1,70	0				

El estudio se corrobora mediante la obtención de estas distancias, más la medida Stress que informa sobre la bondad de ajuste del modelo de escalamiento multidimensional en el plano, alcanza a 0,001, es decir, con una medición calificada de excelente.

El gráfico del modelo de distancia euclídea siguiente corresponde al diagrama de dispersión de puntos en el plano euclídeo de los dos vectores propios de la matriz  $B = H'AH$  denominada matriz de productos escalares de dimensión 32.

Gráfico 1

Configuración de estímulos derivada



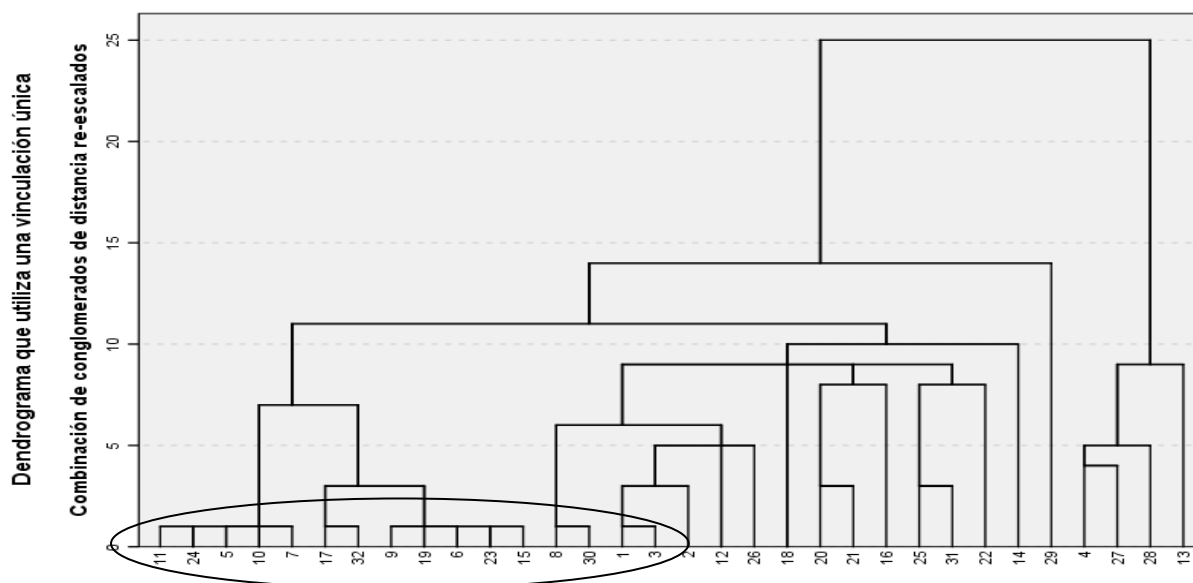
Las variables en el gráfico 1, con los puntos VAR# corresponden a los objetos (municipio). Existen hasta 32 municipios o 32 puntos agrupados alrededor de la línea horizontal y vertical del gráfico 1.

Los municipios concentrados a la derecha del gráfico, son aquellos que mayor similitud o parentesco tienen respecto a la superficie en Ha y producción en qq promedio. Existen puntos

correspondientes a municipios “atípicos” como los números 4, 13, 16, 20, 27, 28, 29 y 30 que por encontrarse aislados en el gráfico frente a los demás, no tienen parentesco y actúan de manera aislada por alguna característica especial que seguro presenta en las variables observadas. Luego se analizará ese aspecto.

Ahora se incorpora un gráfico 2 (dendograma) de un análisis cluster jerárquico para observar la clasificación que se forma entre los 32 municipios estudiados.

Gráfico 2



El método de clasificación corresponde al del vecino más cercano mediante distancia euclidiana. Como se puede notar, los municipios 4 (Achacachi), 13 (Pucarani), 27 (Sica Sica) y 28 (Umala), forman un cluster aparte en el dendrograma, al igual que en el gráfico, estos hacen una especie de línea vertical en el sector izquierdo.

La pregunta es por qué? se marginan de los demás municipios. Entonces se hace un análisis observando los datos de las variables estudiadas, y ordenándolos por superficie cultivada primero de manera ascendente y luego por producción, se

encuentran con color amarillo al final de la tabla. Se detecta que son aquellos que presentan mayor valor en ambas variables, excepto el municipio 29 (Calamarca) y 20 (Coro Coro).

Sin embargo, el municipio de Calamarca, según del dendrograma es municipio de enganche a un “cluster grande”, y los otros municipios identificados con color naranja en el cuadro 2, son municipios agregados en nivel mayor al denominado “cluster grande”. Todos ellos entendidos como los de mayor valor en ambas variables estudiadas.

*“Se puede confiar en las malas personas, no cambian jamás.”.*

*Ralph Waldo Emerson*

Cuadro 2

Nro.	Municipio	Superficie Cultivada (Ha)	Producción Cosechada (qq)
10	Tiquina	3,5	288,0
11	Tito Yupanqui	9,8	917,9
24	Nazacara de Pacajes	11,5	980,2
5	Ancoraimas	12,8	1.237,0
7	Chua Cocani	23,9	1.896,5
19	Desaguadero	55,3	5.717,7
6	Santiago de Huata	55,3	6.464,2
9	Huatajata	57,3	5.985,4
15	Puerto Pérez	73,9	6.996,1
17	Taraco	90,9	8.770,8
23	Comanche	100,5	6.577,2
32	Collana	111,8	8.956,6
18	Guaqui	156,4	15.560,6
1	Puerto Acosta	200,8	20.620,5
3	Escoma	227,0	20.990,5
2	Puerto Carabuco	234,9	22.536,2
12	Copacabana	255,9	25.103,9
26	Jesús de Machaca	256,3	28.207,0
8	Huarina	362,0	31.631,0
22	Calacoto	452,2	37.772,2
30	Patacamaya	476,5	32.275,2
25	Santiago de Callapa	503,0	42.606,7
16	Tiahuanacu	538,2	55.618,0
31	Colquencha	539,8	44.366,8
14	Batallas	568,7	61.551,7
21	Caquiaviri	604,3	49.571,7
20	Coro Coro	809,5	51.248,8
13	Pucarani	859,4	85.392,2
29	Calamarca	872,3	69.980,8
4	Achacachi	882,4	92.816,5
28	Umala	1.101,9	95.653,9
27	Sica Sica	1.135,8	90.730,2

Los municipios que se encuentran en el nivel inicial del gráfico 2 o dendograma (dentro del óvalo), es decir, que son aquellos de mayor similitud con respecto a las variables superficie y producción del producto papa, siendo estos: 1, 3, 5, 6, 7, 8, 9, 10, 11, 15, 17, 19, 23, 24, 30 y 32, se encuentran en su mayoría en los puntos concentrados de la derecha del gráfico 1 (dentro del círculo). El municipio 30 (Patacamaya) se encuentra fuera del círculo rojo, más arriba sobre la línea vertical. Observando si existe alguna diferencia con los datos, se tiene: registra similar superficie cultivada a los municipios del entorno pero registra menor producción a los que lo rodean en casi 10.000 qq (ver cuadro 3 con color verde).

Conclusión del tema. El análisis de escalado multidimensional permite una clasificación de objetos o variables de manera perceptual, gráficamente con visión en dos o tres dimensiones, y decidir de manera exploratoria si estos objetos dibujados en

los ejes coordinados se asemejan o no. Con un agrupamiento cluster, de igual manera se puede analizar a los objetos y su clasificación por cierta similitud. Comparando ambos métodos, uno desde la perspectiva perceptual y el otro desde una clasificación de nivel mediante dendograma, resulta interesante y novedoso a la vez. Se puede incluir a este ejercicio un análisis discriminante más para conseguir mejores resultados.

### 3. Referencias

S.JAMES PRESS (1948): “Applied Multivariate Analysis”, University of Chicago. New York.

TEODORO LUQUE MARTÍNEZ (2000): “Técnicas de Análisis de Datos en Investigación de Mercados”.