

17 MAGGIO 2017

Definizione dei gruppi sociali e loro descrizione

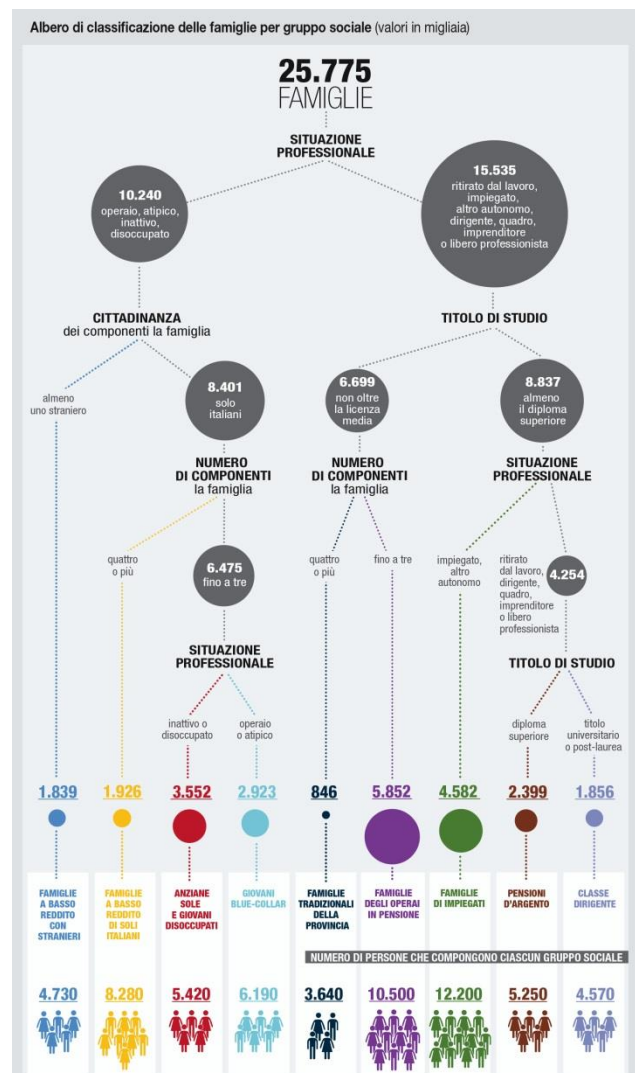
Presentata nel Rapporto annuale 2017, la suddivisione delle famiglie in gruppi è stata condotta sulla base di una metodologia statistica che ha consentito di non definire a priori l'esito della classificazione tramite alcune caratteristiche, quali il ruolo all'interno del processo produttivo o la posizione lavorativa occupata (elementi chiave solitamente utilizzati in modo esclusivo nella definizione dell'appartenenza di classe).

Con la metodologia prescelta, concorrono a definire l'appartenenza delle famiglie all'uno o all'altro dei diversi gruppi sociali una pluralità di caratteristiche. La considerazione di elementi di diversa natura e origine nella definizione dei gruppi consente una rappresentazione più articolata e completa della realtà sociale.

Per scelta, i gruppi sono frutto di un approccio multidimensionale, che è stato in grado di tenere in considerazione, con un differente ruolo nella formazione dei gruppi, aspetti di natura economica (reddito, condizione occupazionale), culturale (titolo di studio) e sociale (cittadinanza, dimensione della famiglia, tipologia del comune di residenza).

Caratterizzando i gruppi così costruiti per aspetti socio-economici e demografici è stato poi possibile tracciare profili e dinamiche di comportamento delle famiglie in ambito di consumi, partecipazione alla vita attiva del Paese, consumi culturali e così via. L'approccio adottato apporta un arricchimento delle dimensioni considerate nello studio delle disuguaglianze sociali, aggiungendo alla dimensione della condizione occupazionale altri aspetti di contesto e soprattutto di diretta ricaduta sull'inclusione ed esclusione sociale.

Le letture rese possibili dall'individuazione dei gruppi sociali a partire dalle loro caratteristiche, rilevate dalla statistica ufficiale, offrono un quadro variegato, aperto ad analisi e interpretazioni anche in termini di eterogeneità delle caratteristiche socio-economiche e strutturali. Esaminate in termini statici, queste eterogeneità consentono una lettura più articolata del presente; esaminate in termini dinamici, offrono indizi sulle direzioni dei cambiamenti in atto nella società italiana.



Per la costruzione dei gruppi sociali è stata adottata una tecnica non parametrica a segmentazione gerarchica, che ha permesso di tralasciare qualsiasi ipotesi sulla distribuzione a priori delle variabili di interesse. Questa classificazione è quindi il risultato di una sperimentazione realizzata appositamente per il Rapporto annuale 2017. Sono stati individuati nove gruppi, relativamente omogenei al loro interno per livello di reddito familiare equivalente. Ciascuno è connotato da una specifica combinazione di modalità delle altre variabili considerate. Le diverse caratteristiche che definiscono i gruppi sociali individuati consentono di interpretare contesti ed effetti dei differenziali di reddito e di fare luce sugli elementi che congiuntamente concorrono a interpretare le diseguaglianze nei redditi equivalenti delle famiglie.

La definizione dei gruppi sociali in funzione del reddito è stata condotta mediante l'ausilio di un metodo di regressione non parametrica che consente di produrre un 'albero di classificazione' e individua un certo numero di gruppi disgiunti, raccogliendo tutte le osservazioni accomunate dalle modalità di una o più delle variabili esplicative considerate. Si tratta di un approccio di classificazione multipla supervisionata nel quale a una variabile – il reddito equivalente – viene attribuito il ruolo di variabile rispetto alla quale costruire gruppi omogenei al loro interno (variabile dipendente); ad altre variabili viene dato il compito di segmentare le famiglie progressivamente in gruppi omogenei secondo il reddito (i nodi dell'albero), consentendo una lettura dei gruppi attraverso una combinazione distintiva di modalità di tali variabili (variabili esplicative o discriminanti); ad altre variabili ancora viene attribuito infine il ruolo di variabili illustrative, al fine di studiare ad ampio raggio l'eterogeneità interna ed esterna ai gruppi individuati. Il metodo non parametrico prescelto ha il pregio di consentire il trattamento congiunto di variabili quantitative e qualitative e di non dover fare le tradizionali ipotesi a priori sui processi di generazione dei dati utilizzati, tipiche dei metodi di regressione parametrici.

Grazie ai dati disponibili dall'indagine sul reddito e le condizioni di vita delle famiglie (Eu-Silc 2015) è stato possibile raggruppare le famiglie residenti in Italia utilizzando il reddito equivalente come variabile dipendente e le variabili esplicative di seguito elencate. La scelta di utilizzare variabili a livello familiare è dovuta alla condivisione delle risorse disponibili all'interno della famiglia. La scelta delle variabili selezionate è stata guidata da un duplice obiettivo: da una parte, la ricerca di quelle caratteristiche che possono essere considerate elementi caratterizzanti della dimensione economica della famiglia, oltre che determinanti di inclusione ed esclusione sociale, di similarità nei gusti di consumo culturale e di propensione alla partecipazione sociale; dall'altra, la possibilità di replicare la costruzione dei gruppi nelle altre indagini sulle famiglie:

- sesso della persona di riferimento (maschio, femmina);
- età della persona di riferimento (fino a 29 anni, 30-39 anni, 40-49 anni, 50-59 anni, 60-69 anni, 70-79 anni, 80 anni e oltre);
- titolo di studio della persona di riferimento (fino alle elementari, licenza di scuola secondaria di I grado, diploma di scuola secondaria di II grado, titolo universitario, titolo post-laurea);
- situazione professionale della persona di riferimento (dirigente o quadro, impiegato, operaio o assimilato, imprenditore o libero professionista, altro autonomo, atipico, disoccupato, inattivo, ritirato dal lavoro);
- numero di componenti della famiglia (1, 2, 3, 4, 5 o più);

- presenza di stranieri (famiglia di soli italiani, famiglia con stranieri);
- tipo di comune di residenza (centro di area metropolitana, periferia di area metropolitana, fino a 2.000 abitanti, da 2.001 a 10.000 abitanti, da 10.001 a 50.000 abitanti, oltre 50.000 abitanti).

I gruppi via via ottenuti attraverso la segmentazione gerarchica operata sono stati 'sfrondata' (pruning), in modo da minimizzare la complessità interpretativa (numero di gruppi finali) garantendo potere discriminatorio. Si è operata dunque una scelta che ha consentito di ottenere nove gruppi sociali.

Le variabili esplicative che offrono un contributo maggiore alla classificazione dei gruppi di famiglie sono la presenza di stranieri, la situazione professionale (che include sia aspetti riferiti al profilo occupazionale sia elementi relativi alla posizione nella professione e il titolo di studio della persona di riferimento, oltre al numero di componenti delle famiglie. Sesso ed età della persona di riferimento e tipologia del comune di residenza offrono un contributo più limitato alla definizione delle classi.

Sfruttando l'approccio gerarchico del metodo di classificazione prescelto, è utile, per una migliore comprensione dei gruppi sociali, descrivere la sequenza con le quali le diverse variabili esplicative intervengono nella partizione progressiva delle famiglie, creando i 'rami' e le 'foglie' dell'albero di classificazione. Il 'tronco' dell'albero (nodo 0), dove sono ancora presenti tutte le famiglie (quasi 26 milioni), fa riferimento al reddito, la variabile guida del modello adottato. La variabile situazione professionale della persona di riferimento della famiglia definisce i primi due 'rami' e quindi discrimina due macro-gruppi: il primo (nodo 1), che include le modalità lavoratore atipico (dipendente con contratto a termine o collaboratore), operaio o assimilato, inattivo e disoccupato, comprende poco più di 10 milioni di famiglie; il secondo (nodo 2), che comprende le modalità dirigente, quadro, imprenditore o libero professionista, impiegato, altro lavoratore autonomo e ritirato dal lavoro, raggruppa 15 milioni e mezzo di famiglie. Al passaggio successivo, il nodo 1, determinato da una situazione professionale cui corrisponde un minore benessere economico, si suddivide sulla base della variabile cittadinanza generando il nodo 3 (famiglie di soli italiani, circa 8 milioni e mezzo) e il nodo 4 (famiglie con stranieri, 1,8 milioni). Il nodo 2, quello della situazione professionale che garantisce maggiore benessere economico, si suddivide ulteriormente, invece, sulla base di una diversa variabile, il titolo di studio della persona di riferimento: nel nodo 5 si trovano le famiglie in cui la persona di riferimento possiede al più la licenza di scuola media (6,7 milioni di famiglie) e il nodo 6 in cui possiede un titolo di studio superiore alla licenza media (quasi 9 milioni di famiglie).

La variabile numero di componenti interviene due volte a discriminare le famiglie.

In primo luogo, scinde il nodo 3, costituito da famiglie di italiani con situazione professionale più disagiata, in due nodi ulteriori: il nodo 7 (poco meno di 6,5 milioni di famiglie) è caratterizzato da famiglie con un numero di componenti minore o uguale a tre; il nodo 8 (circa 2 milioni) da famiglie di quattro componenti o più.

In secondo luogo, anche il nodo 5, (famiglie la cui persona di riferimento ha una buona situazione professionale e un basso titolo di studio) si suddivide nel nodo 11 (quasi 6 milioni di famiglie con un numero di componenti minore di quattro) e nel nodo 12 (meno di 1 milione di famiglie con quattro o più componenti).

Invece, il nodo 6 (famiglie in cui la persona di riferimento ha una buona situazione professionale e un titolo di studio superiore alla licenza media) si divide in due ulteriori nodi sempre sulla base della situazione professionale: nel nodo 13 (oltre 4 milioni di famiglie) la persona di riferimento

è ritirato dal lavoro, dirigente, quadro, imprenditore o libero professionista che, se almeno laureato, va a costituire il nodo 28 (1,8 milioni di famiglie); e altrimenti ricade nel nodo 27 (2,4 milioni di famiglie). Diversamente che nel nodo 13, nel nodo 14 la persona di riferimento è impiegata o svolge un altro lavoro autonomo (oltre 4,5 milioni di famiglie).

La situazione professionale è ancora determinante per la partizione del nodo 7 (lo era già stata al primo livello di questo ramo dell'albero) suddividendolo ulteriormente nel nodo 15 (quasi 3 milioni di famiglie in cui la persona di riferimento è un lavoratore atipico, un operaio o un assimilato) e nel nodo 16 (circa 3,5 milioni di famiglie in cui la persona di riferimento è inattiva o disoccupata).

In definitiva, i gruppi individuati sono ognuno definiti da una sequenza di variabili esplicative discriminanti:

Giovani blue-collar (nodo 15): situazione professionale (atipico, operaio, inattivo, disoccupato) → cittadinanza (famiglie di soli italiani) → numero di componenti (fino a tre) → situazione professionale (operaio o atipico)

Anziane sole e giovani disoccupati (nodo 16): situazione professionale (atipico, operaio, inattivo, disoccupato) → cittadinanza (famiglie di soli italiani) → numero di componenti (fino a tre) → situazione professionale (inattivo o disoccupato)

Famiglie a basso reddito di soli italiani (nodo 8): situazione professionale (atipico, operaio, inattivo, disoccupato) → cittadinanza (famiglie di soli italiani) → numero di componenti (quattro o più)

Famiglie a basso reddito con stranieri (nodo 4): situazione professionale (atipico, operaio, inattivo, disoccupato) → cittadinanza (almeno uno straniero)

Famiglie di operai in pensione (nodo 11): situazione professionale (ritirato dal lavoro, impiegato, altro autonomo, dirigente, quadro, imprenditore o libero professionista) → titolo di studio (non oltre la licenza media) → numero di componenti (fino a tre)

Famiglie tradizionali della provincia (nodo 12): situazione professionale (ritirato dal lavoro, impiegato, altro autonomo, dirigente, quadro, imprenditore o libero professionista) → titolo di studio (non oltre la licenza media) → numero di componenti (quattro o più)

Pensionati d'argento (nodo 27): situazione professionale (ritirato dal lavoro, impiegato, altro autonomo, dirigente, quadro, imprenditore o libero professionista) → titolo di studio (almeno il diploma di media superiore) → situazione professionale (ritirato dal lavoro, dirigente, quadro, imprenditore o libero professionista) → titolo di studio (diploma di media superiore)

Classe dirigente (nodo 28): situazione professionale (ritirato dal lavoro, impiegato, altro autonomo, dirigente, quadro, imprenditore o libero professionista) → titolo di studio (almeno il diploma di media superiore) → situazione professionale (ritirato dal lavoro, dirigente, quadro, imprenditore o libero professionista) → titolo di studio (titolo universitario o post-laurea)

Famiglie di impiegati (nodo 14): situazione professionale (ritirato dal lavoro, impiegato, altro autonomo, dirigente, quadro, imprenditore o libero professionista) → titolo di studio (almeno il diploma di media superiore) → situazione professionale (impiegato, altro autonomo)

La procedura di classificazione ad albero, automatizzata nell'algoritmo applicato e guidata da scelte ragionate tendenti a facilitare la sintesi e la descrizione dei gruppi sociali nel processo di pruning, ha generato nove gruppi sociali, di composizione e consistenza differenti. In generale, con riferimento alla distanza dal reddito medio, due dei nove gruppi possono definirsi a reddito medio, quattro a basso reddito e tre a reddito elevato.

Avendo classificato le famiglie in gruppi tramite l'associazione tra variabili esplicative e reddito equivalente, è utile non solo fornire una descrizione dei gruppi secondo le variabili di classificazione, ma anche una rappresentazione sintetica della loro condizione economica.

Per valutare le condizioni economiche è necessario integrare il quadro di lettura con alcuni indicatori che forniscono informazioni sul disagio economico e sul rischio di esclusione sociale, quali il rischio di povertà, la severa deprivazione materiale e la bassa intensità lavorativa, sintetizzati in un quarto indicatore: il rischio di povertà o esclusione sociale. Questi indicatori, armonizzati a livello europeo, si integrano l'uno con l'altro, poiché forniscono una visione di deprivazione monetaria rispetto al resto della popolazione (il rischio di povertà); una visione di deprivazione non monetaria in termini assoluti (la severa deprivazione che misura il non potersi permettere per motivi economici determinati beni e servizi); e una visione di esclusione sociale (la bassa intensità lavorativa), in un'ottica in cui il lavoro non è solo una forma di approvvigionamento di risorse economiche, ma anche e soprattutto un mezzo di partecipazione sociale e di crescita personale.

Le famiglie a reddito medio sono i **giovani blue-collar** e le **famiglie di operai in pensione**.

Le famiglie a basso reddito comprendono le **famiglie a basso reddito con stranieri**, le **famiglie a basso reddito di soli italiani**, le **famiglie tradizionali della provincia** e le **anziane sole e giovani disoccupati**.

Le famiglie benestanti sono formate da tre gruppi: le **famiglie di impiegati**, i **pensionati d'argento** e la **classe dirigente**.

Per la descrizione dei gruppi si rinvia al [Rapporto annuale 2017](#), in particolare alle pagine 61-70.