

# NOTA METODOLOGICA

## Strategia di campionamento e valutazione degli errori campionari<sup>1</sup>

### 1 - Introduzione

La *popolazione di interesse* dell'indagine è costituita dalle donne italiane di età compresa tra 16 e 75 anni residenti in Italia.

L'indagine ha la finalità di fornire stime con diversi riferimenti territoriali:

- l'intero territorio nazionale;
- le cinque ripartizioni geografiche (Nord-ovest, Nord-est, Centro, Sud e Isole);
- le regioni geografiche;
- sei aree basate sulla tipologia socio-demografica dei comuni, così definite:
  - A, area *metropolitana* suddivisa in :
    - A1, comuni centro dell'area metropolitana: Torino, Milano, Venezia, Genova, Bologna, Firenze, Roma, Napoli, Bari, Palermo, Catania e Cagliari;
    - A2, comuni che gravitano intorno al centro dell'area metropolitana;
  - B, area *non metropolitana* suddivisa in :
    - B1 , comuni aventi fino a 2 mila abitanti;
    - B2 , comuni con 2.001-10 mila abitanti;
    - B3 , comuni con 10.001-50 mila abitanti;
    - B4 , comuni con oltre 50 mila abitanti.

### 2. Disegno di campionamento

La *lista di campionamento* per la selezione del campione è costituita dall'archivio unificato delle liste anagrafiche comunali (LAC). Tale archivio contiene le informazioni degli individui e delle famiglie residenti sul territorio italiano: per ciascun individuo sono riportate, oltre alle variabili identificative – compreso il codice fiscale - l'indirizzo, la data di nascita, il sesso, la cittadinanza, e l'anno di iscrizione in anagrafe. Nella lista è stato pertanto possibile individuare le unità appartenenti alla popolazione di interesse.

È stato utilizzato un disegno a uno stadio stratificato. Gli strati sono stati definiti dall'incrocio della regione e della tipologia comunale, ottenendo un numero complessivo di strati pari a 104.

La numerosità campionaria è stata fissata in 21.000 donne italiane. Tale numerosità è stata allocata tra gli strati in un'ottica di compromesso (tra un'allocazione uniforme e una proporzionale tra i diversi domini di stima) per garantire livelli soddisfacenti degli errori campionari attesi a livello dei diversi domini di stima sopra descritti.

La selezione delle unità campionarie è stata preceduta dall'aggancio tra le unità della popolazione e la lista dei numeri di telefono disponibili e una successiva suddivisione delle donne di ciascuno strato in due sotto-strati sulla base della presenza o meno di almeno un numero di telefono

---

<sup>1</sup> A cura di Claudia De Vitiis, Marco Dionisio Terribili e David Trambusti

agganciato. La numerosità campionaria prevista per ognuno dei 104 strati è stata quindi ripartita proporzionalmente tra i due sotto-strati e l'estrazione degli indirizzi campione da ciascun sotto-strato è stata effettuata con probabilità uguali e senza reimmissione, mediante tecnica di selezione sistematica. Per ogni nominativo selezionato sono stati selezionati altri cinque nominativi dallo stesso strato da utilizzare come sostituzioni in caso di mancato contatto o non disponibilità all'intervista.

Alle donne selezionate è stata inviata una lettera nella quale si chiedeva di confermare il numero presente nella lista oppure fornire un numero di telefono al quale essere contattate per l'intervista. Sulla base dei recapiti complessivamente ottenuti, sono state formate delle quartine di nominativi per le interviste.

Nel Prospetto 1 sono riportate le numerosità campionarie delle donne rispondenti per regione.

#### **Prospetto 1 - Distribuzione regionale del campione delle donne italiane rispondenti (16-75 anni)**

<b>Regioni</b>	<b>Popolazione</b>	<b>Campione</b>
Piemonte	1.405.232	1.043
Valle d'Aosta	42.374	722
Lombardia	3.241.144	1.488
Bolzano	173.209	449
Trento	182.490	738
Veneto	1.597.412	1.070
Friuli-Venezia Giulia	391.337	800
Liguria	494.658	806
Emilia-Romagna	1.439.641	1.021
Toscana	1.193.467	957
Umbria	278.750	727
Marche	497.227	795
Lazio	1.928.937	1.093
Abruzzo	439.718	741
Molise	103.737	664
Campania	2.052.425	778
Puglia	1.423.057	881
Basilicata	192.479	687
Calabria	657.348	588
Sicilia	1.746.877	742
Sardegna	574.050	754
<b>Totale</b>	<b>20.055.569</b>	<b>17.544</b>

#### **3 - Procedimento per il calcolo delle stime**

Le stime sono ottenute mediante uno stimatore di ponderazione vincolata. Il principio su cui è basato ogni metodo di stima campionaria è che le unità appartenenti al campione rappresentino anche le unità della popolazione che non sono incluse nel campione. Questo principio viene realizzato attribuendo ad ogni unità campionaria un peso che indica il numero di unità della popolazione rappresentate dall'unità medesima. Se, ad esempio, ad un'unità campionaria viene attribuito un peso pari a 100, vuol dire che questa unità rappresenta se stessa ed altre 99 unità della popolazione che non sono state incluse nel campione.

Per ottenere la stima del totale di una generica variabile di interesse Y occorre moltiplicare il peso finale associato a ciascuna unità campionaria per il valore della variabile y assunto da tale unità ed effettuare, a livello del dominio di interesse, la somma dei prodotti così ottenuti.

Il peso da attribuire alle unità campionarie è ottenuto per mezzo di una procedura complessa che ha le seguenti finalità:

- correggere l'effetto distorsivo dovuto agli errori di lista e al fenomeno della mancata risposta totale;
- tenere conto della conoscenza di alcuni totali noti sulla popolazione oggetto di studio, nel senso che le stime campionarie di tali totali devono coincidere con i rispettivi valori noti.

La procedura per la costruzione dei pesi finali da attribuire alle unità campionarie è articolata nelle seguenti fasi:

1. viene dapprima calcolato il peso base (o peso diretto), ottenuto come reciproco della probabilità di inclusione di ogni unità campionaria;
2. si calcola quindi il fattore correttivo che consente di soddisfare la condizione di uguaglianza tra i totali noti della popolazione e le corrispondenti stime campionarie;
3. il peso finale è dato dal prodotto del peso base per i fattori correttivi sopra indicati.

I fattori correttivi del punto 3 sono ottenuti dalla risoluzione di un problema di minimo vincolato, in cui la funzione da minimizzare è una funzione di distanza (opportunamente prescelta) tra i pesi base e i pesi finali e i vincoli sono definiti dalla condizione di uguaglianza tra stime campionarie dei totali noti di popolazione e valori noti degli stessi. La funzione di distanza prescelta è la funzione logaritmica troncata; l'adozione di tale funzione garantisce che i pesi finali siano positivi e contenuti in un predeterminato intervallo di valori possibili, eliminando in tal modo i pesi positivi estremi (troppo grandi o troppo piccoli).

Per il calcolo dei pesi la popolazione di riferimento è costituita dalle donne italiane in età 16-75 anni, al netto delle convivenze.

I totali noti<sup>2</sup> imposti a livello di ripartizione geografica sono i seguenti:

- a. popolazione per classi di età (16-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-75) a livello regionale;
- b. popolazione per tipologia comunale (aree A<sub>1</sub>, A<sub>2</sub>, B<sub>1</sub>, B<sub>2</sub>, B<sub>3</sub>, B<sub>4</sub> definite nel paragrafo 1);
- c. popolazione per titolo di studio (fino a licenza media o avviamento professionale, diploma superiore, laurea o titolo superiore);
- d. popolazione per stato civile (nubili, coniugate, separate o divorziate o vedove);
- e. popolazione per dimensione familiare (famiglie mono-componenti per età (16-49, 50-75), 2 componenti, 3, 4, 5 o più componenti).

Tutti i metodi di stima che scaturiscono dalla risoluzione di un problema di minimo vincolato del tipo sopra descritto rientrano in una classe generale di stimatori nota come stimatori di ponderazione

---

<sup>2</sup> I totali noti relativi al titolo di studio, allo stato civile, alla dimensione familiare e alle classi d'età derivano da stime dell'indagine sulle Forze di lavoro.

vincolata<sup>3</sup>. Un importante stimatore appartenente a tale classe, che si ottiene utilizzando la funzione di distanza euclidea, è lo stimatore di regressione generalizzata. Come verrà chiarito meglio nel paragrafo 4, tale stimatore riveste un ruolo centrale perché è possibile dimostrare che tutti gli estimatori di ponderazione vincolata convergono asintoticamente, all'aumentare della numerosità campionaria, allo stimatore di regressione generalizzata.

#### **4 - Valutazione del livello di precisione delle stime e presentazione sintetica degli errori campionari**

Le principali statistiche di interesse per valutare la variabilità campionaria delle stime prodotte dall'indagine sono l'errore di campionamento assoluto e l'errore di campionamento relativo.

Indicando con  $\hat{V}\text{ar}(_d\hat{Y})$  la varianza della stima  $_d\hat{Y}$ , riferita al dominio d, la stima dell'errore di campionamento assoluto di  $_d\hat{Y}$  si può ottenere mediante la seguente espressione:

$$\hat{\sigma}(_d\hat{Y}) = \sqrt{\hat{V}\text{ar}(_d\hat{Y})} \quad (1)$$

La stima dell'errore di campionamento relativo di  $_d\hat{Y}$ , è invece definita dall'espressione:

$$\hat{\varepsilon}(_d\hat{Y}) = \frac{\sqrt{\hat{V}\text{ar}(_d\hat{Y})}}{_d\hat{Y}} \quad (2)$$

Gli errori campionari delle espressioni (1) e (2), consentono di valutare il grado di precisione delle stime; inoltre, l'errore assoluto permette di costruire l'intervallo di confidenza, che, con una certa probabilità, contiene il parametro d'interesse. Con riferimento alla generica stima  $\hat{Y}$  tale intervallo assume la seguente forma:

$$\Pr\left\{\hat{Y} - k \hat{\varepsilon}(\hat{Y}) \leq Y \leq \hat{Y} + k \hat{\varepsilon}(\hat{Y})\right\} = P \quad (3)$$

Nella (3) il valore di k dipende dal valore fissato per la probabilità P; ad esempio, per P=0,95 si ha k=1,96.

Ad ogni stima  $_d\hat{Y}$  è associato un errore campionario relativo  $\hat{\varepsilon}(_d\hat{Y})$ ; quindi, per consentire un uso corretto delle stime fornite dall'indagine, sarebbe necessario fornire, per ogni stima pubblicata, anche il corrispondente errore di campionamento relativo.

Ciò, tuttavia, non è possibile, sia per limiti di tempo e di costi di elaborazione, sia perché le tavole della pubblicazione risulterebbero eccessivamente appesantite e di non agevole consultazione per l'utente finale. Inoltre, non sarebbero in ogni caso disponibili gli errori delle stime non pubblicate, che l'utente può ricavare in modo autonomo.

Per questi motivi, generalmente, si ricorre ad una presentazione sintetica degli errori relativi, basata sul *metodo dei modelli regressivi*. Tale metodo si fonda sulla determinazione di una funzione matematica che mette in relazione ciascuna stima con il proprio errore relativo.

---

3 Nella letteratura in lingua anglosassone sull'argomento tali estimatori sono noti come calibration estimators, si veda come riferimento: Deville J.C. e Sarndäl C.E. (1992), "Calibration Estimators in Survey Sampling", Journal of the American Statistical Association 87: 376-382.

L'approccio utilizzato per la costruzione di questi modelli è diverso a seconda che si tratti di variabili qualitative o quantitative. Infatti, solo nel caso delle stime di frequenze assolute (o relative) riferite alle modalità di variabili qualitative, è possibile utilizzare dei modelli che hanno un fondamento teorico e secondo cui gli errori relativi delle stime di frequenze assolute sono funzione decrescente dei valori delle stime stesse.

Per calcolare gli errori di campionamento è stato utilizzato un software generalizzato, messo a punto presso l'Istat, che consente di calcolare gli errori campionari e gli intervalli di confidenza e permette di costruire dei modelli regressivi per la presentazione sintetica degli errori di campionamento.

Il modello utilizzato per le stime di frequenze assolute, con riferimento al generico dominio d, è il seguente:

$$\log \hat{\varepsilon}^2(d\hat{Y}) = a + b \log(d\hat{Y}) \quad (4)$$

dove i parametri a e b vengono stimati mediante il metodo dei minimi quadrati.

Il prospetto 2 riporta i valori dei coefficienti a e b e dell'indice di determinazione  $R^2$  del modello utilizzato per l'interpolazione degli errori campionari delle stime di frequenze riferite alle famiglie e alle persone, per aree territoriali.

Sulla base delle informazioni contenute nel suddetto prospetto è possibile calcolare l'errore relativo di una determinata stima di frequenza assoluta  $d\hat{Y}^*$ , riferita ai diversi domini, mediante la formula:

$$\hat{\varepsilon}(d\hat{Y}^*) = \sqrt{\exp(a + b \log(d\hat{Y}^*))} \quad (5)$$

e costruire l'intervalllo di confidenza al 95% come:

$$\left\{ d\hat{Y}^* - 1,96 \cdot \hat{\varepsilon}(d\hat{Y}^*) \cdot d\hat{Y}^* ; d\hat{Y}^* + 1,96 \cdot \hat{\varepsilon}(d\hat{Y}^*) \cdot d\hat{Y}^* \right\}.$$

Allo scopo di facilitare il calcolo degli errori campionari, nel prospetto 3 sono riportati gli errori relativi percentuali corrispondenti a valori crescenti di stime di frequenze assolute calcolati introducendo nella (5) i valori di a e b riportati nel prospetto 2.

Le informazioni contenute in tale prospetto consentono di calcolare l'errore relativo di una generica stima di frequenza assoluta mediante due procedimenti di facile applicazione che, tuttavia, conducono a risultati meno precisi di quelli ottenibili applicando direttamente la formula (5).

Il primo metodo consiste nell'approssimare l'errore relativo della stima di interesse  $d\hat{Y}^*$  con quello, riportato nei prospetti, corrispondente al livello di stima che più si avvicina a  $d\hat{Y}^*$ .

Il secondo metodo, più preciso del primo, si basa sull'uso di una formula di interpolazione lineare per il calcolo degli errori di stime non comprese tra i valori forniti nei prospetti. In tal caso, l'errore campionario della stima  $d\hat{Y}^*$ , si ricava mediante l'espressione:

$$\hat{\varepsilon}_{(d\hat{Y}^*)} = \hat{\varepsilon}_{(d\hat{Y}^{k-1})} + \frac{\hat{\varepsilon}_{(d\hat{Y}^k)} - \hat{\varepsilon}_{(d\hat{Y}^{k-1})}}{d\hat{Y}^k - d\hat{Y}^{k-1}} (d\hat{Y}^* - d\hat{Y}^{k-1})$$

dove  $d\hat{Y}^{k-1}$  e  $d\hat{Y}^k$  sono i valori delle stime entro i quali è compresa la stima  $d\hat{Y}^*$ , mentre  $\hat{\varepsilon}_{(d\hat{Y}^{k-1})}$  e  $\hat{\varepsilon}_{(d\hat{Y}^k)}$  sono i corrispondenti errori relativi.

**Prospetto 2 - Valori dei coefficienti a, b e dell'indice di determinazione R<sup>2</sup> (%) delle funzioni utilizzate per le interpolazioni degli errori campionari delle stime di frequenze assolute per aree territoriali delle donne italiane**

	a	b	R <sup>2</sup> (%)
<b>ITALIA</b>	8,67	-1,12	92,7%
<b>RIPARTIZIONI GEOGRAFICHE (a)</b>			
Nord-ovest	7,83	-1,06	90,1%
Nord-est	7,55	-1,07	91,6%
Centro	8,50	-1,14	92,4%
Sud	8,25	-1,09	89,7%
Isole	8,86	-1,14	90,3%
<b>TIPI DI COMUNE (b)</b>			
A1	7,71	-1,01	94,7%
A2	7,20	-0,97	95,2%
B1	5,47	-0,84	90,2%
B2	6,44	-0,93	95,4%
B3	6,95	-0,96	94,5%
B4	6,96	-0,97	93,2%
<b>REGIONI</b>			
Piemonte	7,86	-1,07	95,0%
Valle d'Aosta	4,22	-1,03	94,6%
Lombardia	8,67	-1,11	95,3%
Trentino-Alto Adige	6,18	-1,05	95,2%
Bolzano	6,67	-1,09	93,3%
Trento	5,81	-1,04	94,4%
Veneto	7,92	-1,08	95,1%
Friuli-Venezia Giulia	6,54	-1,04	94,8%
Liguria	6,62	-1,03	94,8%
Emilia-Romagna	8,18	-1,10	95,4%
Toscana	7,71	-1,07	94,9%
Umbria	6,31	-1,05	94,5%
Marche	6,88	-1,05	92,8%
Lazio	8,73	-1,13	94,9%
Abruzzo	5,78	-0,92	90,2%
Molise	4,62	-0,92	89,7%
Campania	8,20	-1,04	91,8%
Puglia	7,10	-0,97	90,8%
Basilicata	5,22	-0,94	89,5%
Calabria	6,98	-1,00	91,7%
Sicilia	8,91	-1,13	92,3%
Sardegna	6,47	-0,96	91,9%

(a) Italia nord-occidentale: Piemonte, Valle d'Aosta, Lombardia, Liguria; Italia nord-orientale: Bolzano, Trento, Veneto, Friuli-Venezia Giulia, Emilia Romagna; Italia centrale: Toscana, Umbria, Marche, Lazio; Italia meridionale: Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria; Italia insulare: Sicilia, Sardegna.

(b) Comuni tipo A1: Area urbana centro; Tipo A2: Area urbana periferia; Tipo B1: comuni fino a 2 mila abitanti; Tipo B2: da 2.001 a 10 mila abitanti; Tipo B3: da 10.001 a 50 mila abitanti; Tipo B4: oltre 50 mila abitanti.

**Prospetto 3 - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute per aree territoriali, donne italiane**

STIME	Italia	Nord-ovest	Nord-est	Centro	Sud	Isole
20,000	29,6	26,3	21,7	24,7	28,6	29,1
30,000	23,6	21,2	17,5	19,6	23,0	23,1
40,000	20,1	18,2	15,0	16,7	19,6	19,6
50,000	17,7	16,2	13,3	14,7	17,4	17,3
60,000	16,0	14,7	12,0	13,2	15,8	15,6
70,000	14,7	13,5	11,1	12,1	14,5	14,2
80,000	13,6	12,6	10,3	11,2	13,5	13,2
90,000	12,7	11,8	9,7	10,5	12,6	12,3
100,000	12,0	11,2	9,2	9,9	11,9	11,6
200,000	8,1	7,7	6,3	6,7	8,2	7,8
300,000	6,5	6,2	5,1	5,3	6,6	6,2
400,000	5,5	5,4	4,4	4,5	5,6	5,3
500,000	4,9	4,8	3,9	3,9	5,0	4,6
750,000	3,9	3,8	3,1	3,1	4,0	3,7
1,000,000	3,3	3,3	2,7	2,7	3,4	3,1
2,000,000	2,2	2,3	1,8	1,8	2,3	2,1
3,000,000	1,8	1,8	1,5	1,4	1,9	1,7
4,000,000	1,5	1,6	1,3	1,2	1,6	-
5,000,000	1,3	1,4	-	-	1,4	-
7,500,000	1,1	-	-	-	-	-
10,000,000	0,9	-	-	-	-	-
15,000,000	0,7	-	-	-	-	-
20,000,000	0,6	-	-	-	-	-

**Prospetto 3 (segue) - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute per aree territoriali, donne italiane**

STIME	1	2	3	4	5	6
20,000	31,1	29,9	24,3	25,1	27,4	27,1
30,000	25,3	24,5	20,5	20,8	22,5	22,3
40,000	21,9	21,3	18,2	18,2	19,6	19,4
50,000	19,5	19,1	16,6	16,4	17,6	17,4
60,000	17,8	17,5	15,3	15,1	16,1	15,9
70,000	16,5	16,3	14,4	14,0	15,0	14,8
80,000	15,4	15,2	13,6	13,2	14,0	13,9
90,000	14,5	14,4	12,9	12,5	13,3	13,1
100,000	13,7	13,7	12,4	11,9	12,6	12,4
200,000	9,7	9,8	9,3	8,6	9,0	8,9
300,000	7,9	8,0	7,8	7,1	7,4	7,3
400,000	6,8	7,0	6,9	6,2	6,5	6,4
500,000	6,1	6,3	6,3	5,6	5,8	5,7
750,000	4,9	5,1	5,3	4,7	4,8	4,7
1,000,000	4,3	4,5	4,7	4,1	4,2	4,1
2,000,000	3,0	3,2	-	2,9	3,0	2,9
3,000,000	2,4	2,6	-	2,4	2,4	2,4
4,000,000	-	-	-	2,1	2,1	2,1
5,000,000	-	-	-	1,9	1,9	-

**Prospetto 3 (segue) - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute per aree territoriali, donne italiane**

STIME	Piemonte	Valle d'Aosta	Lombardia	Trentino-Alto Adige	Bolzano	Trento	Veneto	Friuli-Venezia Giulia
20,000	25,0	5,1	31,4	11,9	12,4	10,4	25,3	15,1
30,000	20,1	4,1	25,1	9,6	10,0	8,4	20,3	12,2
40,000	17,2	3,6	21,4	8,3	8,5	7,3	17,4	10,5
50,000	15,3	3,2	18,9	7,4	7,5	6,5	15,4	9,4
60,000	13,8	-	17,1	6,7	6,8	5,9	14,0	8,5
70,000	12,7	-	15,7	6,2	6,3	5,4	12,9	7,9
80,000	11,9	-	14,6	5,7	5,8	5,1	12,0	7,3
90,000	11,1	-	13,7	5,4	5,5	4,8	11,3	6,9
100,000	10,5	-	12,9	5,1	5,2	4,5	10,6	6,5
200,000	7,2	-	8,8	3,5	3,5	3,1	7,3	4,5
300,000	5,8	-	7,0	2,9	-	-	5,9	3,7
400,000	5,0	-	6,0	2,5	-	-	5,0	15,1
500,000	4,4	-	5,3	-	-	-	4,5	-
750,000	3,6	-	4,2	-	-	-	3,6	-
1,000,000	3,1	-	3,6	-	-	-	3,1	-
2,000,000	2,1	-	2,4	-	-	-	2,1	-
3,000,000	-	-	2,0	-	-	-	-	-

**Prospetto 3 (segue) - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute per aree territoriali, donne italiane**

STIME	Liguria	Emilia-Romagna	Toscana	Umbria	Marche	Lazio	Abruzzo	Molise
20,000	17,0	25,4	23,8	13,2	16,9	29,0	19,2	10,5
30,000	13,8	20,3	19,2	10,7	13,6	23,0	15,9	8,7
40,000	11,9	17,3	16,5	9,2	11,7	19,6	14,0	7,6
50,000	10,6	15,3	14,6	8,2	10,4	17,2	12,6	6,9
60,000	9,7	13,9	13,3	7,4	9,4	15,6	11,6	6,3
70,000	8,9	12,7	12,2	6,9	8,7	14,3	10,8	5,9
80,000	8,3	11,8	11,4	6,4	8,1	13,2	10,2	5,6
90,000	7,9	11,1	10,7	6,0	7,6	12,4	9,6	5,3
100,000	7,4	10,5	10,1	5,7	7,2	11,6	9,2	5,0
200,000	5,2	7,1	7,0	4,0	5,0	7,9	6,7	-
300,000	4,2	5,7	5,6	3,2	4,0	6,3	5,5	-
400,000	3,7	4,9	4,8	-	3,5	5,3	4,9	-
500,000	3,3	4,3	4,3	-	3,1	4,7	4,4	-
750,000	-	3,4	3,4	-	-	3,7	-	-
1,000,000	-	2,9	2,9	-	-	3,2	-	-
2,000,000	-	2,0	-	-	-	2,1	-	-

**Prospetto 3 (segue) - Valori interpolati degli errori relativi percentuali delle stime di frequenze assolute per aree territoriali, donne italiane**

STIME	Campania	Puglia	Basilicata	Calabria	Sicilia	Sardegna
20,000	35,1	28,5	13,2	23,6	32,1	22,0
30,000	28,4	23,4	10,9	19,3	25,5	18,1
40,000	24,5	20,4	9,5	16,7	21,7	15,8
50,000	21,8	18,3	8,6	14,9	19,1	14,2
60,000	19,8	16,7	7,9	13,6	17,3	13,0
70,000	18,3	15,5	7,3	12,6	15,8	12,1
80,000	17,1	14,6	6,9	11,8	14,7	11,3
90,000	16,1	13,8	6,5	11,1	13,7	10,7
100,000	15,2	13,1	6,2	10,6	12,9	10,2
200,000	10,6	9,3	4,5	7,5	8,8	7,3
300,000	8,6	7,7	-	6,1	7,0	6,0
400,000	7,4	6,7	-	5,3	5,9	5,2
500,000	6,6	6,0	-	4,7	5,2	4,7
750,000	5,3	4,9	-	-	4,2	-
1,000,000	4,6	4,3	-	-	3,5	-
2,000,000	3,2	3,1	-	-	2,4	-