5° Convegno Italiano degli Utenti di Stata

5° Italian Stata User Group

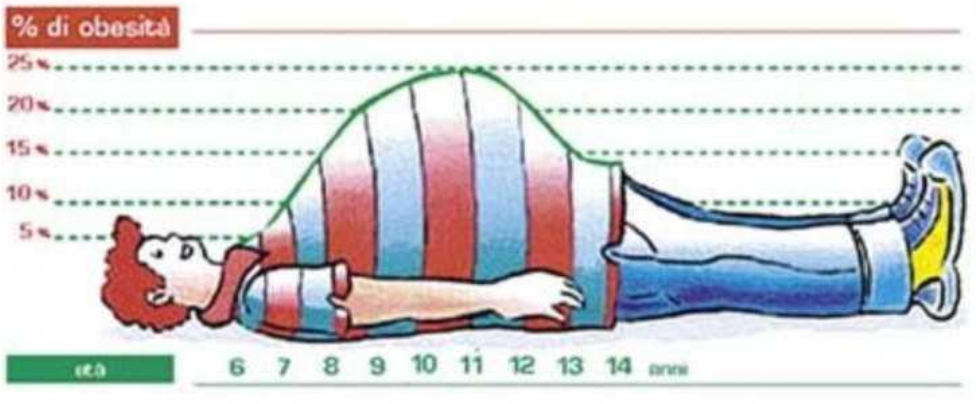
Milano 20-21 ottobre 2008

New wine in old bottles: visualizing the progression over time of the epidemics of tobacco smoking and obesity through the use of population pyramids

Giovanni Capelli (1), Bruno Federico (1), Giuseppe Costa (2)

(1) Dipartimento di Scienze Motorie e della Salute, Università di Cassino

(2) Dipartimento di Sanità Pubblica e Microbiologia, Università di Torino



Tobacco and obesity epidemics

- Tobacco smoking and obesity appear as different phenomena
  - Anyway, similar pyschologic, social and environmental factors, among which media pressure, have an influence on the exposure to these risk factors
- In last decade, smoking prevalence has shown a decrease in many industrialized countries, whereas obesity showed pandemia-like patterns
  - Public Health programs aimed to contrast obesity may take advantage of the experiences in the field of tobacco consumption control

Tobacco and obesity epidemics: not so different after all?

Mickey Chopra and Ian Darnton-Hill

BMJ 2004;328;1558-1560
doi:10.1136/bmj.328.7455.1558

Surveillance on behavioural risk factors

- National and International governmental agencies, epidemiological agencies routinely evaluate prevalence of these risk factors through
  - Standardized rates
    - Fictitious values
    - Strictly depends on the reference population
  - Specific sex and age class rates
    - Hide informations on the total number of at-risk subjects

Sources of data


MOD. ISTAT/IMF-8/A.04-05

ISTAT

SISTEMA STATISTICO NAZIONALE
ISTITUTO NAZIONALE DI STATISTICA

INDAGINE STATISTICA MULTISCOPO SULLE FAMIGLIE

CONDIZIONI DI SALUTE E RICORSO AI SERVIZI SANITARI 2004 - 2005

A MESE

1 Provincia
Comune
Sezione di Censimento

2 *(a cura del Comune)*
Numero d'ordine definitivo della famiglia
Da 0001 al totale dei modelli ISTAT/IMF-8/A compilati dal Comune nel trimestre

3 Numero d'ordine della famiglia nel Mod. ISTAT/IMF-8/2
Da 00001 al totale delle famiglie del Mod. ISTAT/IMF-8/2 nel trimestre

- Exposure data:
  - Raw data from Italian National Statistics Institute (ISTAT) Surveys on Health of the Italian Population (1983, 1990, 1994, 1999-2000 e 2004-2005)
    - Large samples of italian non-istitutionalized population (50,000 - 140,000 enrolled)
    - Direct interviews and self-complied questionnaires
    - About 10% of non-responders
- Population data
  - Annual estimates of the italiana poipulation by age and sex (www.istat.it)


demo istat.it — A cura di

Demografia in Cifre

Istat — English

popolazione residente

Popolazione Residente
per età, sesso e stato civile al 1° gennaio
Anno 2007
Anno 2006
Anno 2005
Anno 2004
Anno 2003
Anno 2002

bilancio demografico

Bilancio Demografico
e popolazione residente per sesso al 31 dicembre
Anno 2006
Anno 2005
Anno 2004
Anno 2003
Anno 2002

Bilancio Demografico Mensile
e popolazione residente per sesso
Anno 2007
Anno 2006
Anno 2005
Anno 2004
Anno 2003

L'ISTAT mette a disposizione i dati ufficiali più recenti sulla popolazione residente nei Comuni italiani derivanti dalle indagini effettuate presso gli Uffici di Anagrafe. Interrogazioni personalizzate (per anno, territorio, cittadinanza, ecc.) permettono di costruire le tabelle di interesse e scaricare i dati in formato rielaborabile.
È possibile trovare anche informazioni sui principali fenomeni demografici, come i tassi di natalità e mortalità, le previsioni della popolazione residente, l'indice di vecchiaia, l'età media.

Ultimi aggiornamenti

11 dicembre 2007 - **Popolazione residente**
Popolazione residente per età, sesso e stato civile al 1° gennaio 2007

11 dicembre 2007 - **Cittadini stranieri**
Popolazione straniera residente per età e sesso al 1° gennaio 2007

14 novembre 2007 - **I permessi di soggiorno**
I dati aggiornati al 1° gennaio 2007

5 novembre 2007 - **I matrimoni**
Dati relativi all'anno 2005

elaborazioni

Ricostruzione Intercensuaria della popolazione
per età e sesso al 1° gennaio
Anni 1992-2001
Anni 1982-1991

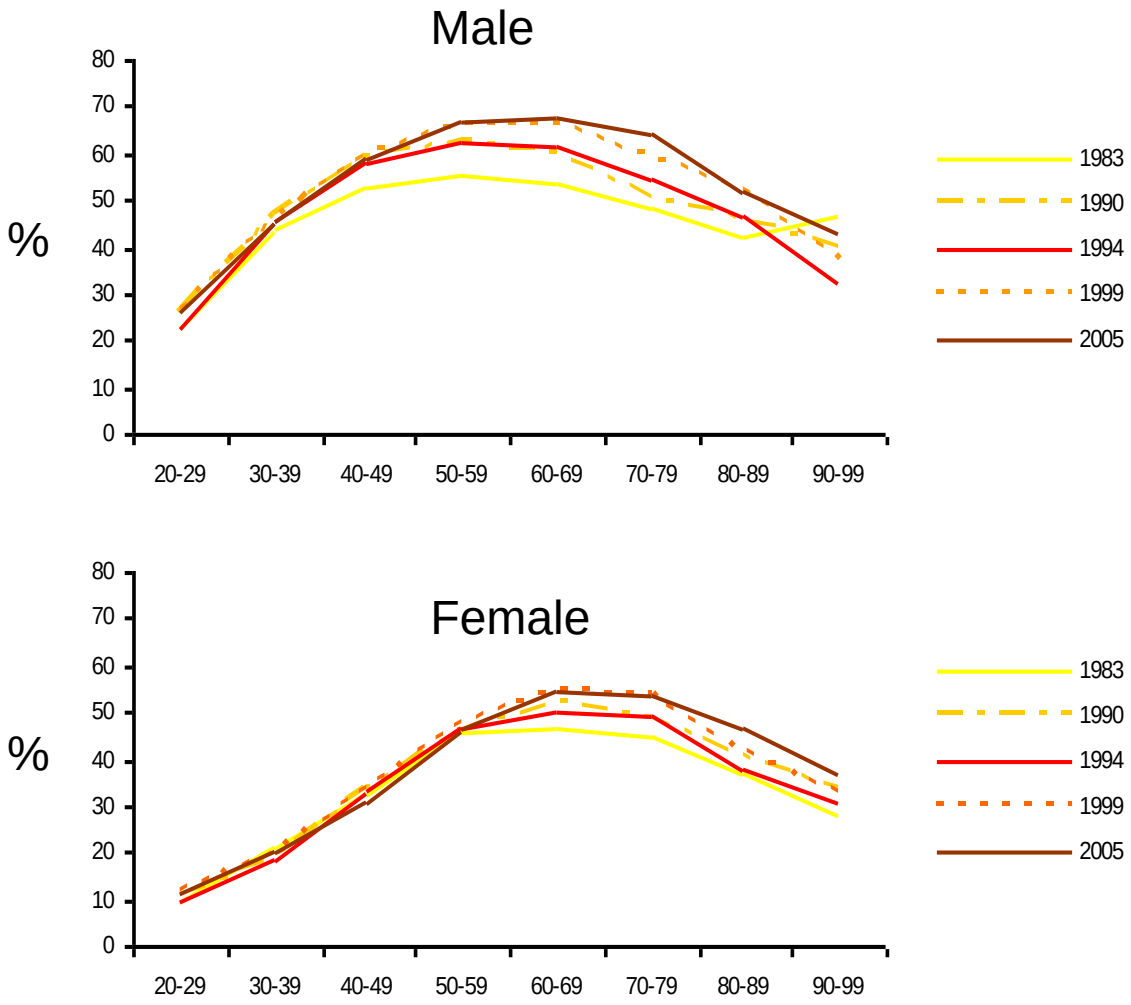
Ricostruzione Intercensuaria del bilancio demografico
per sesso
Anni 1991 - 2001

Previsioni della Popolazione
per età, sesso e regione
Anni 2001-2051

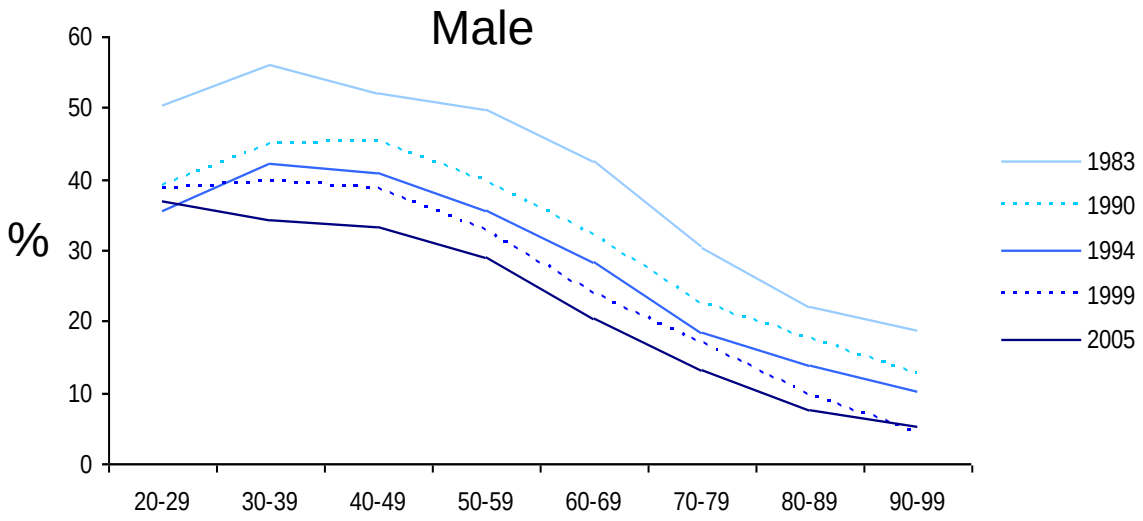
Tavole di Mortalità
della popolazione italiana per provincia e regione di residenza
Anno 2004
Anno 2003
Anno 2002
Anno 2001

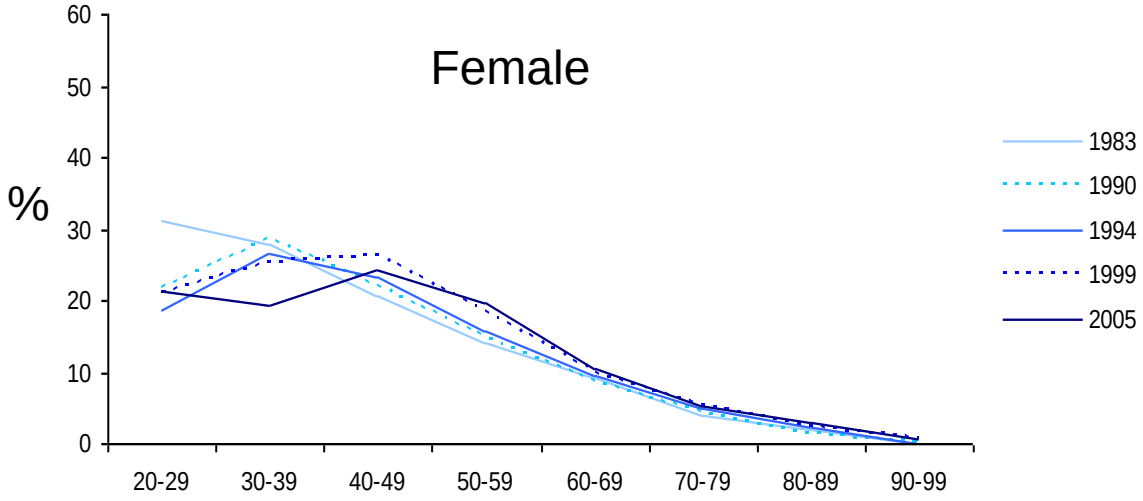
Tavole di Fecondità
della popolazione italiana per regione di residenza
Anni 1952-2004

Sex- and age- specific prevalence rates (overweight+obese)



Sex- and age- specific prevalence rates (tobacco use)





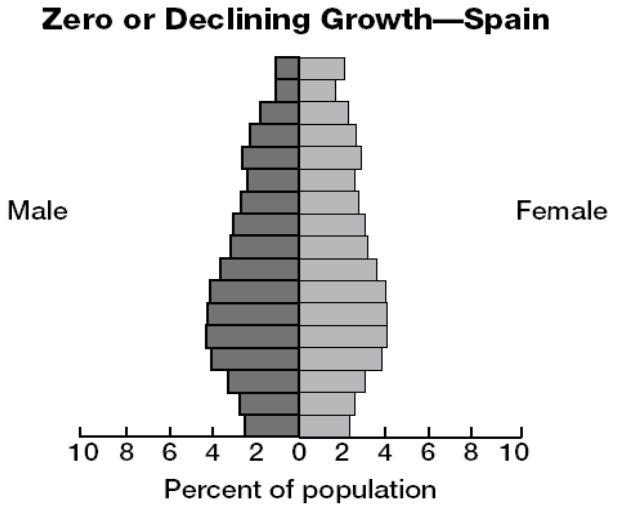
Age-standardized rates (obese/overweight)

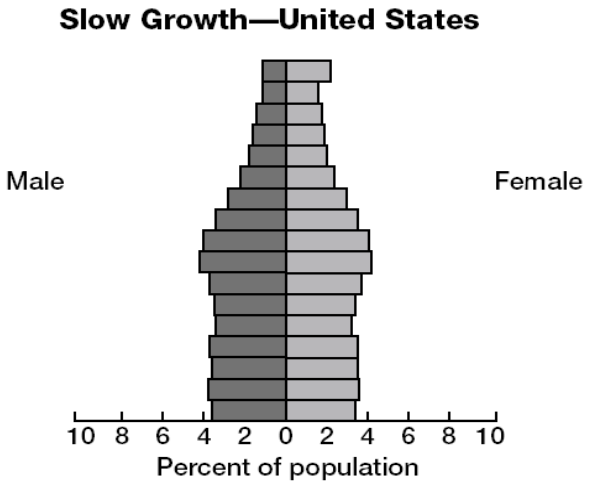
	Male		Female	
	% obese	**% overwgt+ obese**	**% obese**	**% overwgt+ obese**
1983	6.8	45.3	7.6	31.8
1990	7.9	48.9	7.2	31.7
1994	8.5	49.0	7.6	31.8
2000	10.3	52.4	9.2	35.3
2005	10.6	54.0	9.2	36.0

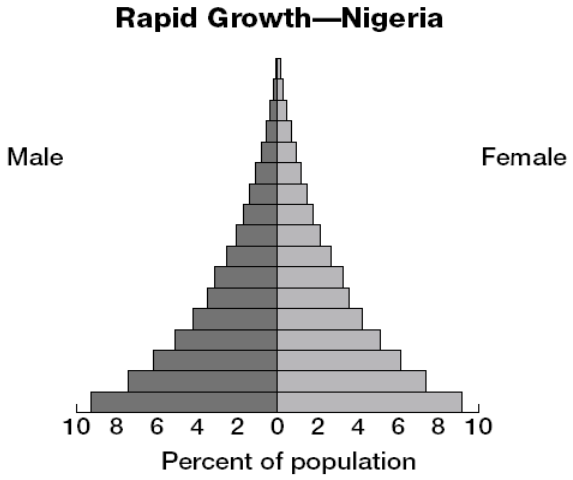
Age-standardized rates (Tobacco use)

	Male		Female	
	% smokers	**% smok+ former**	**smokers**	**% smok+ former**
1983	46.6	63.9	17.8	20.5
1990	37.2	64.6	16.9	23.7
1994	33.5	NA	16.7	NA
2000	31.5	62.4	17.8	30.4
2005	27.6	60.4	16.8	31.9

What about using… the old good population pyramids?







- An historical method to see "at a glance" the overall structure of a population
  - Can we present exposure factor as part of the structure?

```
.help twoway_bar

Advanced use:  Population pyramid
    We have the following aggregate data from the U.S. 2000
    Census recording total population by age and
    sex.  From this, we produce a population pyramid:
        . sysuse pop2000, clear
        . list agegrp maletotal femtotal
```

	agegrp	maletotal	femtotal
1.	Under 5	9,810,733	9,365,065
2.	5 to 9	10,523,277	10,026,228
3.	10 to 14	10,520,197	10,007,875
4.	15 to 19	10,391,004	9,828,886
5.	20 to 24	9,687,814	9,276,187
6.	25 to 29	9,798,760	9,582,576
7.	30 to 34	10,321,769	10,188,619
8.	35 to 39	11,318,696	11,387,968
9.	40 to 44	11,129,102	11,312,761
10.	45 to 49	9,889,506	10,202,898
11.	50 to 54	8,607,724	8,977,824
12.	55 to 59	6,508,729	6,960,508
13.	60 to 64	5,136,627	5,668,820
14.	65 to 69	4,400,362	5,133,183
15.	70 to 74	3,902,912	4,954,529
16.	75 to 79	3,044,456	4,371,357
17.	80 to 84	1,834,897	3,110,470

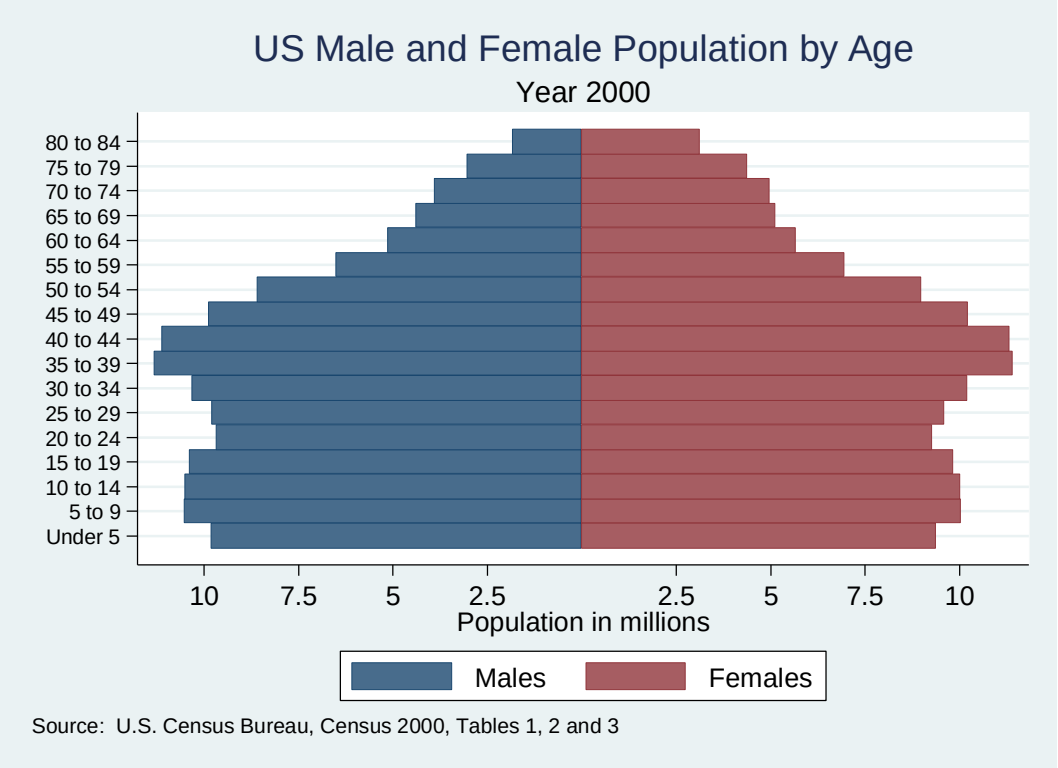
```
        . replace maletotal = -maletotal/1e+6
        . replace femtotal = femtotal/1e+6
```

Population pyramids with Stata

see [G] graph twoway bar, pag 181

```
. twoway
                bar maletotal agegrp, horizontal xvarlab(Males)
          ||
                bar  femtotal agegrp, horizontal xvarlab(Females)
          ||
          , ylabel(1(1)17, angle(horizontal) valuelabel labsize(*.8))
            xtitle("Population in millions") ytitle("")
            xlabel(-10 "10" -7.5 "7.5" -5 "5" -2.5 "2.5" 2.5 5 7.5 10)
            legend(label(1 Males) label(2 Females))
            title("US Male and Female Population by Age")
            subtitle("Year 2000")
            note("Source:  U.S. Census Bureau, Census 2000, Tables 1, 2 and 3",
                  span)
```

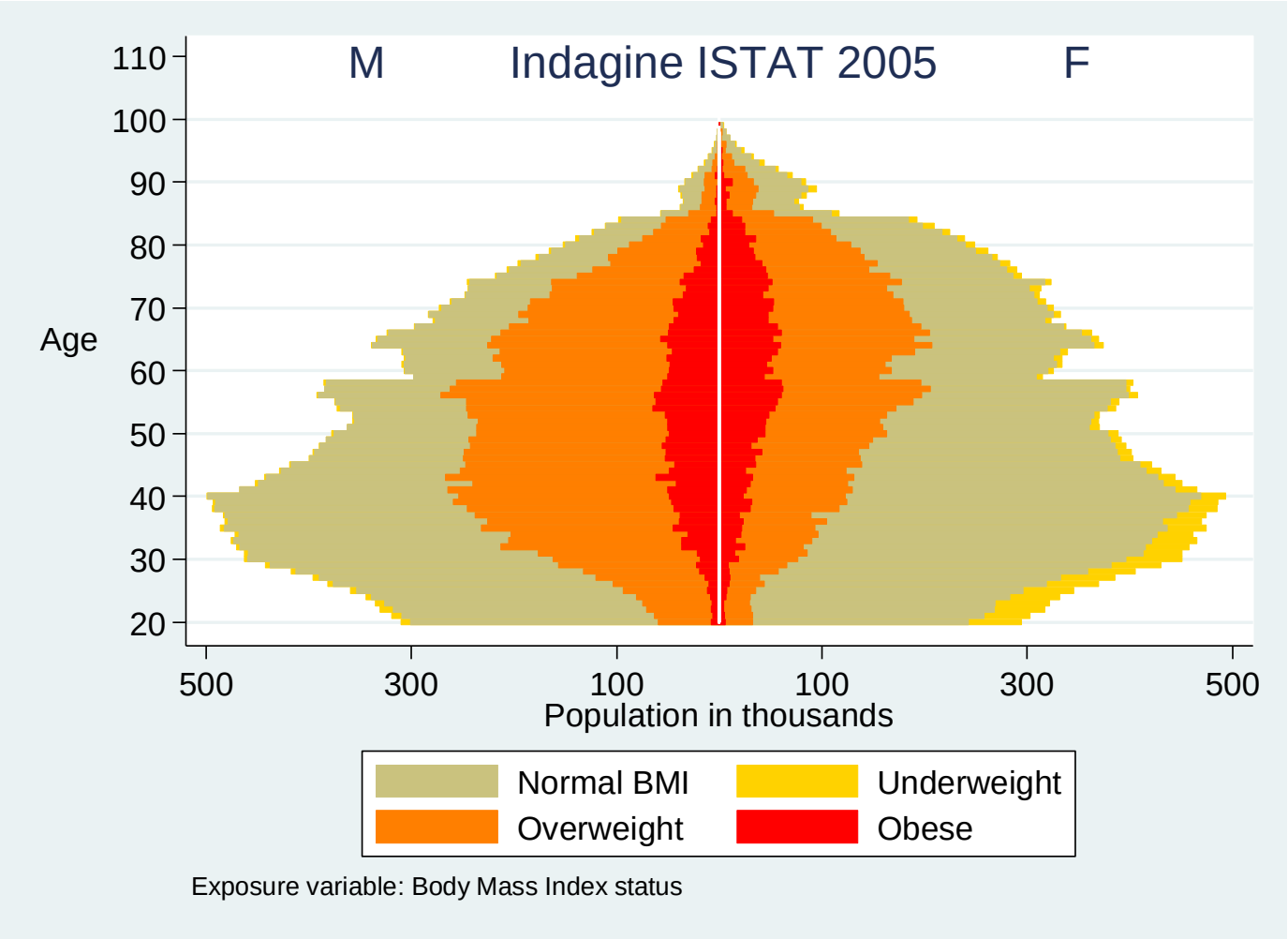
see [G] graph
twoway bar, pag 181



Our Aim

- To evaluate the diffusion of tobacco use and obesity/overweight
  - Using a graphic method easy readable
  - Which can show "at a glance" the **overall impact** of the risk factors in the population
    - Visualizing not only the proportion but the numerical quantity of people in each status
- A modified form of the population pyramid, subidiving each age class in exposure strata, can be used
  - Realizing something like a "section" of the population age and sex structure on the plan of the exposure factor

The “stratified” population pyramid (overweight/obese)



This plot **adds to the pyramid a third dimension**: you can see at one time the distribution by age, sex and exposure level of the studied population

We are starting with detail data like these:

```
. Describe

  obs:       564,828
 vars:             6                          21 Dec 2007 16:49
 size:     7,907,592 (24.6% of memory free)
-------------------------------------------------------------------------------
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
sesso           byte   %8.0g       sessolbl   SESSO
eta             int    %9.0g
fumo            byte   %8.0g       fumolbl    ABITFUMO
peso            int    %8.0g                  PESO
statura         int    %8.0g                  STATURA
anno            int    %9.0g
-------------------------------------------------------------------------------
```

How to do ?

```
. list in 1/10

     +-------------------------------------------------+
     |  sesso   eta      fumo   peso   statura   anno |
     |-------------------------------------------------|
  1. |   male    68   current     58       165   1983 |
  2. |   male    77   current     36       155   1983 |
  3. | female    68     never     50       160   1983 |
  4. |   male    53    former     70       165   1983 |
  5. | female    47     never     58       160   1983 |
  6. |   male    44     never     65       165   1983 |
  7. | female    36     never     66       165   1983 |
  8. | female    14     never     50       150   1983 |
  9. |   male    11     never     40       130   1983 |
 10. |   male    32   current     72       168   1983 |
     +-------------------------------------------------+
```

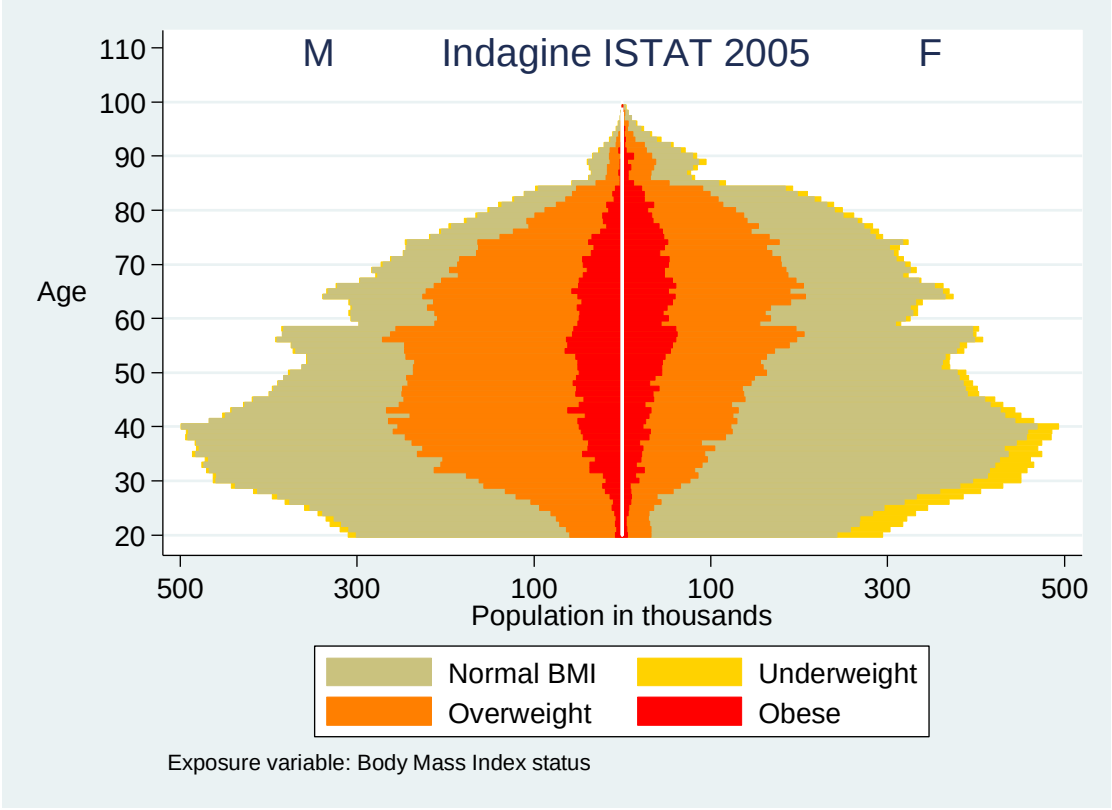
```
. tab anno sesso

           |         SESSO
      anno |      male     female |     Total
-----------+----------------------+----------
      1983 |    43,568     46,197 |    89,765
      1987 |    37,772     39,379 |    77,151
      1990 |    33,025     34,375 |    67,400
      1994 |    30,631     31,830 |    62,461
      2000 |    68,342     71,669 |   140,011
      2005 |    61,917     66,123 |   128,040
-----------+----------------------+----------
     Total |   275,255    289,573 |   564,828
```

```
gen bmi=peso/((statura/100)^2)
drop if bmi==.
gen clasbmi=bmi
recode clasbmi min/18.5=0 18.50001/24.999999=1 25/29.9999=2 30/max=3
label define clasbmilbl 0 underweight 1 normal 2 overweight 3 obese
label values clasbmi clasbmilbl
gen underweight=clasbmi==0
gen normal=clasbmi==1
gen overweight=clasbmi==2
gen obese=clasbmi==3
drop if eta<20 | eta>99

collapse (count) bmi, by(eta sesso clasbmi)
gen onumbmiclas=bmi
replace onumbmiclas=-bmi if sesso==1
drop bmi

reshape wide onumbmiclas, i(eta sesso) j(clasbmi)
mvencode onumbmiclas*, mv(0)

gen tot=onumbmiclas3+onumbmiclas2+onumbmiclas1+onumbmiclas0
local i=0
while `i' <4 {
    gen percclas`i'=onumbmiclas`i'/tot
    replace percclas`i'=-percclas`i' if sesso==1
    local i=`i'+1
    }
gen linkcode=sesso*1000+eta
sort linkcode
```

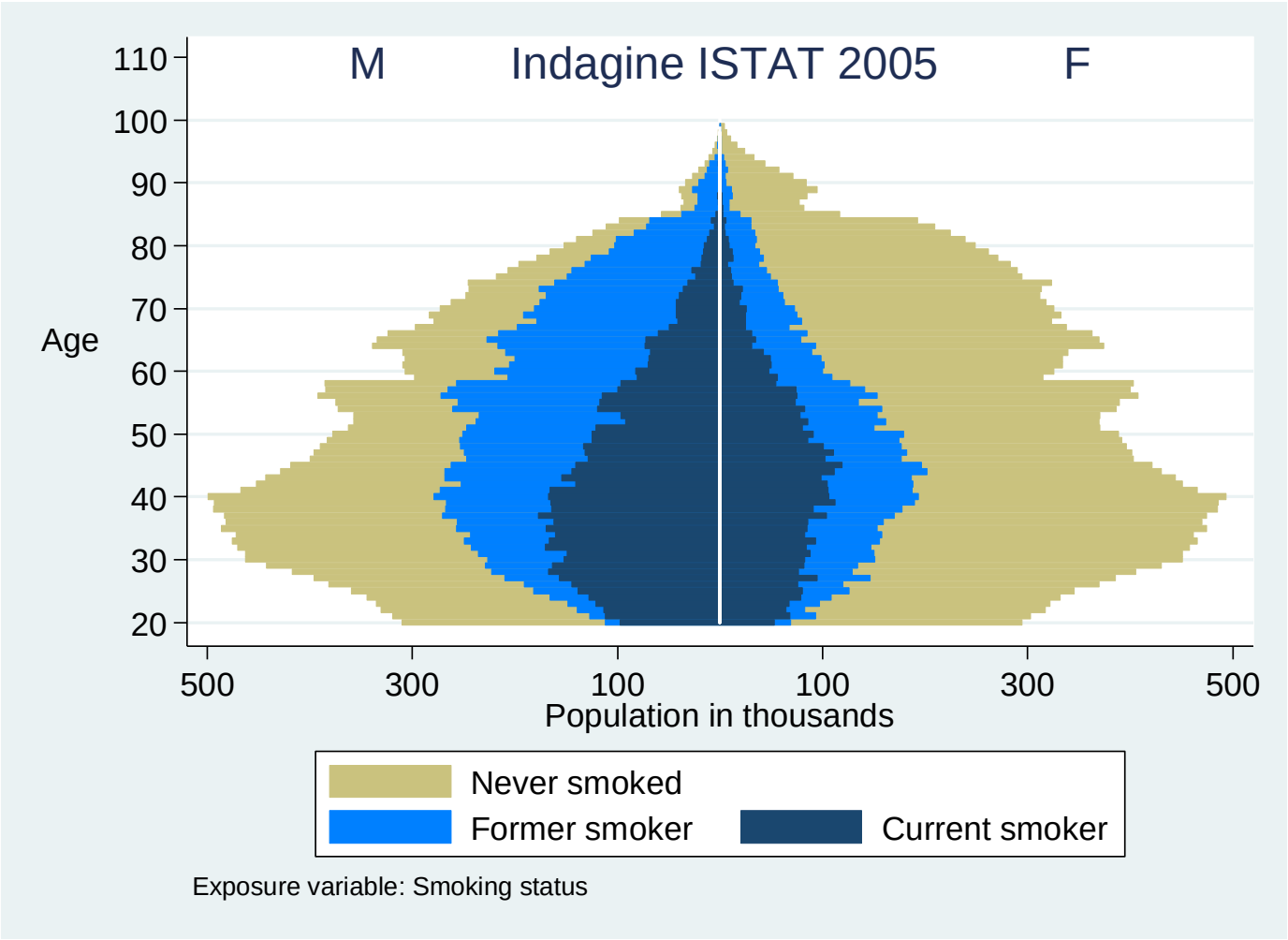
```
. describe

Contains data
  obs:           159
 vars:            12
 size:         7,473 (99.9% of memory free)
-------------------------------------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------------------------------------
sesso           byte   %8.0g       sessolbl   SESSO
eta             int    %9.0g
onumbmiclas0    float  %9.0g                  0 onumbmiclas
onumbmiclas1    float  %9.0g                  1 onumbmiclas
onumbmiclas2    float  %9.0g                  2 onumbmiclas
onumbmiclas3    float  %9.0g                  3 onumbmiclas
tot             float  %9.0g
percclas0       float  %9.0g
percclas1       float  %9.0g
percclas2       float  %9.0g
percclas3       float  %9.0g
linkcode        float  %9.0g
-------------------------------------------------------------------------------------------------------------
Sorted by:  linkcode
     Note:  dataset has changed since last saved

. list in 1/2

     +-------------------------------------------------------------------------------------------------------+
  1. | sesso | eta | onumbm~0 | onumbm~1 | onumbm~2 | onumbm~3 |  tot | percclas0 | percclas1 | percclas2 |
     |  male |  20 |      -23 |     -579 |     -123 |      -19 | -744 |  -.030914 | -.7782258 | -.1653226 |
     |-------------------------------------------------------------------------------------------------------|
     |                   percclas3                     |                   linkcode                      |
     |                   -.0255376                     |                       1020                      |
     +-------------------------------------------------------------------------------------------------------+

     +-------------------------------------------------------------------------------------------------------+
  2. | sesso | eta | onumbm~0 | onumbm~1 | onumbm~2 | onumbm~3 |  tot | percclas0 | percclas1 | percclas2 |
     |  male |  21 |      -21 |     -527 |     -122 |      -14 | -684 | -.0307018 | -.7704678 | -.1783626 |
     |-------------------------------------------------------------------------------------------------------|
     |                   percclas3                     |                   linkcode                      |
     |                   -.0204678                     |                       1021                      |
     +-------------------------------------------------------------------------------------------------------+
```

```
merge linkcode using popitaly.dta , keep(anno`anno')
keep if _merge==3

local i=0
while `i' <4 {
    gen numbmiclas`i'=percclas`i'*anno`anno'
    local i=`i'+1
    }
gen obese=(numbmiclas3)/1000
gen overweight=(numbmiclas3+numbmiclas2)/1000
gen normal=(numbmiclas3+numbmiclas2+numbmiclas1)/1000
gen underweight=(numbmiclas3+numbmiclas2+numbmiclas1+numbmiclas0)/1000

label variable obese "Obese"
label variable overweight "Overweight"
label variable normal "Normal BMI"
label variable underweight "Underweight“

local m=500
local m2=300
local m3=100

gen zero=0
```

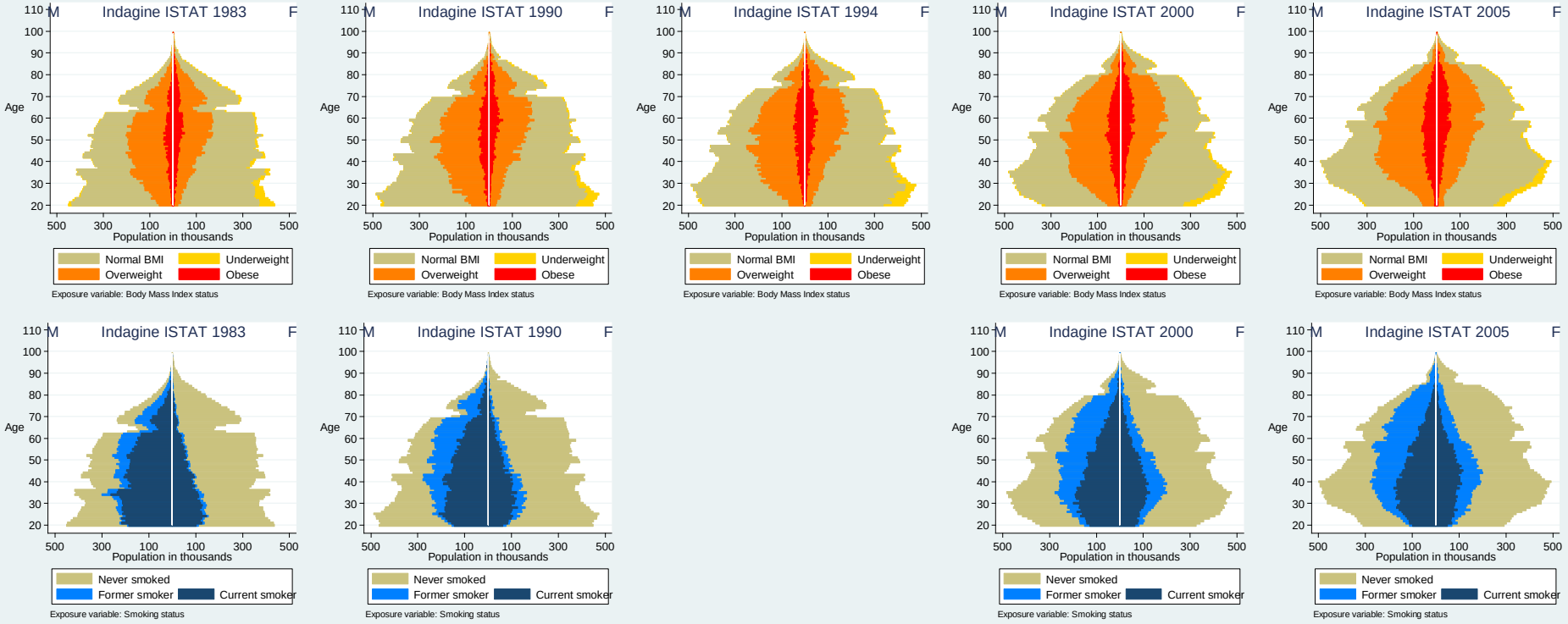
```
twoway (bar  underweight normal overweight obese eta, horizontal
   bcolor(gold khaki orange red) fintensity(100 100 100 100) blwidth(0 0 0
   0)) (line eta zero, lcolor(white)), ylabel(20(10)110, angle(horizontal))
   xlabel(-`m' "`m'" -`m2' "`m2'" -`m3' "`m3'" `m3'  `m2' `m')
   xline(0,lcolor(white)) legend(order(2 1 3 4) cols(2)) note("Exposure
   variable: Body Mass Index status") ytitle("Age",
   orientation(horizontal)) xtitle("Population in thousands") title("M
   Indagine ISTAT `anno'          F", ring(0))
   saving(bmiclaspopestimate_`anno'test.gph, replace)
```



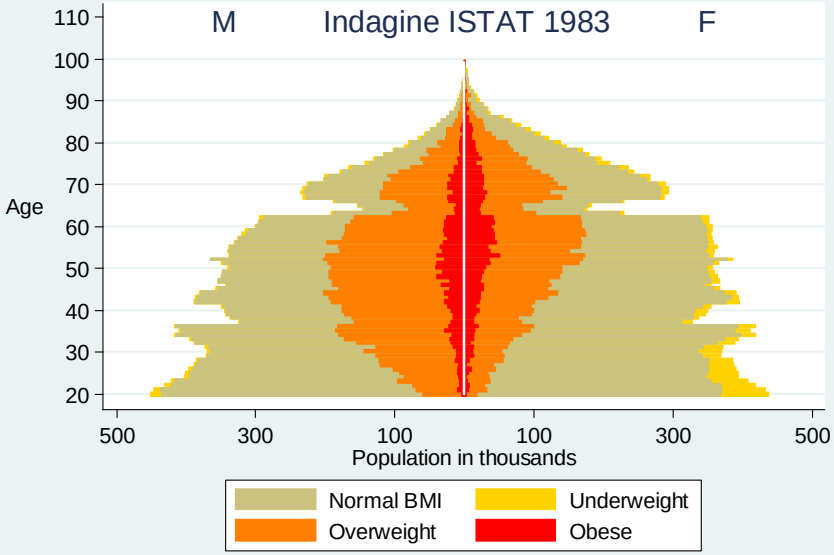
Stratified population pyramid for tobacco smoking



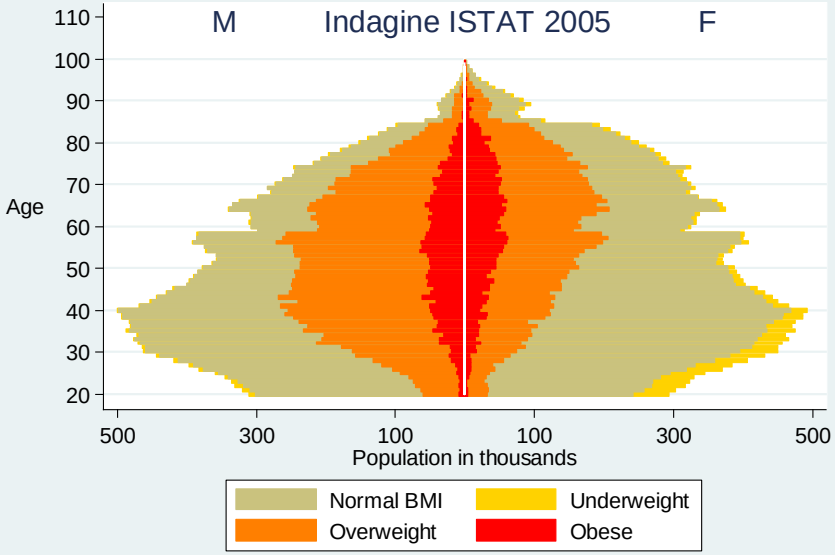
Time evolution of tobacco consumption and overweight in Italy (1983-2005)



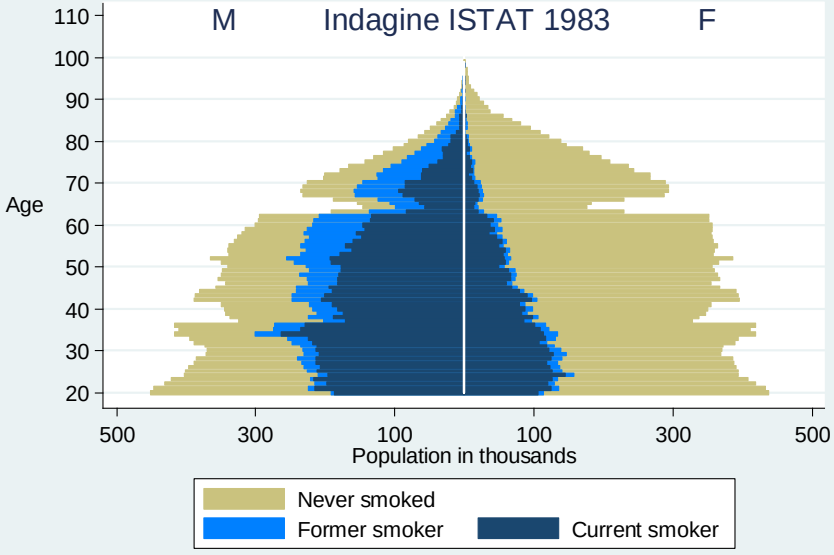

```
graph combine bmiclaspopestimate_1983test.gph bmiclaspopestimate_1990test.gph
bmiclaspopestimate_1994test.gph bmiclaspopestimate_2000test.gph
bmiclaspopestimate_2005test.gph fumoclaspopestimate_1983test.gph
fumoclaspopestimate_1990test.gph fumoclaspopestimate_2000test.gph
fumoclaspopestimate_2005test.gph , rows(2) cols(5) holes(8) ycommon xcommon altshrink
xsize(12) ysize(5)
```



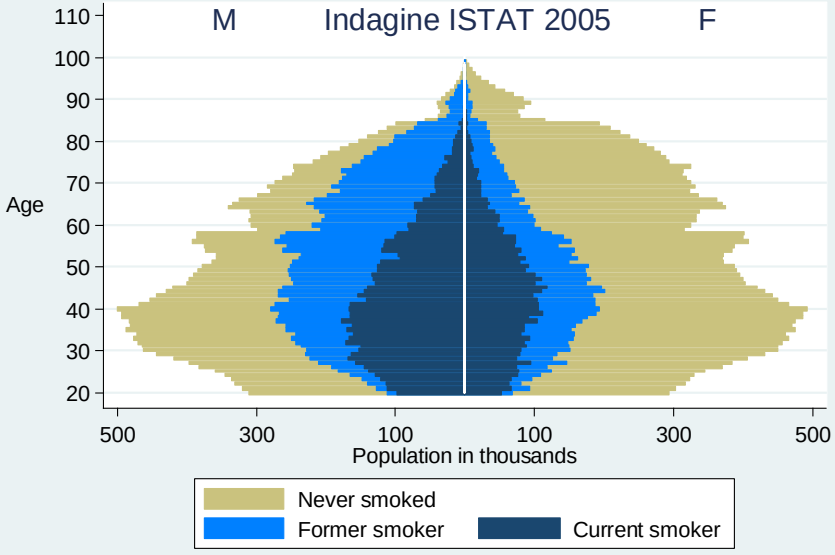
Exposure variable: Body Mass Index status



Exposure variable: Body Mass Index status



Exposure variable: Smoking status



Exposure variable: Smoking status

Summary of the main results

- Tobacco consumption and overweight affect a large proportion of Italian population
  - Up to 50% some age and sex class
- Tobacco consumption is decreasing in males not in females
  - Number of current smokers was reduced of about 20% (from13.3 to 10.5 millions)
- Obesity and overweight are in constant increase
  - Obese people is almost doubled (from 2.7 to 4.7 millions)
- Pyramids show a huge gender asymmetry
  - Look at the overweighted men in their ’40s and women in their ’70s

Discussion issues

- The proposed stratified pyramids show many data at same time, allowing the reader to:
  - See the absolute burden of risk factors in the structure of the population
  - Estimate the number of exposed subject for sex and age
- Any attribute or exposure level in the population can be reported in such form of pyramids. May apply also to:
  - Perceived health status, disability level, severity of disease, …
- The *small multiples* format allows immediate comparison of risk factor distribution in different nations/regions or different times