

Actas del V Coloquio de Geografía Cuantitativa
Universidad de Zaragoza
1992, Zaragoza

PROGRAMACION DE BASES DE DATOS (DBASEIII.PLUS Y DBASEIV): APLICACION A EJERCICIOS DE RANGOS Y CORRELACIONES.

Enrique LOPEZ LARA

*Departamento de Geografía Física y Análisis Geográfico Regional.
Universidad de Sevilla.*

Los cálculos de rangos y de correlaciones son de amplia utilidad en los análisis geográficos y territoriales, siendo cada vez más difícil su uso en función de la multiplicación de las fuentes y de los datos estadísticos, que hace ingobernable su tratamiento sin la aplicación de la informática (especialmente, bases de datos y hojas de cálculo).

Entre las diversas utilidades de las bases de datos (DBA SEIII.PLUS y DBASEIV), en las que no entramos por ser de sobras conocidas (1), ha de destacarse su capacidad de programación. Es decir, la potencialidad de implementar programas que hagan operaciones de diversa índole con los datos recogidos en la base (CEBRIAN, 1988).

Se ofrece en esta comunicación la formulación y la sistemática para introducir en las bases de datos (DBASEIII.PLUS y DBA SEIV) dos utilidades: el establecimiento de rangos y la realización de correlaciones (de Spearman o por rangos (HAMMOND y McCULLAGH, 1974)). Con la primera, se pueden establecer análisis de rangos entre las variables puestas en lista en la base de datos, sean estas del tipo que fueren. Con la segunda, en parte derivada de la primera, se pueden establecer correlaciones entre las variables que se deseen en función de las características de éstas, de las hipótesis y de los objetivos de trabajo (2).

1.- PROGRAMACION DE LA BASE DE DATOS (DBASEIII.PLUS y DBASEIV) PARA RANGOS Y CORRELACIONES:

Las bases de datos (DBASEIII.PLUS y DBASEIV) ofrecen la potencialidad de su programación. Esto es, aparte de las utilidades que posee el programa en sí, permiten la introducción de formulaciones, estructuras y tratamientos diversos de la información a través de la programación. Programación que se abre a través de la orden "MODIFY COMMAND" (con la que se accede al procesador de textos de la base de datos).

Es en este fichero, que se ha de nombrar (3), en donde se han de introducir las órdenes de programación, fáciles de ejecutar, y que han de adecuarse a la estructura y a las variables de la/s base/s de dato/s que se quieran utilizar. Se ofrece a continuación las órdenes a introducir para programar "RANGOS" y "CORRELACIONES", apareciendo en **negrilla** lo susceptible de ser cambiado (en función de las variables y de la estructura de las bases de datos), y en letra normal, lo básico (órdenes del programa). Una vez introducidas las órdenes se cierra el programador de la base de datos pulsando simultáneamente Control-Fin.

1.1.- PROGRAMACION DE "RANGOS":

```
CLEAR ALL
SET TALK OFF
SET STAT OFF
SET SAFETY OFF
ACCE 'INTRODUZCA NOMBRE FICHERO BASE DE DATOS: '
TO NBD
ACCE 'INTRODUZCA NOMBRE DE BASE DE DATOS PARA
RANGOS: ' TO NРАН
ACCE 'INTRODUZCA NOMBRE DE VARIABLE DISCRIMANTE: '
TO DISC
INPUT 'NUMERO DE VARIABLES PARA CALCULO DE RANGO:
' TO NVAR
ACCEP 'INTRODUZCA LISTA DE VARIABLES SEPARADAS
POR COMAS: ' TO LVAR
LIVAR='COD,NOMBRE,'+LVAR+','+DISC
CLEAR
? 'ESPERE, POR FAVOR...'
USE &NBD
```

```
NREG=RECCO()
COPY TO DATOS FIELDS &LIVAR
?TIME(),FICHERO DE DATOS CREADO,...'
USE DATOS
COPY STRUC EXTEN TO CABEZA
USE CABEZA
DELE FOR FIELD_NAME=UPPER(DISC)
PACK
REPL FIELD_LEN WITH 3,FIELD_DEC WITH 0 FOR
FIELD_TYPE='N'
APPE BLANK
REPL FIELD_NAME WITH 'SUMA',FIELD_TYPE WITH
'N',FIELD_LEN WITH 7,FIELD_DEC WITH 0
APPE BLANK
REPL FIELD_NAME WITH 'RANGO',FIELD_TYPE WITH
'N',FIELD_LEN WITH 3,FIELD_DEC WITH 0
USE
SELE 2
CREATE &NRAN FROM CABEZA
APPE FROM DATOS
SELE 1
USE DATOS
SET RELA TO RECNO() INTO B
CONTA=3
DO WHILE CONTA<=NVAR+2
CAMPO=FIELD(CONTA)
SET EXACT ON
INDEX ON STR(&CAMPO,10)+STR(&DISC,10) TO INDICE
SET EXACT OFF
RANGO=NREG
CAMPOB='B->'+CAMPO
GO TOP
DO WHILE RANGO>0
REPLA &CAMPOB WITH RANGO
RANGO=RANGO-1
SKIP
ENDDO
CONTA=CONTA+1
ENDDO
SET INDEX TO
```

```
SET RELA TO
SELE 2
GO TOP
DO WHILE .NOT. EOF()
CONTA=3
SU=0
DO WHILE CONTA<=NVAR+2
CAMPO=FIELD(CONTA)
SU=&CAMPO+SU
CONTA=CONTA+1
ENDDO
REPL SUMA WITH SU
SKIP
ENDDO
SET RELA TO RECNO() INTO A
DISC='A->'+DISC
SET EXACT ON
INDEX ON STR(SUMA,7)+STR(&DISC,10) TO INDICE
SET EXACT OFF
RELEASE RANGO
RAN=1
GO TOP
DO WHILE RAN<=NREG
REPL RANGO WITH RAN
RAN=RAN+1
SKIP
ENDDO
CLOSE ALL
?CHR(7)
ERASE CABEZA.DBF
SET STAT ON
? 'FIN. '
SET SAFETY ON
SET TALK ON
```

1.2.- PROGRAMACION DE "CORRELACIONES":

```
CLEAR ALL
SET TALK OFF
SET STAT OFF
SET SAFETY OFF
```

```
ACCE 'INTRODUZCA NOMBRE FICHERO BASE DE DATOS: '
TO NBD
ACCE 'INTRODUZCA NOMBRE DE BASE DE DATOS PARA
RANGOS: ' TO NRAN
ACCE 'INTRODUZCA NOMBRE DE LA VARIABLE
DISCRIMINATORIA: ' TO DISC
acce 'INTRODUZCA CONDICION PARA SELECCION DE
MUNICIPIOS: ' TO CONDIC
INPUT 'NUMERO DE VARIABLES PARA CALCULO DE
CORRELACION: ' TO NVAR
INPUT 'NUMERO DE ELEMENTOS DE LA SERIE: ' TO NREG
ACCEP 'INTRODUZCA LISTA DE VARIABLES SEPARADAS
POR COMAS: ' TO LVAR
ACCE 'NOMBRE DE LA MATRIZ DE CORRELACIONES: ' TO
CORR
LIVAR='COD,NOMBRE,'+LVAR+','+DISC
CLEAR
? 'ESPERE, POR FAVOR...'
USE &NBD
COPY TO DATOS FIELDS &LIVAR &CONDIC
DIVISOR=NREG*(NREG^2-1)/6
USE DATOS
COPY STRUC EXTEN TO CABEZA
USE CABEZA
DELE FOR FIELD_NAME=DISC
PACK
REPLA FIELD_LEN WITH 3,FIELD_DEC WITH 0 FOR
FIELD_TYPE='N'
USE
SELE 2
CREATE &NRAN FROM CABEZA
APPE FROM DATOS
SELE 1
USE DATOS
SET RELA TO RECNO() INTO B
CONTA=3
DO WHILE CONTA<=NVAR+2
CAMPO=FIELD(CONTA)
SET EXACT ON
INDEX ON STR(&CAMPO,10)+STR(&DISC,10) TO INDICE
```

```
SET EXACT OFF
RANGO=NREG
CAMPOB='B->'+CAMPO
DO WHILE RANGO>0
REPLA &CAMPOB WITH RANGO
RANGO=RANGO-1
SKIP
ENDDO
CONTA=CONTA+1
ENDDO
CLOSE ALL
USE CABEZA
DELE FOR FIELD_NAME='COD'
PACK
REPLA FIELD_LEN WITH 10 FOR FIELD_TYPE='C'
REPLA FIELD_LEN WITH 5,FIELD_DEC WITH 2 FOR
FIELD_TYPE='N'
USE
CREATE &CORR FROM CABEZA
SELE 2
USE &NRAN
SELE 1
CONTAF=2
DO WHILE CONTAF<=NVAR+1
APPE BLAN
REPLA NOMBRE WITH FIELD(CONTAF)
VARX=NOMBRE
CONTAC=CONTAF+1
DO WHILE CONTAC<=NVAR+1
VARY=FIELD(CONTAC)
DO SPEARMAN
REPLA &VARY WITH RO
CONTAC=CONTAC+1
ENDDO
CONTAF=CONTAF+1
ENDDO
CLEAR ALL
?CHR(7)
ERASE CABEZA.DBF
? 'FIN.'
```

SET SAFETY ON
SET TALK ON
SET STAT ON

1.3.- PROGRAMACION DE "SPEARMAN":

SELE 2
SUM (&VARX-&VARY)^2 TO DIVIDEN
COCIENTE=DIVIDEN/DIVISOR
PUBLIC RO
RO=1-COCIENTE
SELE 1

2.- APLICACION PRACTICA: RANGOS Y CORRELACIONES EN UNA BASE DE DATOS MUNICIPAL.

Los programas anteriores se han utilizado en varios ejercicios (4), mostrándose de gran utilidad y potencialidad operativa. Se ofrece a continuación un ejercicio práctico y sencillo, a efectos de demostrar la facilidad de aplicación y la bondad de los resultados. El ejercicio es simple: se ha recogido por cada término municipal de la provincia de Sevilla (un total de 102) los siguientes datos (variables):

- 1.- porcentaje de población vieja (>64 años); Padrón de Habitantes de 1986.
- 2.- altitud (sobre el nivel del mar); Nomenclátor de 1981.
- 3.- porcentaje de pendientes superiores al 45%; (I.M.A., 1991).
- 4.- porcentaje de población activa en el sector primario; Censo de 1981.

Trátase de demostrar el grado de correlación entre estas variables, así como el rango de cada municipio según éstas, en su conjunto. Para ello hemos aplicado la programación apuntada en el anterior apartado, llegándose a los siguientes resultados:

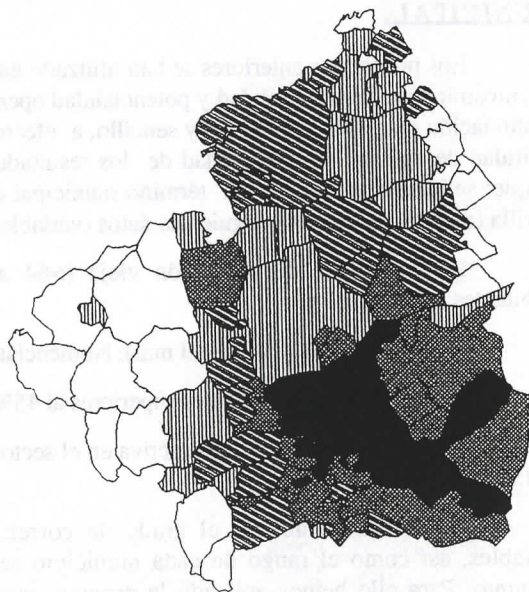
- base datos s: rango (MAPA 1) y CUADRO I: Zonas geográficas.

- base datos sr1: grado de correlación según Spearman.

CUADRO II.

RANGOS/"RANKING" ENTRE CINCO VARIABLES.
PROVINCIA DE SEVILLA. 1986.

Fuente: Elaboración propia.



INTERVALOS DE RANGOS.

□	1 a	21
▨	22 a	41
▩	42 a	61
▧	62 a	81
■	82 a	103

CUADRO I. RANGOS DE LOS MUNICIPIOS DE LA PROVINCIA DE SEVILLA SEGUN ALTITUD (1), PORCENTAJE DE PENDIENTES (2), PORCENTAJE DE POBLACION VIEJA (3), PORCENTAJE DE ACTIVOS EN EL SECTOR PRIMARIO (4).

COD	MUNICIPIO	R	1	2	3	4	SUMA	R
2	ALANIS	1	3	6	6	7	22	1
54	LORA DE ESTEPA	2	14	5	21	12	52	2
80	REAL DE LA JARA (EL)	3	8	2	12	35	57	3
76	PRUNA	4	1	9	30	19	59	4
100	VILLANUEVA DE SAN JUAN	5	11	4	23	25	63	5
32	CAZALLA DE LA SIERRA	6	5	10	7	44	66	6
9	ALMADEN DE LA PLATA	7	10	7	4	46	67	7
66	NAVAS DE LA CONCEPCION (LAS)	8	15	15	19	21	70	8
35	CORIFE	9	26	1	35	10	72	9
48	GUADALCANAL	10	2	28	8	34	72	10
73	PEDROSO (EL)	11	17	14	5	41	77	11
90	SAUCEJO (EL)	12	9	34	28	15	86	12
78	PUEBLA DE LOS INFANTES (LA)	13	36	12	16	24	88	13
57	MADROÑO (EL)	14	21	17	1	54	93	14
14	BADOLATOSA	15	34	40	18	4	96	15
33	CONSTANTINA	16	7	19	15	59	100	16
8	ALGAMITAS	17	16	3	75	11	105	17
31	CASTILLO DE LAS GUARDAS (EL)	18	23	22	2	64	111	18
64	MONTELLANO	19	28	30	25	28	111	19
37	CORRALES (LOS)	20	20	24	71	1	116	20
27	CASTILBLANCO DE LOS ARROYOS	21	24	25	17	51	117	21
46	GILENA	22	13	20	31	55	119	22
43	GARROBO (EL)	23	29	13	20	74	136	23
83	RONQUILLO (EL)	24	22	33	11	75	141	24
88	SAN NICOLAS DEL PUERTO	25	6	41	13	86	146	25
99	VILLANUEVA DEL RIO Y MINAS	26	42	27	3	82	154	26
42	FUENTES DE ANDALUCIA	27	55	70	22	8	155	27
1	AGUADULCE	28	30	98	10	18	156	28
74	PEÑAFLOR	29	79	39	29	16	163	29
77	PUEBLA DE CAZALLA (LA)	30	39	11	52	63	165	30
41	ESTEPA	31	4	29	53	80	166	31
22	CAMPANA (LA)	32	52	76	39	3	170	32
36	CORONIL (EL)	33	38	77	27	29	171	33
61	MARINALEDA	34	35	94	44	2	175	34
26	CASARICHE	35	27	79	56	14	176	35
72	PEDRERA	36	12	31	69	65	177	36
68	OSUNA	37	25	43	50	60	178	37
24	CARMONA	38	33	54	49	43	179	38
30	CASTILLEJA DEL CAMPO	39	59	102	14	5	180	39
84	RUBIO (EL)	40	37	85	42	17	181	40
62	MARTIN DE LA JARA	41	19	45	89	31	184	41
65	MORON DE LA FRONTERA	42	31	26	48	84	189	42
45	GERENA	43	64	44	40	45	193	43
15	BENACAZON	44	61	38	57	38	194	44
52	LANTEJUELA (LA)	45	45	90	59	9	203	45
25	CARRION DE LOS CESPEDES	46	65	93	9	37	204	46
13	AZNALCOLLAR	47	50	16	55	83	204	47
50	HERRERA	48	32	75	26	71	204	48
23	CANTILLANA	49	77	67	38	23	205	49

Cuadro 1 cont

71	PARADAS	50	92	71	24	26	213	50
82	RODA DE ANDALUCIA (LA)	51	18	83	41	72	214	51
60	MARCHENA	52	53	57	37	67	214	52
97	VILLAMANRIQUE DE LA CONDESA	53	88	89	33	6	216	53
39	ECIJA	54	62	51	61	42	216	54
6	ALCOLEA DEL RIO	55	86	88	32	13	219	55
19	BURGUILLOS	56	70	35	88	30	223	56
56	LUISIANA (LA)	57	41	72	68	50	231	57
101	VILLAVERDE DEL RIO	58	96	21	84	33	234	58
49	GUILLENA	59	91	8	67	68	234	59
92	TOCINA	60	89	69	54	22	234	60
28	CASTILLEJA DE GUZMAN	61	56	46	47	89	238	61
85	SALTERAS	62	46	95	45	52	238	62
11	ARAHAL (EL)	63	60	58	63	58	239	63
58	MAIRENA DEL ALCOR	64	51	62	51	76	240	64
55	LORA DEL RIO	65	84	42	62	57	245	65
67	OLIVARES	66	40	73	82	61	256	66
18	BRENES	67	94	65	65	32	256	67
63	MOLARES (LOS)	68	73	92	74	20	259	68
12	AZNALCAZAR	69	78	91	43	49	261	69
5	ALCALA DEL RIO	70	90	66	66	40	262	70
102	VISO DEL ALCOR (EL)	71	49	60	79	79	267	71
95	UTRERA	72	80	50	76	62	268	72
53	LEBRIJA	73	85	52	96	36	269	73
10	ALMENSILLA	74	82	101	36	53	272	74
87	SANLUCAR LA MAYOR	75	47	68	81	77	273	75
16	BOLLULLOS DE LA MITACION	76	69	81	58	66	274	76
75	PILAS	77	76	64	64	70	274	77
69	PALACIOS Y VILLAFRANCA (LOS)	78	97	53	97	27	274	78
51	HUEVAR	79	71	97	34	73	275	79
94	UMBRETE	80	57	82	70	69	278	80
20	CABEZAS DE SAN JUAN (LAS)	81	74	61	100	48	283	81
79	PUEBLA DEL RIO (LA)	82	93	59	92	39	283	82
3	ALBAIDA DEL ALJARAFE	83	43	100	94	47	284	83
40	ESPARTINAS	84	54	96	46	88	284	84
98	VILLANUEVA DEL ARISCAL	85	44	87	73	81	285	85
93	TOMARES	86	72	18	101	98	289	86
86	SAN JUAN DE AZNALFARACHE	87	81	23	91	102	297	87
59	MAIRENA DEL ALJARAFE	88	63	36	102	97	298	88
96	VALENCINA DE LA CONCEPCION	89	48	84	72	96	300	89
29	CASTILLEJA DE LA CUESTA	90	66	37	99	101	303	90
38	DOS HERMANAS	91	68	48	95	94	305	91
21	CAMAS	92	98	32	77	100	307	92
91	SEVILLA	93	101	47	60	99	307	93
81	RINCONADA (LA)	94	99	56	98	56	309	94
4	ALCALA DE GUADAIRA	95	75	49	90	95	309	95
47	GINES	96	58	80	86	92	316	96
17	BORMUJOS	97	67	78	85	90	320	97
34	CORIA DEL RIO	98	102	55	78	85	320	98
7	ALGABA (LA)	99	100	63	80	87	330	99
89	SANTIPONCE	100	95	74	83	91	343	100
44	GELVES	101	83	86	87	93	349	101
70	PALOMARES DEL RIO	102	87	99	93	78	357	102

Fuente: Elaboración Propia, 1991.

El ejemplo es trivial, pero ayuda a comprender la potencia lidad de la herramienta de programación de la base de datos. Los resultados señalan un grado de relación , mayor o menor, según variables, superando el nivel de significación del 99 % (HAMMOND y McCULLAGH, 1974).

3.- CONCLUSIONES.

Múltiples son las perspectivas analíticas que permite la programación de las bases de datos. En este caso se han adaptado dos programas, uno para el establecimiento de rangos, y otro para la realización de correlaciones. Los resultados son óptimos, la rapidez alta y el volumen de datos utilizable prácticamente ilimitado. Son múltiples los programas adaptables, en función de las necesidades, estructura de las bases de datos así como de las hipótesis y de los objetivos de trabajo.

CUADRO II

GRADO DE CORRELACION SEGUN EL COEFICIENTE DE SPEARMAN ENTRE VARIABLES DE LOS MUNICIPIOS DE SEVILLA.

ALTITUD	PENDIENTES	PORC.>64	PORC.AGRI.	
ALTITUD	-	0.43	0.57	0.24
PENDIENTES	0.43	-	0.24	0.02
PORCENTAJE > 64	0.57	0.24	-	0.40
PORCENTAJE AGRI	0.24	0.02	0.40	-

Fuente: Elaboración Propia.

Las bases de datos DBASEIII.PLUS y DBASEIV permiten, asimis mo, la transferencia de información a hojas de cálculo (por ejemplo, LOTUS123), a través de las cuales se pueden calcular y representar gráficamente coeficientes de correlación/regresión y rectas de regresión. De la misma manera, se puede transferir información de las Hojas de Cálculos a programas de gráficos (por ejemplo, Havard Graphics), donde se pueden establecer representa ciones adecuadas a los fines perseguidos, además de a paquetes estadísticos de gran potencialidad, como, por ejemplo, el CSS Statistics.

BIBLIOGRAFIA.

- CANO GARCIA, G. (Director) Geografía de Andalucía. 8 Tomos. Ed. Tartessos. Sevilla, 1987-1990.
- CEBRIAN DE MIGUEL, J.A. "Hojas de cálculo y programas de gestión de bases de datos" Aplicaciones de la Informática a la Geografía y Ciencias Sociales. Ed. Síntesis. Madrid, 1988. pp. 87-112.
- HAMMOND, R. y McCULLAGH, P. Técnicas cuantitativas en Geografía. Ed. Saltés. Madrid, 1974. pp. 375.
- LOPEZ LARA, E. Salud y Territorio: Geografía Médica y Asistencial de Andalucía. Tesis Doctoral (Julio de 1991). (En prensa).
- LOPEZ LARA, E. "Distribución territorial del envejecimiento de la población andaluza. Implicaciones administrativas y socio- sanitarias" R.E.A., 17. Sevilla (En prensa, 1992).

NOTAS:

(1) Las bases de datos (DBASEIII.PLUS y DBASEIV) consisten en programas muy útiles para el tratamiento de un gran volumen de datos (McGraw-Hill).

(2) Ha de apuntarse que las hojas de cálculos, especialmente la LOTUS123, permiten la realización de estos ejercicios, previa programación, así como la ejecución de rectas de regresión. El paquete estadístico CSS.Statistics, asimismo, aunque con ciertas variantes y potencialidades, permite el ejercicio anterior.

(3) Es decir, se le ha de dar un nombre, tanto para el programa de rangos (que puede ser "rangos") como para el de correlaciones ("correl") y, adosado a este, el "Spearman".

(4) Especialmente en el tomo VIII de la Geografía de Andalucía (CANO GARCIA, 1990 (Director)), en una tesis doctoral y en un artículo sobre el proceso de envejecimiento en Andalucía (LOPEZ LARA, 1991 y 1992, respectivamente).