

FONAMENTS MATEMÀTICS I ESTADÍSTICS

(GRUP 52)

Natàlia Castellana

Despatx C1-216, Facultat de Ciències

Horari de consulta:

- Dimarts 17:30-19:30
- Dijous 11:15-13:15

**FONAMENTS MATEMÀTICS I ESTADÍSTICS
per a la Sociologia i la Ciència Política**

Curs 2005-06

Codi 25222. 6 crèdits. Primer curs. Primer semestre

Professorat responsable: Xavier Bardina, Natàlia Castellana, Mercè Farré, Pedro López Roldán i Joan Porti

Objectius

L'objectiu del curs és dotar l'estudiantat de Sociologia i de Ciència Política dels conceptes i dels instruments bàsics de tractament i anàlisi de les dades estadístiques per tal de preparar-lo en l'estudi i l'aplicació de les tècniques d'investigació d'orientació quantitativa específiques de les Ciències Socials que veurà durant els estudis de grau.

En particular, amb l'assignatura es pretén que l'alumnat construeixi els seu aprenentatge en base a:

- El coneixement dels fonaments de la investigació científica des del punt de vista de la utilització conceptual, tècnica i instrumental de l'estadística.
- La comprensió dels conceptes bàsics, introductoris i instrumentals de l'estadística, que s'adrecen a l'estudiantat de ciències socials en relació a l'estadística descriptiva univariable i bivivariable, i l'estadística inferencial.
- Saber escollir i aplicar les tècniques estadístiques bàsiques per a cada tipus de problema d'anàlisi de la realitat social a partir d'informació quantitativa.
- Saber utilitzar de forma bàsica i instrumental el programari estadístic SPSS.
- Saber interpretar fonamentalment els resultats estadístics d'una anàlisi de dades des d'un punt de vista tècnic i substantiu.
- Assolir una capacitat crítica en relació a la informació científica.

Temari

Tema 0. Introducció al contingut del curs. Preliminars.

- L'estadística a les Ciències Socials: estadística descriptiva i inferencial. Exemples d'aplicació
- Escalles de mesura de les variables: qualitatives (nominals i ordinals), quantitatives (discretes i contínues). Càlculs aritmètics i arrodoniments.
- Representació gràfica: sistemes de referència al pla, funcions lineals-equació d'una recta, altres funcions.

Tema 1. Estadística descriptiva d'una variable.

- Distribucions de freqüències: absolutes, relatives i percentatges; acumulades i no acumulades. Agrupament de valors en intervals o classes.
- Representacions gràfiques: el cas de les variables qualitatives (barres, sectors i Pareto) i el cas de les variables quantitatives (histogrames, polígons i ogives). Gràfiques deformades i/o manipulades. Histogrames amb intervals de distinta amplitud.
- Les característiques de posició central (moda, mediana, mitjana). Càlculs usant freqüències relatives, percentatges, o pesos: la mitjana ponderada. Les característiques de posició no-central (percentils, decils i quartils). Comparació de les característiques.
- Les característiques de dispersió (rang, desviació típica, i variància) i una característica de dispersió relativa, el coeficient de variació.
- Les característiques de forma : simetria i curtosi.
- Anàlisi exploratòria: valors extrems (màxim i mínim), valors anòmals (outliers), diagrama de caixa, diagrama de tija i fulles i mitjanes retallades.
- Transformacions de variables. Propietats de les mesures de posició i de dispersió.
- Les puntuacions tipificades.
- Índex i taxa.

Tema 2. Estudi descriptiu de les relacions entre variables.

- Comparació d'una variable numèrica en dos o més grups: el cas de dades independents i el cas de dades emparellades (anàlisi exploratòria descriptiva i gràfica).
- Relació entre dues variables categòriques (taules de contingència i diagrames agrupats).
- Relació entre dues variables numèriques (correlació i regressió).

Tema 3. Fonaments d'estadística inferencial.

- Mostra i població. Mostreig aleatori.
- Teoria elemental de la probabilitat: Introducció al concepte de probabilitat a partir de les freqüències relatives. Definició i propietats. Esdeveniments incompatibles i independents.
- Distribució de probabilitat d'una variable-població. Esperança (o mitjana) i variància poblacionals.
- La distribució Normal: La distribució normal com a model de certes poblacions. Càlculs amb la distribució normal amb l'SPSS. Alguns valors crítics de la distribució normal.
- Algunes distribucions importants derivades de la normal: t-de Student, khi-quadrat de Pearson, i F de Fisher-Snedecor. La significació d'un valor.

Tema 4. Inferència estadística.

- Paràmetres i estadístics: estimacions puntuals i per intervals.
- Interval de confiança per a la mitjana poblacional: nivell de confiança, risc, marge d'error i mida de la mostra.
- Interval de confiança per a la proporció poblacional: nivell de confiança, risc, marge d'error i mida de la mostra.
- La idea dels tests d'hipòtesis: la formulació de la hipòtesi d'investigació, la decisió i el risc d'error. Tests bilaterals i unilaterals. El test t de comparació de dues mitjanes.
- Test khi-quadrat d'independència de dues variables categòriques.

Bibliografia comentada

a) Bibliografia bàsica

La bibliografia bàsica de l'assignatura correspon a dos textos bàsics en procés d'elaboració per part del professorat de l'assignatura que es publicaran per el proper curs a la col·lecció Materials de la Universitat Autònoma de Barcelona. Es tracta de textos elaborats en base als dossiers de teoria de l'assignatura que s'han utilitzat en cursos anteriors. El primer contindrà l'estadística descriptiva univariable, i el segon l'anàlisi bivivariable i inferencial.

A més s'han elaborat els dossiers de pràctiques, documents que es publiquen al Campus Virtual de l'assignatura, que mitjançant exemples i exercicis pràctics il·lustren i apliquen els principals conceptes de l'assignatura i la seva realització amb el programari estadístic SPSS.

b) Bibliografia complementària

- SPSS INC. (2001). *Guía para el análisis de datos*. Madrid: SPSS Hispanoportuguesa. (CD-Rom)

Manual del programari estadístic SPSS que permet ampliar i treballar de forma autònoma el coneixement i domini del programari. N'existeix una versió inicial en format digital, disponible al Campus Virtual de l'assignatura. Posteriorment se'n publicà un text de referència: Pardo, A.; Ruíz, M.A. (2002). SPSS 11. *Guía para el análisis de datos*. McGraw-Hill. Manual complet, ben presentat i de fàcil comprensió, del qual en destaquem la presentació del programari i de les seves funcionalitats; la introducció i identificació de les dades, i l'anàlisi de les dades estadístiques de tipus univariable i bivivariable.

- K. D. Hopkins, B. R. Hopkins, G. V. Glass (1997). *Estadística Básica para las ciencias sociales i del comportamiento*. 3a ed. Naucalpan de Juárez : Prentice-Hall Hispanoamericana.

Text de referència complementària de l'assignatura que cobreix tot el contingut de la matèria, des de l'estadística descriptiva univariable, l'anàlisi de relacions entre dues variables, a l'estadística inferencial. Inclou capítols amb temes més avançats que poden ser útils en cursos posteriors. Conté un disquet amb dades.

- López Roldan, P.; Lozares Colina, C. (1999). *Anàlisi bivivariable de dades estadístiques*. Bellaterra (Barcelona): Universitat Autònoma de Barcelona. Col·lecció Materials, 79.

Manual complementari de l'assignatura per als temes d'anàlisi de la relació entre dues variables (taules de contingència i anàlisi de regressió) amb un contingut que es presenta equilibrat entre la necessària formulació del procediments i la interpretació dels resultats obtinguts amb l'ajut del programari SPSS. El text és fruit de l'experiència docent dels autors en assignatures de tècniques d'investigació i d'estadística en l'àmbit de les ciències socials.

- Sánchez Carrión, J.J. (1999) *Manual de análisis estadístico de los datos*. Madrid: Alianza. Manuales 055.

Manual de referència complementària, que cobreix tot el contingut de la matèria, des de l'estadística descriptiva univariable, l'anàlisi de relacions entre dues variables, a l'estadística inferencial. Es tracta d'un manual senzill i concís adient per a l'estudiantat de ciències socials. El text s'acompanya d'aplicacions i interpretacions dels resultats generats amb l'SPSS . Inclou un disquet amb les dades i la sintaxi dels exemples presentats en el text.

- TERMCAT, Centre de Terminologia (1994). *Diccionari d'estadística*. Barcelona: Termcat.
- <http://www.termcat.net/> Recurs textual i virtual del diccionari terminològic d'estadística en català i altres llengües per a la seva consulta.

Avaluació

L'avaluació de l'assignatura, per tal d'aconseguir els objectius de formació, es planteja en la modalitat d'avaluació continuada. Aquesta consisteix en 6 proves, 4 de continguts pràctics i 2 de continguts més teòrics, d'acord amb la divisió en blocs i subblocs en què s'estructura l'assignatura. A la taula següent es presenta el sistema d'avaluació amb les característiques descriptives bàsiques i el pes de cada prova d'avaluació:

Bloc	Subbloc	Avaluació pràctica	Avaluació teòrica
1. Estadística descriptiva univariable	1.1. Anàlisi dels diferents tipus de distribució de variables	Pes: 1/6 Descripció: Examen a l'aula d'informàtica sobre els continguts corresponents a les pràctiques 1,2 i 3.	Pes: 1/6 Descripció: Examen basat en qüestions teòriques i problemes a l'aula sobre els continguts del temes 0 i 1.
	1.2. Manipulació i transformació d'arxius de dades	Pes: 1/6 Descripció: Examen a l'aula d'informàtica sobre els continguts corresponents a les pràctiques 4 i 5.	
2. Estadística descriptiva bivariàble i estadística inferencial	2.1. Estadística descriptiva bivariàble	Pes: 1/6 Descripció: Examen a l'aula d'informàtica sobre els continguts corresponents a les pràctiques 6, 7 i 8.	Pes: 1/6 Descripció: Examen basat en qüestions teòriques i problemes a l'aula sobre els continguts del temes 2, 3 i 4.
	2.2. Estadística Inferencial	Pes: 1/6 Descripció: Examen a l'aula d'informàtica sobre els continguts corresponents a les pràctiques 9, 10 i 11.	

L'avaluació continuada és l'eina avaluativa de l'aprenentatge progressiu de l'alumnat, per la qual cosa no es contempla una examen final de l'assignatura a la primera convocatòria.

De forma excepcional i transitòria, en aquest proper curs acadèmic, es preveu la realització d'un examen final a la segona convocatòria per a totes les persones que han obtingut una qualificació de suspès en convocatòries de cursos anteriors.

Serà un requisit per a tot l'alumnat de nou ingrés haver seguit l'avaluació continuada per tal de superar l'assignatura a la segona convocatòria.

Pràctiques de laboratori amb SPSS:

1. 30 Setembre
2. 7 d'Octubre
3. 14 d'Octubre
4. 28 d'Octubre
5. 11 de Novembre
6. 25 de Novembre
7. 2 de Desembre
8. 16 de Desembre

9. 13 de Gener

10. 20 de gener

11. 27 de Gener

Avaluació continuada:

- Exàmens de Teoria
 1. 25 Novembre
 2. Entre 30 de Gener i 10 de Febrer

- Exàmens de Pràctiques
 1. 21 d'Octubre
 2. 18 de Novembre
 3. Entre 30 de Gener i 10 de Febrer

Objectius bàsics en la investigació en ciències socials:

- Descripció
- Explicació
- Predicció

Mètodes: Tècniques d'observació, enquestes i experiments

que produeixen un conjunt de dades (números)

Problema real, curiositat

→

Questions concretes sobre el món real

→

Mètode per obtenir informació, dades

→

Obtenció de dades

→

Resumir i analitzar les dades

→

Respostes a la pregunta

Estadística: tècniques (matemàtiques) per organitzar, resumir o sintetitzar, i interpretar dades.

Obtenir resultats minimitzant l'efecte de la incertesa.

Una **població** és el conjunt de tots els individus u objectes d'interés. (Totes les observacions possibles).

Per exemple, si ens preguntem per la intenció de vots dels habitants dels EEUU en les properes eleccions, la població estarà formada per tots els ciutadans censats als EEUU.

Per exemple, si ens preguntem per l'alçada dels eucaliptus, la població estarà per tots els eucaliptus vius del planeta.

Una **mostra** és un subconjunt de la població.

Per exemple, si realitzem una enquesta a 1000 habitants dels EEUU per saber la seva intenció de vot, el conjunt d'individus que han contestat l'enquesta és una mostra.

Per què no estudiem sempre tota la població?

En molts casos, el procés seria massa costós o inviable en qüestió de temps.

A l'escollir una mostra cal fer-ho de manera aleatòria de manera que tots els individus tinguin les mateixes possibilitats de ser inclosos a la mostra. Així és més representativa de la població.

Estadística descriptiva: resumir i explorar les dades d'una mostra (que pot coincidir amb la població).

Estadística inferencial: extrapolar els resultats obtinguts en una mostra aleatòria a la població.

Estadística descriptiva:

Mètode: recollir dades, presentar dades, caracteritzar les dades numèricament.

Objectiu: descriure un conjunt de dades. (fer afirmacions sobre la mostra).

Estadística inferencial:

Mètode: mostreig aleatori, estimacions, test d'hipòtesis.

Objectiu: prendre decisions sobre una població basades en les característiques d'una mostra.

Variables i dades:

Una variable és una mesura que pot prendre valors diferents en individus i circumstàncies diferents. Quan el seu grau d'incertesa és molt gran direm que és una variable aleatòria.

Les dades són els valors que prenen una o més variables en la realització d'un estudi. Les dades les obtindrem a partir de publicacions, enquestes, observacions,.....

Exemple:

Curiositat: qui serà el proper president de la Generalitat?

Pregunta concreta (variable): Intenció de vot dels habitants censats de Catalunya.

Mètode: selecció d'una mostra aleatòria representativa a la que es pregunta per la seva intenció de vot.

Estadística descriptiva: resumir i presentar les dades obtingudes de la mostra.

Estadística inferencial: extrapolar els resultats convenientment de la mostra per preveure el comportament de la població.

Tipus de variables:

- Numèriques o quantitatives
 - Discretes
 - Contínues
- No-numèriques o qualitatives
 - Nominal
 - Ordinal

De quin tipus és la variable de l'exemple anterior?

Tipus d'escala de mesura: (depèn de la natura de les observacions)

- Escala nominal: només permet una classificació dels individus.
- Escala Ordinal: classifica els individus i permet establir ordres.
- Escala numèrica: a més de classificar els individus, ens permet operar amb els valors.

El tipus de variable i escala determinarà els mètodes descriptius que podrem utilitzar en cada cas. Té sentit calcular la mitja d'una variable qualitativa nominal?

Exemples:

De quin tipus són les següents variables?

1. Recompte de bactèries en un cultiu.
2. Pes de les taronges de la collita del 2004 a València.
3. Posicionament a favor o en contra de la llei contra el tabac.
4. Categoria laboral en una empresa.

Condiciones de Vida



Índice de precios de consumo (IPC). Base 2001

Índice general y por grupos

	Índice octubre 2002	Variación enero-octubre 2002
INDICE GENERAL	104,9	3,5
Alimentos y bebidas no alcohólicas		
bebidas no alcohólicas	106,1	3,8
Bebidas alcohólicas y tabaco	106,9	4,6
Vestido y calzado	110,0	2,0
Vivienda	103,0	2,7
Menaje	102,7	1,8
Medicina	103,1	2,4
Transporte	103,4	5,8
Comunicaciones	97,3	-1,7
Ocio y cultura	102,3	0,9
Enseñanza	106,3	3,9
Hoteles, cafes y restaurantes	106,7	5,5
Otros	104,7	3,6

Las que más han subido de enero a octubre

Rúbricas	% Variación
Legumbres y hortalizas frescas	18,1
Carne de ave	9,6
Aceites y grasas	8,8
Frutas frescas	7,7
Transporte público urbano	6,2
Complementos y reparaciones de prendas de vestir	5,8
Transporte personal	5,7
Tabaco	5,5
Educación universitaria	5,4
Turismo y hostelería	5,2

Las que menos han subido de enero a octubre

Rúbricas	% Variación
Carne de ovino	-10,0
Objetos recreativos	-3,5
Comunicaciones	-1,7
Textiles y accesorios para el hogar	-0,7
Café, cacao e infusiones	-0,3
Electrodomésticos y reparaciones	0,0
Patatas y sus preparados	0,2
Huevos	0,4
Medicamentos y material terapéutico	0,5
Leche	0,6

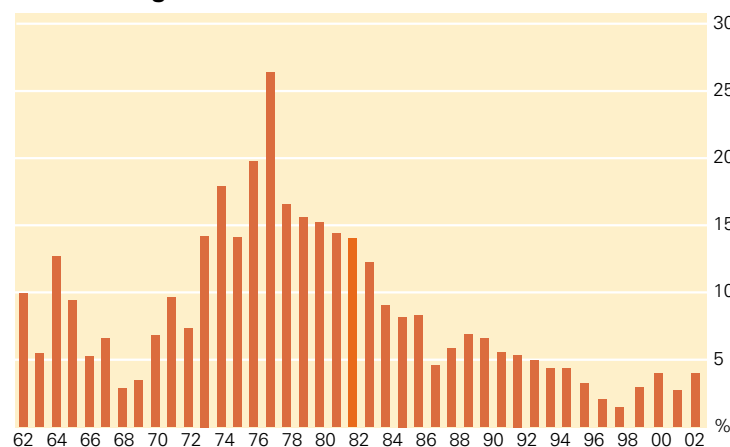
Índice de Precios de Consumo (IPC)

En enero de 2002 se implantó definitivamente el nuevo Sistema IPC-2001. Se ha conseguido un indicador más dinámico, la plena adaptación a la clasificación COICOP (Classification Of Individual Consumption by Purpose), ampliación de la muestra de municipios y establecimientos, actualización de la cesta de la compra (484 artículos), inclusión de precios rebajados y mejoras técnicas en el tratamiento de precios.

El índice se elabora con aproximadamente 200.000 precios, de los cuales informan cerca de 30.000 establecimientos distribuidos en 141 municipios.

La Encuesta Continua de Presupuestos Familiares (ECPF) proporcionó la información básica sobre los gastos de los hogares en bienes y servicios de consumo.

IPC. Índice general. Variación interanual (en diciembre)



Año 2002: variación interanual en octubre



Índice de precios de consumo armonizado

Variación interanual

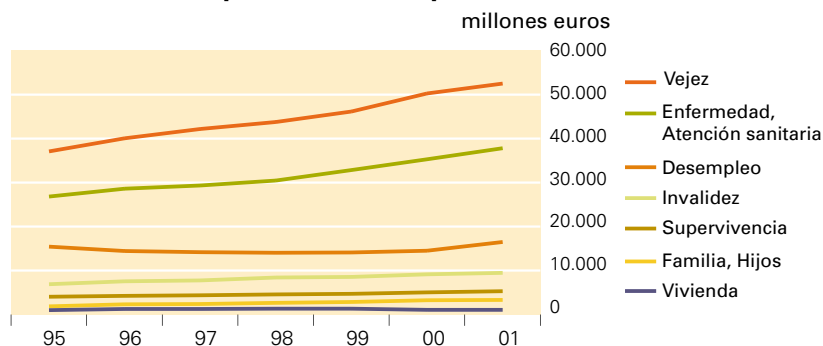
	2000	2001	Oct/01 Oct/02
UE-15	2,1	2,3	2,1
Alemania	2,1	2,4	1,3
Austria	2,0	2,3	1,8
Bélgica	2,7	2,4	1,3
España	3,5	2,8	4,0
Finlandia	3,0	2,7	1,7
Francia	1,8	1,8	1,9
Grecia	2,9	3,7	3,9
Irlanda	5,3	4,0	4,4
Italia	2,6	2,3	2,8
Luxemburgo	3,8	2,4	2,5
Países Bajos	2,3	5,1	3,6
Portugal	2,8	4,4	4,1
UNIÓN MONETARIA	2,3	2,5	2,3
Dinamarca	2,7	2,3	2,7
Reino Unido	0,8	1,2	1,4
Suecia	1,3	2,7	1,7

Fuente: Eurostat

Prestaciones de protección social. 2001

	Euros	Variación interanual
Prestaciones de protección social por habitante	3.151	5,44
Prestaciones de asistencia sanitaria por habitante	778	6,01
Prestaciones en dinero por persona de 65 años o más	7.500	2,84
Prestaciones de <i>desempleo</i> en dinero por persona en situación de paro	9.571	17,31

Evolución de las prestaciones de protección social



Gastos totales de protección social. 2000 (% del PIB)

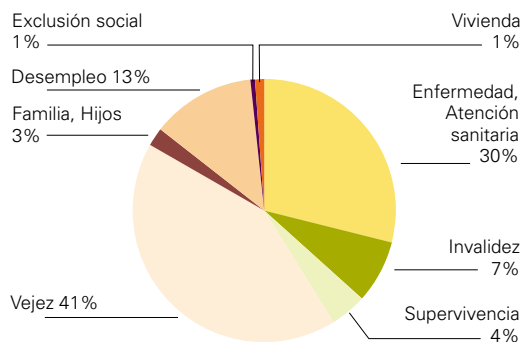
UE-15 ¹	27,5
Suecia	32,3
Francia ¹	30,2
Alemania	29,5
Dinamarca	28,8
Austria	28,7
Países Bajos	27,4
Bélgica	26,7
Reino Unido ¹	26,6
Grecia	26,4
Italia	25,2
Finlandia	25,2
Portugal ¹	22,7
Luxemburgo	21,0
España	20,1
Irlanda	14,1
1 Año 1999	

Fuente: Eurostat

Pensiones del sistema de la Seguridad Social. 2001

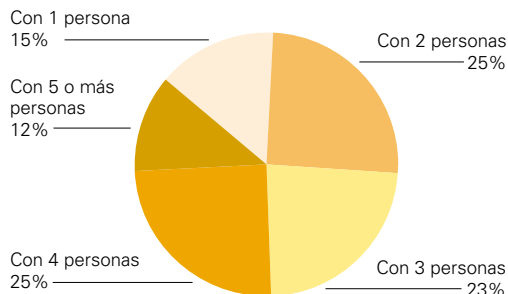
	Miles pensiones	€ / mes
Pensiones contributivas	7.677,9	494,5
Incapacidad permanente	788,6	571,2
Jubilación	4.545,6	565,0
Viudedad	2.042,4	348,8
Orfandad	257,8	210,6
Favor familiar	43,4	262,2
Pensiones no contributivas	481,6	241,3

Gastos en prestaciones de protección social. 2001. Total 126.881 millones €



Hogares según el número de personas

Total: 13.457,7 miles de hogares



Otras fuentes

Las cuentas de protección social las elabora el Ministerio de Trabajo y Asuntos Sociales y están adaptadas al sistema diseñado por EUROSTAT para hacer posible la comparación entre los países miembros.

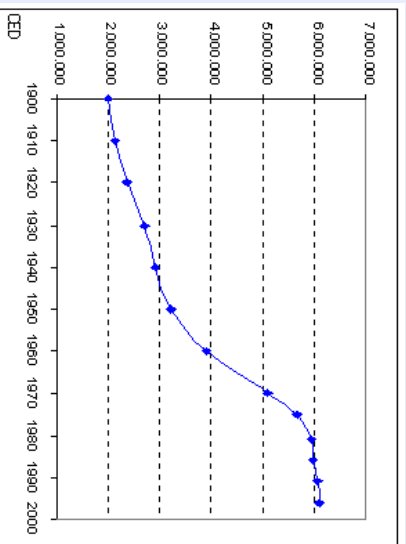


Catalunya

- [Presentació](#) - [Docència](#) - [Personal](#) - [Recerca](#) - [Publicacions](#) - [Activitats](#) - [Dades](#) - [Biblió](#) - [Llin](#)
- [Evolució de la població al segle XX](#)
- [Evolució recent de la població](#)
- [Piràmide població any 2000](#)
- [Estructura per edat](#)
- [Moviment natural](#)
- [Mirigració](#)

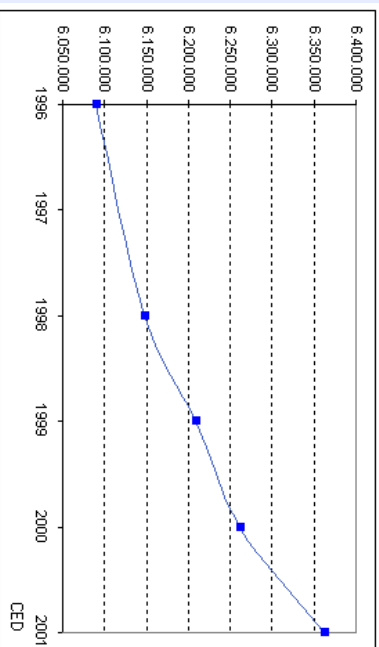
Tomar a selecció d'àmbit

Evolució de la població (1900-1996)



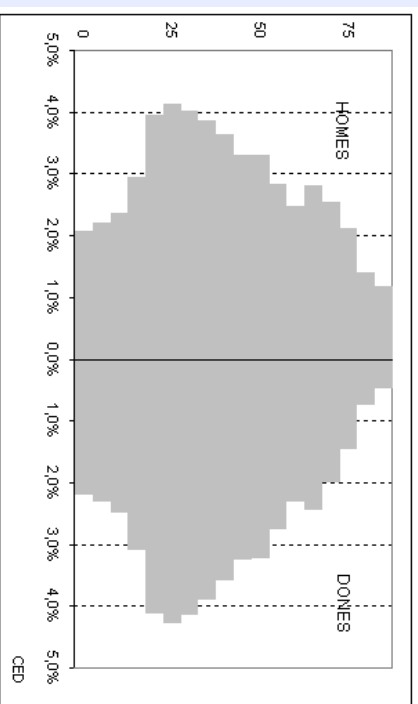
Font: Censos i padrons

Evolució recent de la població



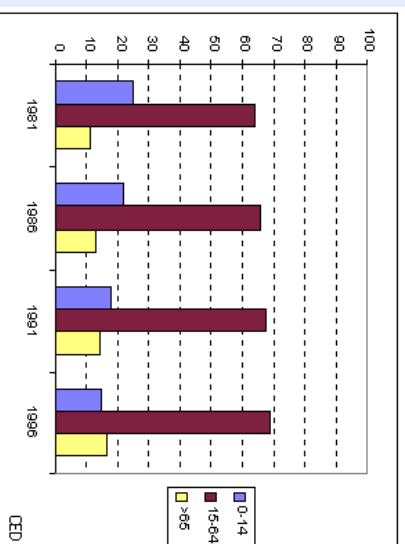
Font: Padró Continu, INE

Piràmide de població 2000



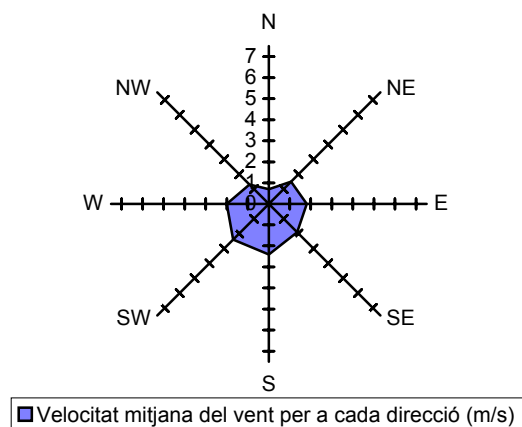
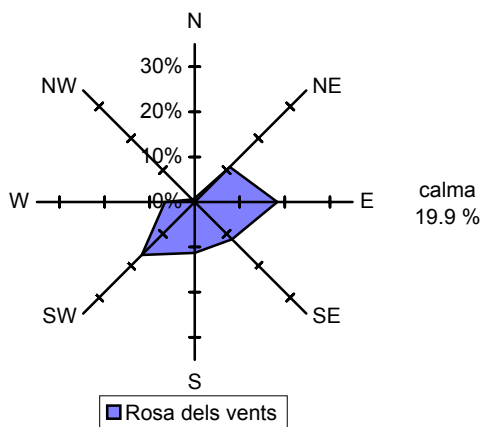
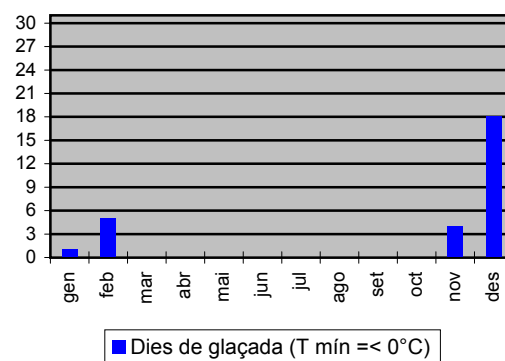
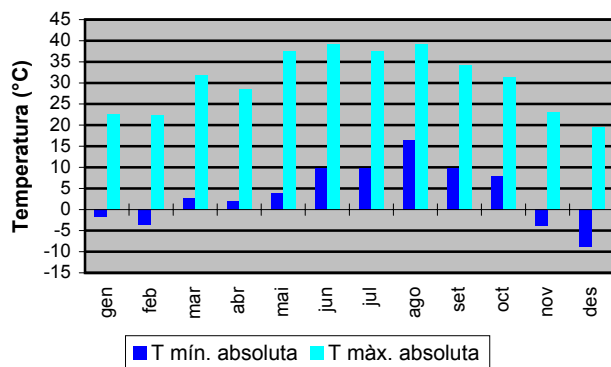
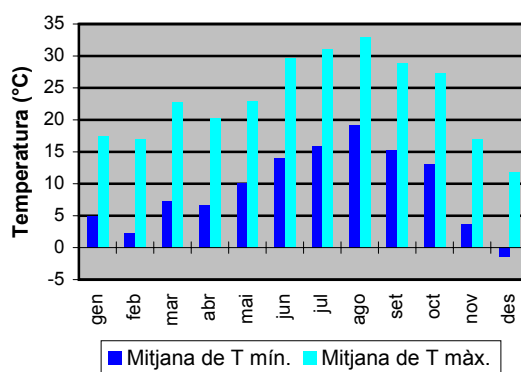
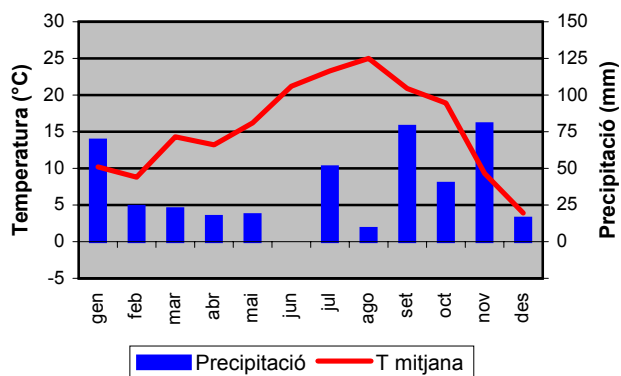
Font: Padró Continu, INE

Estructura de la població per grans grups d'edat (1981-1996)



Font: Censos i padrons i WEB Institut d'Estadística de Catalunya

MONTMELÓ (el Vallès Oriental)



Resum any 2001

Precipitació total acumulada:	428,2 mm
Temperatura mitjana:	15,4 °C
Mitjana de les temperatures màximes:	23,2 °C
Mitjana de les temperatures mínimes:	9,2 °C
Temperatura màxima absoluta:	39,1 °C (25/06/01)
Temperatura mínima absoluta:	-8,8 °C (25/12/01)
Velocitat mitjana del vent:	1,6 m/s
Direcció dominant:	E
Humitat relativa mitjana:	68 %
Irradiació global mitjana diària:	

0.5 Annex B: Exemples d'inferència

Font: IDESCAT, Baròmetre polític, (idescat.net)

0.5.1 Fitxa tècnica d'una enquesta



Quadre-resum:

Nivell de confiança:	95,5% (2 sigmes)
Variància:	$p=q=0,5$
Grandària de la mostra:	1.900 individus
Marge d'error (conjunt de la mostra)	$\pm 2'94\%$

Àmbit:

Catalunya

Univers:

Població catalana de 18 i més anys resident a Catalunya

Mètode de recollida de la informació:

Enquesta telefònica assistida per ordinador (mètode C.A.T.I.).

Procediment de mostreig:

Estratificat per província i dimensió de municipi, amb selecció de les unitats primàries de mostreig (municipis) de forma aleatòria proporcional, i de les unitats últimes (individus) aleatòriament per quotes de sexe i edat.

La grandària dels municipis s'ha dividit en 6 categories:

- menor o igual a 2.000 habitants
- de 2.001 a 10.000
- de 10.001 a 50.000
- de 50.001 a 150.000
- de 150.001 a 1.000.000
- més d'1 milió

Les quotes d'edat han estat:

- de 18 a 34 anys
- de 35 a 49 anys
- de 50 a 64 anys
- de 65 anys o més

Per al càlcul de les quotes s'ha utilitzat com a base el Padró de Població 2004.

0.5.2 Alguns resultats de l'enquesta



34. Ara li citaré el nom d'alguns líders polítics. Em podria dir, per a cadascun d'ells, si els coneix? I si el coneix, com valora l'actuació política d'aquests líders. (NOMÉS UNA RESPOSTA)

Coneix a Pasqual Maragall

	Percentatge	Freqüències
Coneix	96,7	1837
No coneix	1,0	20
No ho sap	0,4	7
No contesta	1,9	36
Total	100	1900

Puntuació per Pasqual Maragall

	Percentatge	Freqüències
[0] Molt deficient	7,2	133
[1]	2,7	50
[2]	2,9	53
[3]	6,9	128
[4]	10,4	192
[5]	21,8	401
[6]	15,9	291
[7]	12,2	224
[8]	9,0	166
[9]	3,0	55
[10] Excel·lent	3,1	58
No ho sap	4,2	78
No contesta	0,4	8
Mitjana	5,16	5,16
Desviació típica	2,42	2,42
Total coneix Pasqual Maragall	100	1837