

Estadística en Excel

Introducció

- Usar el full de càlcul
 - per preparar les nostres classes
 - per orientar
 - per veure què podem demanar que facin per ells
- Esquema
 - Introducció al full de càlcul (fórmules i còpia)
 - Explotació d'una base de dades segons el tipus de variables que conté. Està al fitxer estsec.xls
 - Fórmules
 - Menú anàlisi de dades
 - Taules pivot (o dinàmiques)
 - Comencem amb la simulació
 - Etc.....

El full de càlcul

- Format per files (1,2,...) i columnes (A,B,...Z,AA,AB,...).
- Cel·les: identificades amb ColumnaFila (A1,A2,B3)
- Contingut de les cel·les
 - Nombres
 - Lletres
 - Expressions
 - Fórmules
 - Comencen per =
 - Poden tenir nombres, lletres i **referències a altres cel·les**

Referències a cel·les

- Si en B2 escrivim =A2

	A	B
1		
2		=A2

a la cel·la B2 sempre es mostra el que hi ha a A2 (encara que canviem el contingut de A2):

	A	B
1		
2	2	2

	A	B
1		
2	300	300

Copiar cel·les o rangs (rectangles de cel·les)

- Marcar la cel·la o el rang i fer “Edición->copiar” i després “Edición->pegar”.
- Arrossegant la cantonada inferior dreta de la cel·la o el rang.



*Arrosseguem
avall*

Referències absolutes i relatives

- Les referències a la fila o la columna poden ser absolutes:
 - S'indiquen amb un \$ davant la fila o la columna
 - En copiar la cel·la la fila o la columna que té un \$ no es modifica.
 - Exemples: \$A4, A\$4, \$A\$4
- Les referències a la fila o la columna poden ser relatives:
 - Exemples: A4
 - En copiar-les traslladem el “vector lliure” (ex. 2 esquerra, 3 amunt....)

Còpia de referències absolutes i relatives

	A	B	C	D		A	B	C	D
1					1				
2	-234	Pep	=A2	=B2	2	-234	Pep	-234	Pep
3	2,71828183	Gos	=A3	=B3	3	2,71828183	Gos	2,71828183	Gos
4					4				
5			=A\$2	=B\$2	5			-234	Pep
6			=A\$2	=B\$2	6			-234	Pep
7					7				
8			=\$A2	=\$A2	8			-234	-234
9			=\$A3	=\$A3	9			2,71828183	2,71828183
10					10				
11			=\$A\$2	=\$A\$2	11			-234	-234
12			=\$A\$2	=\$A\$2	12			-234	-234
13					13				
14					14				
15					15				
16									

→ Vol dir copiar-enganxar

Pantalla 7

Estadística en Excel

Exemples de còpia absoluta relativa

- Nombres correlatius 1,2,3,...
- Nombres de la forma $1+kn$ amb k fixat
- Taula de valors d'una funció (p.ex. x^2)
- Successió de Fibonacci $a_1=1, a_2=1, a_{n+1}=a_{n-1}+a_{n-2} \dots$ (qui ocupa el lloc 49?)
- Totes les taules de multiplicar AMB FÓRMULES EN LES CEL·LES ENCERCLADES i en la resta còpies (excepte l'1 inicial)

1	2	3	4	5	6	7	8	9	10
2	4	6	8	10	12	14	16	18	20
3	6	9	12	15	18	21	24	27	30
4	8	12	16	20	24	28	32	36	40
5	10	15	20	25	30	35	40	45	50
6	12	18	24	30	36	42	48	54	60
7	14	21	28	35	42	49	56	63	70
8	16	24	32	40	48	56	64	72	80
9	18	27	36	45	54	63	72	81	90
10	20	30	40	50	60	70	80	90	100

Pantalla 8

Estadística en Excel

Altres eines

- Còpia màgica!
- Omplir una sèrie
 - Edición->Rellenar->Serie
 - Podem escollir files o columnes
 - Omplir linealment, geomètricament, amb dates...

Explotació d'una base de dades

Variablers del fitxer estsec.xls

- Categòrica: SEXE (1=home, 2=dona)
- Numèrica: EDAT
- Categòrica: SITUACIÓ LABORAL
- Numèrica: INGRÉS

Explotació d'una base de dades:
UNA VARIABLE CATEGÒRICA

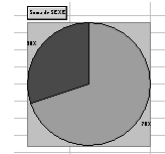
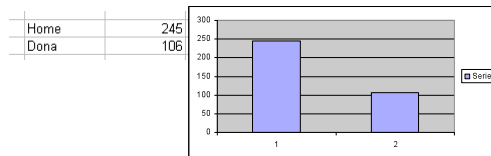
SEXE

- Datos->Asistente tabla... dinàmica

Contar de SEXE	
SEXE	Total
1	245
2	53
Total general	298

- Gràfic de sectors a partir de la taula dinàmica (o de dades agrupades)

- Gràfic de barres a partir de la taula



Pantalla 11

Estadística en Excel

Explotació d'una base de dades:
Exercici: Una altra VARIABLE CATEGÒRICA

SITUACIÓ LABORAL:

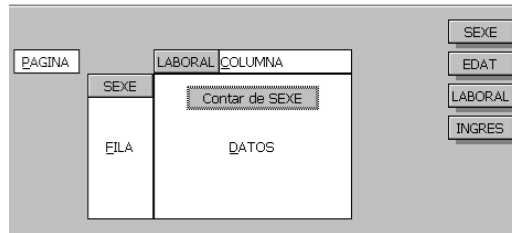
- Freqüència de cada categoria
- Percentatge (freqüència relativa de cada categoria)
- Gràficament

Pantalla 12

Estadística en Excel

Explotació d'una base de dades:
DUES VARIABLES CATEGÒRIQUES: SEXE i LABORAL

- Taula dinàmica: permet calcular freqüències per a cada combinació de valors de les variables (home-fix...): Datos->... Tabla dinámicos...



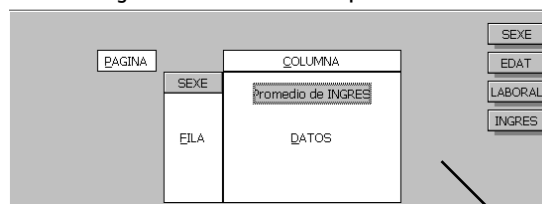
Contar de SEXE	LABORAL					
SEXE	ATURAT	ESTUDIANT	EVENTUAL	FIX		Total general
1	3	12		29		201
2	3	6		12		32
Total general	6	18		41		233

Pantalla 13

Estadística en Excel

Explotació d'una base de dades:
Una VARIABLE CATEGÒRICA I UNA NUMÈRICA

- Taula pivot: permet calcular mitjanes...de la numèrica en cada grup de la categòrica:
 - Mitjanes de INGRES per SEXE



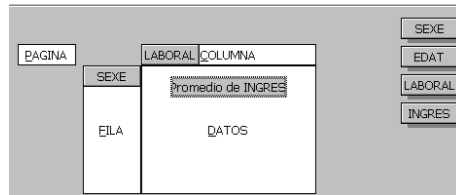
Promedio de INGRES	
SEXE	Total
1	1399,289796
2	1079,471698
Total general	1342,409396

Pantalla 14

Estadística en Excel

Explotació d'una base de dades: DUES VARIABLE CATEGÒRIQUES I UNA NUMÈRICA

- Taula pivot: permet calcular mitjanes...de la numèrica per a cada combinació de valors de les categòriques:
 - Mitjanes de INGRÉS per sexe i categoria laboral



Promedio de INGRES	LABORAL				
SEXE	ATURAT	ESTUDIANT	EVENTUAL	FIX	Total general
1	1012,666667	883,666667	1055	1485,517413	1399,289796
2	653	865,333333	651,25	1320,1875	1079,471698
Total general	832,833333	877,555556	936,829268	1462,811159	1342,409396

Pantalla 15

Estadística en Excel

Explotació d'una base de dades: Una variable numèrica: INGRÉS

INGRÉS:

- Resums numèrics
 - Fórmules (promedio(), mediana(), desvest(), desvestp()....).
 - Menú Herramientas->Análisis de datos (Si no hi és cal afegir-lo des de Herramientas->complementos->herramientas para análisis). Si no hi és a "complementos" cal usar el CD d'instal·lació.
 - Estadística descriptiva
- Gràfics
 - Menú Herramientas->Análisis de datos). Histograma (atenció a "crear gráfico")
 - amb classes "automàtiques"
 - amb classes personalitzades (taula amb extrems drets dels intervals 1000-1500....)

Pantalla 16

Estadística en Excel

Explotació d'una base de dades: Exercici: una variable numèrica: EDAT

INGRÉS:

- Resums numèrics
 - Fórmules (promedio() -mitjana-, mediana(), desvest() -desviació mostral-, desvestp() -desviació poblacional...).
 - Menú Herramientas->Análisis de datos->Estadística descriptiva
- Gràfics
 - Menú Herramientas->Análisis de datos->Histograma amb classes 15-25,25-35.....

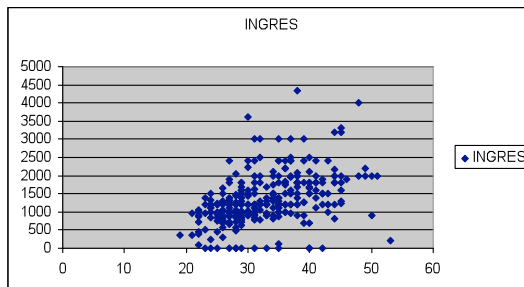
Pantalla 17

Estadística en Excel

Explotació d'una base de dades: DUES VARIABLES NUMÈRIQUES: gràficament

- Podem començar directament pel gràfic XY ("dispersión"=núvol de punts), després d'haver seleccionat les columnes EDAT i SOU amb el que obtenim

Si volem seleccionar dues columnes no consecutives marquem una i fem CONTROL+clik a l'altra

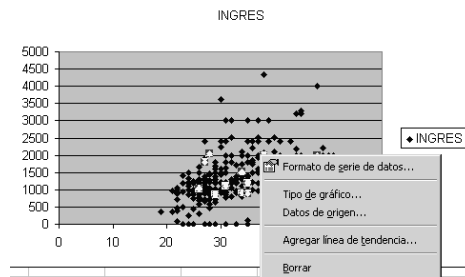


Pantalla 18

Estadística en Excel

Explotació d'una base de dades:
DUES VARIABLES NUMÈRIQUES: gràficament

- A sobre d'aquest gràfic, si cliquem a sobre dels punts del núvol, i un cop seleccionats piquem amb el botó dret ens apareix el quadre

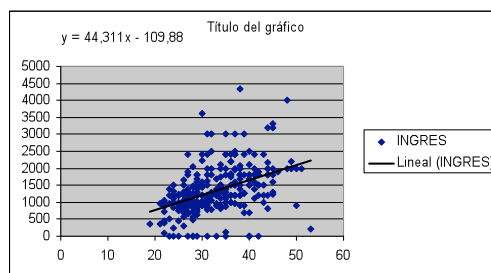


- A través del que podem, des de “agregar línea de tendencia”, fer que aparegui la recta de regressió i la seva equació, amb el que s'obté:

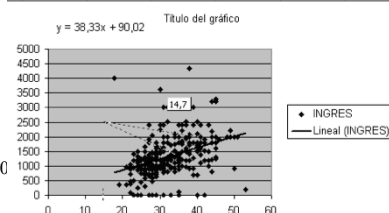
Pantalla 19

Estadística en Excel

Explotació d'una base de dades:
DUES VARIABLES NUMÈRIQUES: gràficament



Cal destacar que si seleccionem un punt el podem arrossegar per veure l'efecte dels punts extrems, per exemple



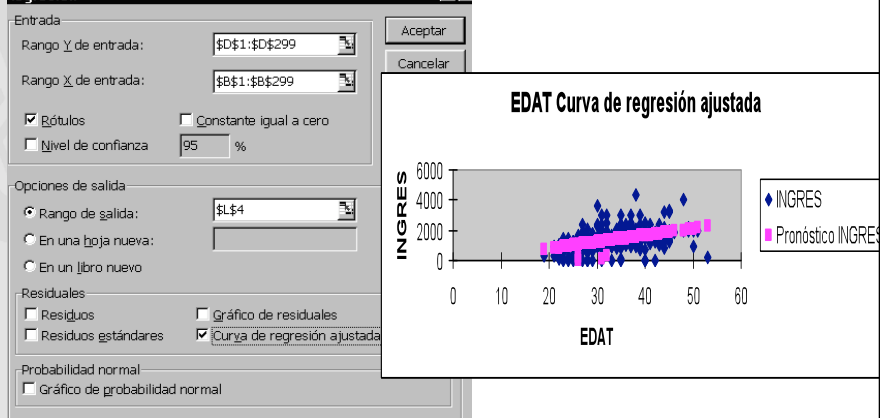
Pantalla 20

Això té l'inconvenient que es modifiquen les dades originals

Estadística en Excel

Explotació d'una base de dades:
DUES VARIABLES NUMÈRIQUES: amb anàlisi de dades

- Anàlisis de datos -> Regresión EDAT-INGRÉS



Pantalla 21

Aquest gràfic també permet afegir l'equació de la recta i veure l'efecte dels extrems ja que de fet és el gràfic XY que hem vist abans

Estadística en Excel

Explotació d'una base de dades:
DUES VARIABLES NUMÈRIQUES

- Anàlisis de datos -> Regresión EDAT-INGRÉS

Estadísticas de la regresión							
Coefficiente de correlación múltiple	0,436606753						
Coefficiente de determinación R ²	0,190625456						
R ² ajustado	0,187891083						
Error típico	613,2238636						
Observaciones	298						
ANÁLISIS DE VARIANZA							
		Grados de libertad	suma de cuadrados	medio de los cuadrados	F	Valor crítico de F	
Regresión		1	26215682,03	26215682,03	69,7144919	2,66677E-15	
Residuos		296	1111308878	376043,5069			
Total		297	137524560,1				
		Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%
Intercepción		-109,8807116	177,52733	-0,618950961	0,53642472	-459,2562919	239,49487
EDAT		44,31068415	5,306973538	8,349520461	2,6668E-15	33,86650657	54,754862

Pantalla 22

Estadística en Excel

Simulació: generació de nombres aleatoris

- Funcions d'Excel
 - =aleatorio() -> nombre aleatori $0 < x < 1$
 - =aleatorio()*10 -> nombre aleatori $0 < x < 10$
 - =aleatorio()*10+1 -> nombre aleatori $1 < x < 11$
 - =entero(aleatorio()*10+1) -> convertit a enter.
 - Finalment tenim nombres enters entre 1 i 10
- Canvien cada cop: es poden fixar amb copiar->pegar especial->pegar valors (a sobre del que tenim).

Dates en Excel

- La data és una qüestió de format.
- Un nombre es pot convertir en una data amb formato->celda->número->fecha
 - 1 és 1-1-1900 o bé 1/1/1900
 - =hoy() és la data d'avui
 - 25/11/2006 és (fent formato->celda->número) 39045