

Progetto Stock & Flussi

Formato Standard dei Dati di Input della Procedura di Bilanciamento

1. Simboli ed Equazioni

$$P_2 = P_1 + N + \text{rowSums}(M)$$

$$N = B - D$$

$$M = F^t - F$$

- P_2 Popolazione al tempo T2
- P_1 Popolazione al tempo T1
- B Nati fra T1 e T2
- D Morti fra T1 e T2
- N Saldo naturale fra T1 e T2
- F Flussi migratori fra T1 e T2 (F_{ij} è il flusso dalla cella i alla cella j)
- M Saldi dei flussi migratori fra T1 e T2

2. Directory e File

- **Standard_Input:** Directory contenente i dati da bilanciare
- **Standard_Output:** Directory contenente i dati bilanciati

I dati da bilanciare contenuti nella directory **Standard_Input** sono memorizzati in file **.csv**. I file devono avere i seguenti nomi:

- **P2.csv**
- **P1.csv**
- **B.csv**
- **D.csv**
- **F.csv**

3. Schemi e Formati delle Tabelle P2.csv, P1.csv, B.csv, D.csv

Schema di P2.csv

Sex	Nation	AgeCl	Region	N
1	1	N	1	88824
...
1	1	1	1	1629936
1	1	1	2	1439411
1	1	2	1	7750119
1	1	2	2	3341710
...

Schema di P1.csv

Sex	Nation	AgeCl	Region	N
1	1	N	1	0
...
1	1	1	1	1639728
1	1	1	2	1466616
1	1	2	1	7773961
1	1	2	2	3358056
...

Schema di B.csv

Sex	Nation	AgeCl	Region	N
1	1	N	1	88982
1	1	N	2	39607
1	1	N	3	81664
1	2	N	1	24121
1	2	N	2	8131
...

Schema di D.csv

Sex	Nation	AgeCl	Region	N
1	1	N	1	184
1	1	N	2	96
1	1	N	3	233
1	1	1	1	177
1	1	1	2	85
1	1	1	3	204
...

In quanto segue per “tabella” si intende una qualsiasi delle tabelle contenute nei file **P2.csv**, **P1.csv**, **B.csv**, **D.csv**.

1) I nomi delle colonne della tabella sono fissi:

- **Sex**: sesso;
- **Nation**: cittadinanza (dicotomica “italiano” / “straniero”, o politomica e.g. “italiano” / “straniero UE” / “straniero non UE”);
- **AgeCl**: classe di coorte di nascita;
- **Region**: classificazione territoriale (e.g. ripartizione, o regione, o provincia, ...);

Se occorresse aggiungere altre variabili di classificazione, ne andrebbe concordato il nome;

- 2) La tabella deve contenere sempre almeno una variabile di classificazione (e.g. **Region**) e la colonna dei conteggi **N**;
- 3) Le modalità di tutte le variabili di classificazione devono essere numeri interi (e.g. **Sex** = **1** per i maschi e **Sex** = **2** per le femmine) o caratteri alfabetici speciali (vedi punto 6 sotto);
- 4) Per variabili di classificazione ordinabili (e.g. **AgeCl**), i codici interi associati alle modalità devono essere correttamente ordinati (e.g. **AgeCl** = **1** per age = [1, 5), **AgeCl** = **2** per age = [5, 10), ...) e consecutivi (e.g. se per age = [5, 10) è **AgeCl** = **2**, allora per age = [10, 15) deve essere **AgeCl** = **3**);
- 5) La colonna dei conteggi **N** deve avere solo valori numerici;
- 6) In **B.csv** e **D.csv**, la modalità ‘**N**’ della variabile di classificazione **AgeCl** deve rappresentare i **neonati** (i neonati sono gli individui nati fra T1 e T2, ovvero individui non ancora nati a T1 che a T2 o sono vivi – e hanno ancora coorte di nascita ‘N’ – o non sono sopravvissuti);
- 7) In **B.csv**, la variabile di classificazione **AgeCl** può assumere la sola modalità ‘**N**’;
- 8) In **D.csv**, la variabile di classificazione **AgeCl** deve fornire la coorte degli individui (deceduti fra T1 e T2);
- 9) In **P2.csv**, la variabile di classificazione **AgeCl** deve fornire la coorte degli individui; la coorte ‘N’ deve essere presente in P2;
- 10) In **P1.csv**, la variabile di classificazione **AgeCl** deve fornire la coorte degli individui; la coorte ‘N’ deve essere presente in P1 ma con N (conteggi) pari a 0;
- 11) Non è richiesto uno specifico ordinamento delle righe della tabella;
- 12) Non è richiesto uno specifico ordinamento delle colonne della tabella;
- 13) Per motivi di leggibilità e comparabilità, si consiglia: (a) di fissare un ordinamento delle colonne e di mantenerlo coerentemente in tutte le tabelle, (b) di scegliere come ultima colonna l’obbligatoria colonna dei conteggi **N**;

4. Schema e Formato della Tabella F.csv

Schema di F.csv

Sex1	Sex2	Nation1	Nation2	AgeCI1	AgeCI2	Region1	Region2	N
1	1	1	1	N	N	1	2	117
1	1	1	1	N	N	1	3	152
1	1	1	1	N	N	1	E	90
1	1	1	1	N	N	2	1	133
...
2	2	2	2	N	N	3	E	8
2	2	2	2	N	N	E	1	177
2	2	2	2	N	N	E	2	75
2	2	2	2	N	N	E	3	122
2	2	2	2	1	1	1	2	73
...
2	2	2	2	3	3	3	2	41
2	2	2	2	3	3	3	E	199
2	2	2	2	3	3	E	1	2668
2	2	2	2	3	3	E	2	992
2	2	2	2	3	3	E	3	505

- 1) I nomi delle colonne della tabella sono fissi;
- 2) Le modalità di tutte le variabili di classificazione devono essere numeri interi (e.g. **Sex** = **1** per i maschi e **Sex** = **2** per le femmine) o caratteri alfabetici speciali (vedi punto 3 sotto);
- 3) Per le variabili di classificazione territoriale **Region1** e **Region2**, la modalità '**E**' deve rappresentare l'**estero**;
- 4) Le variabili di classificazione **AgeCI1** ed **AgeCI2** devono fornire la coorte degli individui e sono sempre identiche;
- 5) Per le variabili di classificazione **AgeCI1** ed **AgeCI2**, la modalità '**N**' deve rappresentare i **neonati**;
- 6) Non è richiesto uno specifico ordinamento della tabella;
- 7) Non è richiesto uno specifico ordinamento delle colonne della tabella;
- 8) Per motivi di leggibilità e comparabilità, si consiglia: (a) di rendere adiacenti tutte le colonne relative alla stessa variabile ai tempi T1 e T2 (e.g. **Nation1** adiacente a **Nation2**), (b) mantenere l'ordinamento delle colonne delle altre tabelle (**P2.csv**, **P1.csv**, **B.csv**, **D.csv**), (c) di scegliere come ultima colonna l'obbligatoria colonna dei conteggi **N**;

9) La modalità **Sex1, Sex2, AgeCl1, AgeCl2** si riferiscono alle caratteristiche degli individui alla nascita, cioè non cambiano mai¹.

Le modalità delle variabili di classificazione **Nation1, Nation2, Region1, Region2** che, in una specifica riga, variano identificano la natura dell'**evento anagrafico** (e.g. trasferimento di residenza, acquisizione di cittadinanza, ...), mentre quelle che non variano si riferiscono alle caratteristiche degli individui coinvolti **al momento dell'evento anagrafico**.

- Es.1: se in una riga si ha: **Nation1=Nation2=1, Region1=1, Region2=2, Sex1=Sex2=1, AgeCl1=AgeCl2=2**, allora **N** si riferisce al numero di individui di sesso maschile nella classe di coorte "2" che si trasferiscono dalla regione 1 alla regione 2 e che in quel momento sono cittadini italiani.
- Es.2: se in una riga si ha: **Nation1=2, Nation2=1, Region1=Region2=3, Sex1=Sex2=2, AgeCl1=AgeCl2=4**, allora **N** si riferisce al numero di individui di sesso femminile nella classe di coorte "4" che diventano cittadini italiani e che in quel momento risiedono nella regione 3.

¹ Seppur raramente, gli individui potrebbero cambiare 'sesso' fra T1 e T2. Dati di input con conteggi non nulli di questi flussi sono dunque leciti. In ogni caso, la procedura di Bilanciamento non sarebbe autorizzata ad alterare i valori grezzi di questi flussi.

5. Variabili di Classificazione e Granularità della Procedura di Bilanciamento

La procedura di Bilanciamento conduce al rispetto delle Equazioni di Bilancio Demografico (DBE) per tutte le sottopopolazioni identificate incrociando le modalità delle variabili di classificazione specificate.

Se le classificazioni (e.g. territoriali) in vigore a T1 e T2 dovessero essere diverse, occorrerà “attualizzare” tutte le tabelle coinvolte (inclusa **P1.csv**) ri-calcolandole rispetto alla classificazione in vigore a T2.

Esempio 1

- Sex: 1 = Maschio, 2 = Femmina
- Nation: 1 = Italiano, 2 = Straniero
- AgeCl: N = 2017, 1 = 2003-2016, 2 = 1953-2002, 3 = 1900-1952
- Region: 1 = Nord, 2 = Centro, 3 = Mezzogiorno, E = Estero

✓ **Numero di DBE atomiche: $64 = 2 * 2 * 4 * 4$**

Esempio 2

- Sex: 1 = Maschio, 2 = Femmina
- Nation: 1 = Italiano, 2 = Straniero
- AgeCl: Classi di coorte di nascita quinquennali (19 modalità + ‘N’ Neonati)
- Region: Province (107 modalità + ‘E’ Estero)

✓ **Numero di DBE atomiche: $8640 = 2 * 2 * 20 * 108$**