



Ejercicios con ordenador

Descriptiva y Componentes principales

Prof. Esperanza Ayuga Téllez

Tareas a realizar en la práctica



- *Análisis descriptivo de las variables*

En este caso concreto:

resumen estadístico con medidas de centralización, dispersión y

representación gráfica de cada variable y en conjunto.

- *Análisis de los Componentes Principales*

En este caso concreto:

Selección del número de componentes principales (CP)

Obtener la ecuación de dichas CP

Obtener y registrar los valores de las CP para las observaciones

Ejemplo con datos del sector forestal

Información



Información sobre los datos y resultados de las pruebas realizadas

De cada una

Comentario explicativo

Preparación de Datos

Selección de CP

Se ha utilizado el criterio de la varianza explicada mayor del 80% y se obtienen.....

Información



Preparación de Datos

STATGRAPHICS Plus - StatFolio sin Nombre - [Análisis Multivariable]

Archivo Edición Gráficos Descripción Comparación Dependencia Avanzado SnapStats!! Ver Ventana Ayuda

Etiqu.: Filas: 2

Resumen del Procedimiento

	Muestra 1:	Muestra 2:	n
Frecuencia	17	17	1
Media	2,97596E6	439694,0	1
Desviación típica	3,04819E6	464620,0	1
Coef. de variación	102,427%	105,669%	9

	Muestra 5:	Muestra 6:	n
Frecuencia	17	17	1
Media	27,6229	0,412941	1
Desviación típica	22,5815	0,387052	1
Coef. de variación	81,7491%	93,7306%	1

Covarianzas

	Superficie	SF pública	S
Superficie	9,29148E12 (17)	1,2775E12 (17)	3
SF pública	1,2775E12 (17)	2,15871E11 (17)	3
SF privada	3,09178E12 (17)	3,77542E11 (17)	1
Población total	9,13996E11 (17)	8,22582E10 (17)	4
Inversiones por ha F	-1,21788E7	-2,64125E6	-

Utilice el botón alternativo para seleccionar opciones

NUM

Inicio Bandeja de entrada ... Urban studies (refer... Prácticas Microsoft PowerPoint... STATGRAPHICS Plus ... ES 12:59

Información



Componentes Principales

STATGRAPHICS Plus - StatFolio sin Nombre - [Análisis de Componentes Principales]

Archivo Edición Gráficos Descripción Comparación Dependencia Avanzado SnapStats!! Ver Ventana Ayuda

Control de Calidad
Diseño Experimental
Análisis Series Temporales
Métodos Multivariantes
Regresión Avanzada

Componentes Principales...
Análisis Factorial...
Análisis Cluster...
Análisis Discriminante...
Correlaciones Canónicas...

Número de casos completos: 17
Tratamiento de valor perdido: lista considerada
Estandarizado: si

Número de componentes extraídos: 3

Análisis de Componentes Principales

Componente	Autovalor	Porcentaje de Varianza	Acumulado Porcentaje
1	5,44076	45,340	45,340
2	3,25231	27,103	72,442

Tabla de Pesos de los Componentes

	Componentes 1	Componentes 2	Componentes 3
Superficie	0,219324	-0,453028	0,0382444
SF pública	0,178156	-0,433496	0,146876
SF privada	0,241187	-0,427964	-0,0417792
Población total	0,367764	0,173313	-0,241123
Inversiones por ha F	-0,00184632	0,150816	-0,011412
Contratos F	0,012363	-0,401973	-0,201413
E aserrado	0,206315	-0,175499	-0,384792

Tabla de Componentes Principales

Fila	Etiqueta	Componentes 1	Componentes 2	Componentes 3
1	Andalucía	3,57689	-1,39825	0,560666
2	Aragón	-0,700707	-1,56484	0,957767
3	Asturias	-1,05837	0,117671	0,0144567
4	Baleares	-2,01657	1,2834	-0,793101
5	Canarias	-1,66527	0,795153	0,693725

Autovalor

Componente

Extraer componentes principales

NUM

Inicio Bandeja de entrada ... Urban studies (refer... Prácticas Microsoft PowerPoint... STATGRAPHICS Plus ... ES 13:05

Información



Preparación de Datos

STATGRAPHICS Plus - StatFolio sin Nombre

Archivo Edición Gráficos Descripción Comparación Dependencia Avanzado SnapStats! Ver Ventana Ayuda

Análisis de Componentes Principales

Etq.: Fil:

Opciones Componentes Principales

Tratamiento de Valores Perdidos:
 Lista Acertada
 Par Acertado

Extraído Por:
 Autovalor Mínimo
 Número de Componentes

Autovalor Mínimo: 1

Números de Componentes: 12

Estandarizado

Aceptar Cancelar Ayuda

Entrada de Datos
Número de Variables: 12
Tratamiento de Valores Perdidos: Lista Acertada
Número de Componentes: 12
Análisis de Componentes Principales

Variable	Componente 1	Componente 2	Componente 3
Superficie	0,219324	-0,453028	0,0382444
SF pública	0,178156	-0,433496	0,146876
SF privada	0,241187	-0,427964	-0,0417792
Población total	0,367764	0,173313	-0,241123
Inversiones por ha F	-0,00184632	0,150816	-0,811412

Tabla de Componentes Principales

Fila	Etiqueta	Componentes 1	Componentes 2	Componentes 3
1	Andalucía	3,57689	-1,39825	0,560666
2	Aragón	-0,700707	-1,56484	0,957767
3	Asturias	-1,85837	0,117671	0,0144567
4	Baleares	-2,01657	1,2834	-0,793101

Gráfico de Sedimentación

Diagrama de dispersión

Gráfico de Pesos del Componente

Biplot

Sin No... StatAd... StatGal... StatRe... Datos ... Gráfico...

Listo NUM

Inicio Bandeja de entrada ... Urban studies (refer ... Prácticas Microsoft PowerPoint... STATGRAPHICS Plus ... ES 13:23

Resultados gráficos



Incluir en descripción

Histogramas

Diagramas de cajas

Gráficos de estrella

Gráficos bidimensionales

Incluir en análisis de CP

Gráfico de autovalores

Observaciones frente a componentes

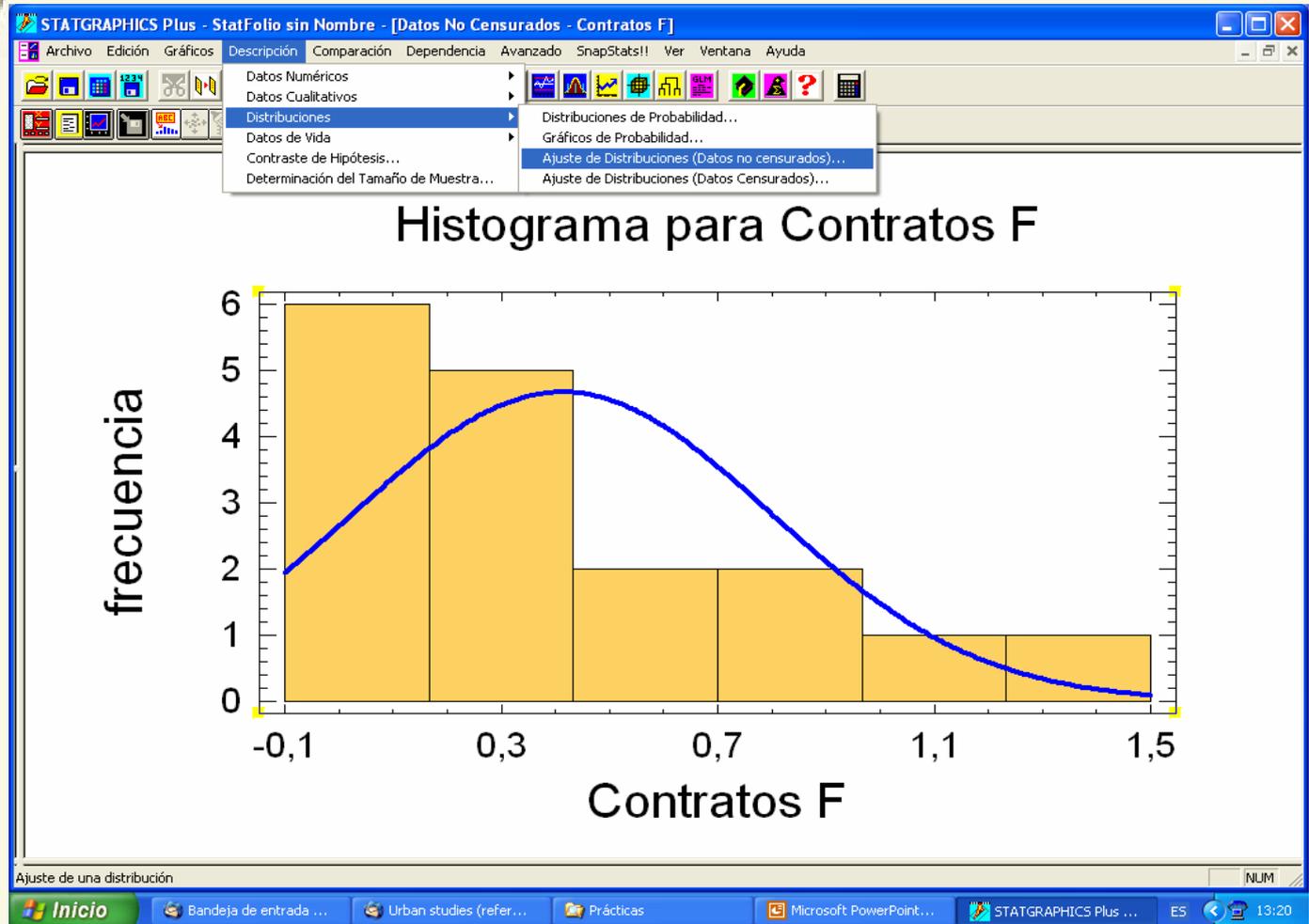
Variables frente a componentes

Variables y observaciones frente a componentes

Resultados gráficos



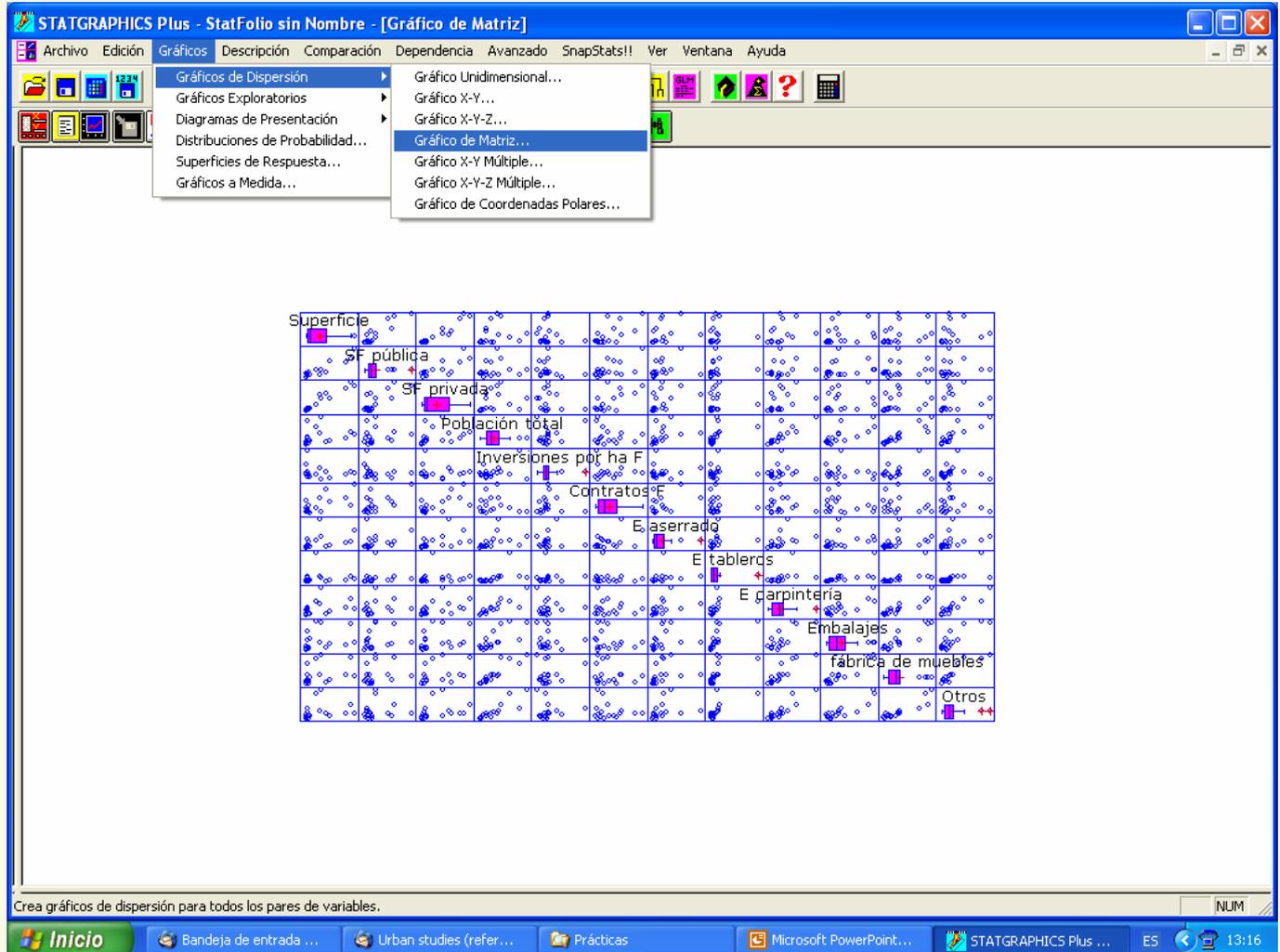
- Se obtienen en resultados de Descripción



Resultados gráficos



- Se obtienen en resultados Gráficos



Resultados gráficos



- Se obtienen en opciones gráficas de CP

STATGRAPHICS Plus - StatFolio sin Nombre - [Análisis de Componentes Principales]

Archivo Edición Gráficos Descripción Comparación Dependencia Avanzado SnapStats!! Ver Ventana Ayuda

Entrada de datos: observaciones
Número de casos completos: 17
Tratamiento de valor perdido: lista consider
Estandarizado: si

Opciones Gráficas

- Gráfico de Sedimentación
- Gráfico de Dispersión 2D
- Gráfico de Dispersión 3D
- Gráfico de Componentes 2D
- Gráfico de Componentes 3D
- Bigráfico 2D
- Bigráfico 3D

Aceptar Cancelar Todos Ayuda

12 0,00409286 0,034 100,01

El StatAdvisor

Este procedimiento realiza un análisis de... El propósito del análisis es obtener un pequeño número de combinaciones lineales de las 12 variables que expliquen la variabilidad en los datos. En este caso, se obtienen 12 componentes, ya que los 2089955858 componentes mayores o iguales a 1,0. Juntos explican el 100% de la variabilidad en los datos originales.

Utilice el botón alternativo para seleccionar opciones

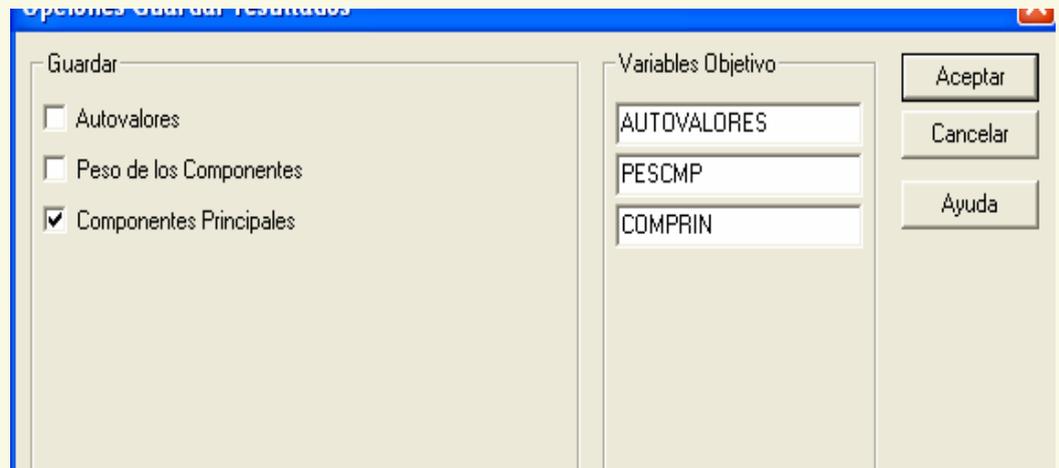
NUM

Inicio Bandeja de entrada ... Urban studies (refer... Prácticas Microsoft PowerPoint... STATGRAPHICS Plus ... ES 13:08

Resultados



- Validación de resultados



Resultados



- Validación de resultados

STATGRAPHICS Plus - StatFolio sin Nombre - [Análisis de los Componentes de la Varianza]

Archivo Edición Gráficos Descripción Comparación Dependencia Avanzado SnapStatsII Ver Ventana Ayuda

Análisis de los Componentes de la Varianza

Comunidad a
Superficie
SF pública
SF privada
Población tot
Inversiones p
Contratos F
E aserrado
E tableros
E carpintería
Embalajes
fábrica de m
Otros
PCOMP_1
PCOMP_2
PCOMP_3
Nº cluster
E

Variable Dependiente:

Factores por Orden de Nivel

(Selección:)

Ordenar

Aceptar Cancelar Borrar Transformar... Ayuda

Análisis de la Varianza para PCOMP_1

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio
TOTAL (CORREGIDO)	87,0522	16	
fábrica de muebles	87,0522	16	5,44076
Otros	-1, #IND	0	
E carpintería	-1, #IND	0	
Embalajes	-1, #IND	0	
Superficie	-1, #IND	0	
SF pública	-1, #IND	0	
SF privada	-1, #IND	0	
Población total	-1, #IND	0	
Inversiones por ha	-1, #IND	0	
Contratos F	-1, #IND	0	
E aserrado	-1, #IND	0	
E tableros	-1, #IND	0	

Gráfico de Componentes de la Varianza

Gráfico de Residuos para PCOMP_1

Gráfico de Residuos para PCOMP_1

Gráfico de Residuos para PCOMP_1

Listo

Inicio Bandeja de entrada... Urban studies (refe... Prácticas Microsoft PowerPol... STATGRAPHICS Plus... ES 13:42

Resultados



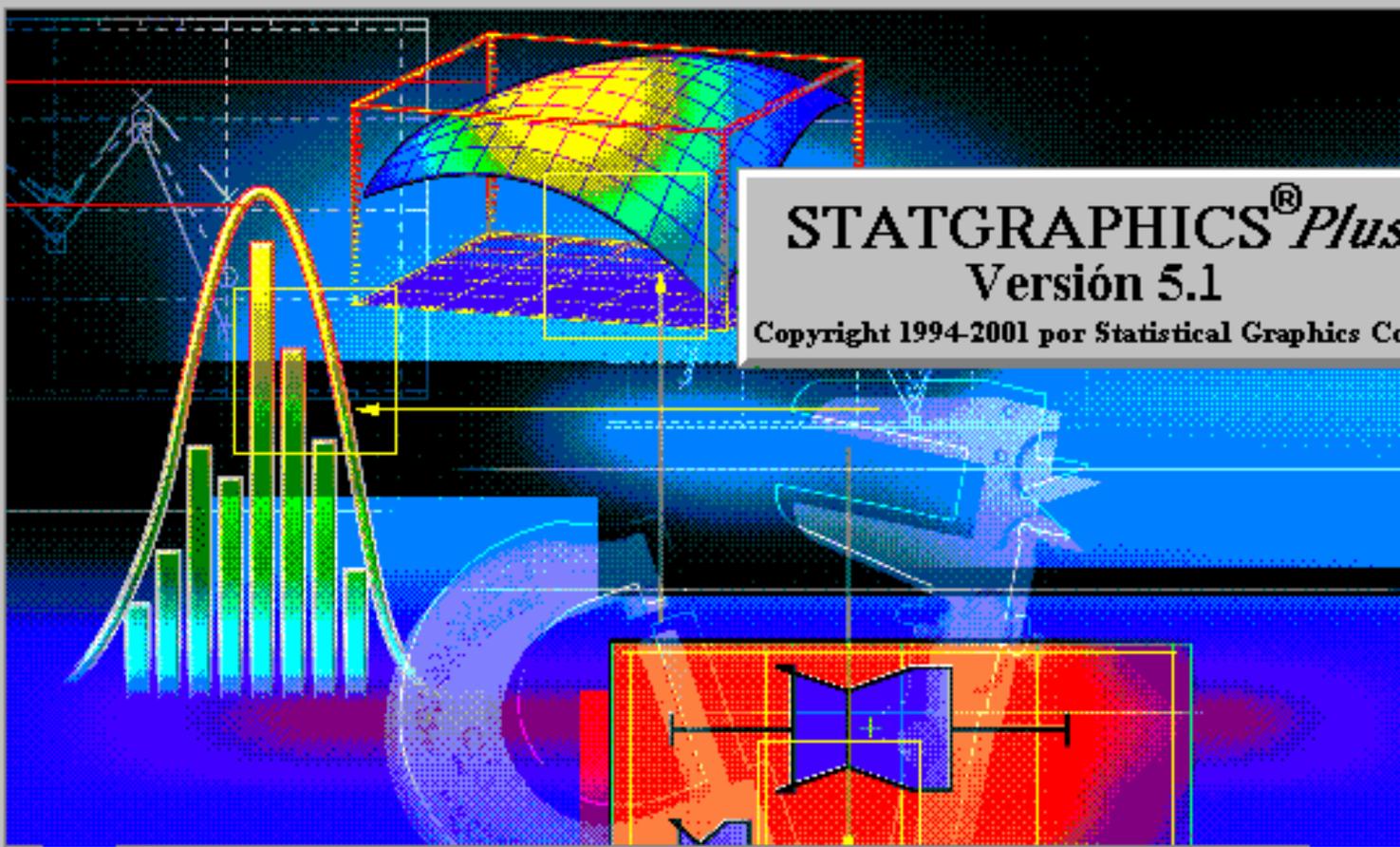
- Interpretación de resultados

Las 12 variables se resumen en 3 CP
La 1ª componente representa

- Consecuencias prácticas

El empleo en el sector depende de las CP y mejorando la fabricación de muebles y las inversiones generales aumenta la empleabilidad.

EJERCICIO



The image features a central text box with the following text:

STATGRAPHICS® *Plus*
Versión 5.1
Copyright 1994-2001 por Statistical Graphics Corp.

The background is a collage of various statistical charts and plots:

- A 3D surface plot with a color gradient from blue to yellow, showing a curved surface within a wireframe box.
- A bar chart with green bars of varying heights, overlaid with a smooth, bell-shaped curve.
- A pie chart with a blue and white pattern, partially obscured by other elements.
- A box plot with a blue box and whiskers, set against a red and orange background.
- A 3D wireframe structure, possibly representing a network or a complex geometric shape.

At the bottom of the image, there is a line of text:

Este programa está protegido por las leyes internacionales y de Estados Unidos de América referentes a los derechos de propiedad según se indica en la caja.