

# DEPRIVAZIONE IN ARGENTINA

Analisi multidimensionale della deprivazione in Argentina utilizzando i  
dati del Indagine Permanente di Famiglie (EPH)

UNIVERSITÀ DEGLI STUDI DI GENOVA - MAGGIO 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

“Ogni scoperta contiene un  
elemento irrazionale, o  
un'intuizione creativa”

Karl Popper

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Dove viene misurato?

- Uno strumento abbastanza semplice per misurare i differenziali socioeconomici. In generale, si basano su dati di sondaggi o censimenti facilmente disponibili.
- Inizialmente si pensava di utilizzare i dati del censimento del 2010 e l'Indagine annuale sulle famiglie della città di Buenos Aires.
- Per questo lavoro si ha usato le variabili selezionate per un precedente studio su Genova, dal Prof. Ivaldi, selezionati tra i diversi modelli teorici.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Dove viene misurato?

- Basato sui dati dei 32 cluster esaminati dell'Indagine Permanente di Famiglie (EPH) corrispondente al terzo trimestre del 2017.
- L'EPH copre 32 agglomerati urbani, dove vive il 70% della popolazione di Argentina.
- Copre tutti i capoluoghi di provincia e le città di oltre 100.000 abitanti.
- Non coprono l'intera popolazione argentina, solo la popolazione urbana, non copre la popolazione rurale.
- La costruzione dell'insieme di variabili iniziali ha comportato la trasformazione dei dati grezzi dell'EPH.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Come viene misurato?

- Il modello include un complesso di tecniche per identificare le dimensioni sottostanti di un insieme di variabili.
- Procedure dall'algebra della matrice a una matrice di correlazioni tra variabili al fine di sintetizzare la varianza dell'insieme di variabili in un numero minore di vettori, chiamati "componenti".
- L'obiettivo: misurare gli effetti della deprivazione materiale e sociale:
  - Determinare una suddivisione delle variabili basata sull'ACP, al fine di trovare le combinazioni lineari tra le variabili studiate che meglio riproducono la varianza all'interno del proprio set.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Come viene misurato?

- Le associazioni statistiche tra coppie di variabili sono interpretate come prova dell'esistenza (o non) di una significativa relazione semantica tra (alcune di) queste variabili e, infine, come una giustificazione (o mancanza di giustificazione) che ci permette di considerare (alcune delle loro) variabili come indicatori dello stesso concetto generale.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Come viene misurato?

- Il principale criterio di selezione: la possibilità di identificare i gruppi di popolazione che potrebbero avere una maggiore probabilità di mostrare una situazione socioeconomica svantaggiata.
- Si ha fatto una prima selezione delle variabili a cui è stato applicato l'ACP.
- Le dimensioni che sono interessanti per l'analisi possono essere molteplici e ciascuna può essere rappresentata da molte variabili, che allo stesso tempo possono essere definite empiricamente e quantificate, oltre a essere interpretate semanticamente.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Costruzione e lettura di diagrammi

- Per interpretare i componenti, è necessario osservare la rappresentazione grafica delle variabili su piani cartesiani definiti da coppie di componenti.
- L'ispezione della matrice di correlazione consente di valutare l'opportunità di applicare il ACP, nonché di rilevare la presenza di variabili collineari ed eliminarne alcune.
- Se la maggior parte delle correlazioni fosse molto bassa, non sarebbe opportuno eseguire un ACP. Viceversa, correlazioni molto alte possono indicare la presenza di variabili collineari.
- Sull'asse delle ascisse si trova il primo componente estratto e sull'asse delle ordinate quello che è stato estratto al secondo posto.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Procedura

- a) Calcolare una matrice di correlazioni con le variabili precedentemente selezionate;
- b) Ispezionare la matrice per identificare le correlazioni troppo alte (superiori a 0,9 in valore assoluto) ed eliminare dal cestino una delle due variabili coinvolte al fine di evitare le conseguenze della colinearità;
- c) Ripetere l'ACP; generare diagrammi del primo componente rispetto al secondo, poi vs il terzo, e così via;
- d) Dall'ispezione di questi diagrammi, identificare i cluster di variabili che occupano posizioni adiacenti nello spazio rispetto alle dimensioni  $v-1$ , e dare loro un'etichetta semantica alla luce delle variabili che le saturano;

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Procedura

- e) Dalla loro posizione nei diagrammi, assegnare ciascuna variabile a uno di questi componenti, o a più di uno o a nessuno;
- f) Realizzare un ACP separato per ciascuna delle dimensioni identificate nella fase d); ogni analisi opera su uno dei cluster individuati nella fase e); in ogni analisi, due componenti sono necessariamente estratti per ispezionare la posizione delle variabili nel piano così formato;
- g) L'analisi di questi diagrammi elimina le variabili che possono essere considerate marginali rispetto a quelle che costituiscono il nucleo della dimensione e i cui contributi semantici non sono considerati importanti

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Procedura

- h) In ciascuno degli ACP, le operazioni (f) e (g) sono iterate fino al raggiungimento di un risultato soddisfacente;
- i) Quando ha raggiunto un risultato soddisfacente per una dimensione, l'ACP è ripetuto, ma questa volta estraendo solo un componente, in modo di condensare in un singolo vettore la massima varianza estraibili, e per calcolare i coefficienti relativi e componenziali;
- j) Per ogni componente (che rappresenta una dimensione), se una o più variabili risultano avere coefficienti componenziali molto bassi o di segno opposto ai pesi relativi vengono eliminate dal canestro ripetendo l'ACP;

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Procedura

- k) Per ciascun componente, i punti «i» e «j» sono ripetuti fino a quando la dimensione è rappresentata da un piccolo numero di variabili con coefficienti di componente bilanciati;
- l) Per ogni dimensione viene prodotto un indice, utilizzando i coefficienti prodotti dall'ultimo ACP per pesare i punteggi standardizzati di tutti i casi (individui, aggregati territoriali, ecc.) sulle variabili che sono state mantenute nell'ultimo ACP della serie;
- n) Questi indici correlano con l'altro (per determinare il grado di associazione tra i componenti che rappresentano le diverse dimensioni) e altre variabili della stessa matrice come in ogni analisi.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Procedura

- Per la realizzazione della versione alternativa dell'ACP è necessario scollegare il piano metodologico dalle esigenze della procedura di algebra della matrice dell'ACP tradizionale.

*«Va notato che quando si vuole investigare il numero e la natura delle dimensioni sottostanti di un insieme di variabili, la rappresentazione delle distanze reciproche di queste variabili in uno spazio bidimensionale è solo una fase intermedia, che serve per identificare le dimensioni e assegnare a ciascuna di esse le variabili semanticamente rilevanti tra quelle del paniere iniziale. È anche necessario identificare e interpretare i sottoinsiemi semanticamente omogenei di variabili (cluster) all'interno del paniere iniziale»*

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Procedura

- Le restrizioni che l'ortogonalità introduce a livello semantico condizionano le scelte del ricercatore.

*«L'unica cosa che conta davvero ... è la rispettiva posizione dei punti che rappresentano le variabili nello spazio multidimensionale. ... Una volta ottenuto questo risultato, il sistema di assi ha svolto la sua funzione: lontano dal reificarlo o dall'introdurlo, può essere dimenticato. Come nella fotografia di una stanza, le diverse posizioni in cui può essere posizionato l'obiettivo sono arbitrarie e prive di significato: vengono utilizzate solo per determinare le posizioni relative degli oggetti nella stanza»*

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina

### Procedura

- Questa procedura, lontano dal costituire un abbandono dell'oggettività dell'analisi, al contrario, organizza una rappresentazione più completa, dove non vengono presi solo gli aspetti matematico-spaziali, ma anche le sfumature semanticamente pertinenti.

«...il vantaggio più importante dal punto di vista metodologico della strategia a due stati è restituire al ricercatore un pieno controllo concettuale di ogni singolo passaggio della procedura. Il ricercatore deve usare questa, come ogni altra tecnica di analisi dei dati, al fine di ottenere risultati che soddisfano i suoi obiettivi cognitivi»

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina

### Analisi

- Applichiamo un'ACP multipla eseguita con i criteri abituali al paniere di variabili.
- Fase di carattere prettamente esplorativo: per individuare le dimensioni salienti.
- Ispezionando i diagrammi in cui una componente viene posta in ascissa e l'altra in ordinata, si decide quante dimensioni vale la pena di interpretare.
- Al termine del primo stadio, l'originario paniere delle variabili viene diviso in sotto-insiemi, uno per ciascuna dimensione individuata.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO



## Deprivazione in Argentina Analisi

- Secondo stadio dell'ACP: per ciascun sotto-insieme individuato nel primo stadio si effettua un'ACP e si ripete l'analisi fino a che la dimensione non sia affinata in modo soddisfacente.
- L'attenzione è rivolta alla prima componente, sulla quale tutte le variabili del sotto-insieme presentano pesi alti.
- Anche se ci interessa solo la prima componente, conviene chiedere l'estrazione di due componenti per avere la possibilità di esaminare le proiezioni delle variabili del sotto-insieme in diagrammi bidimensionali.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

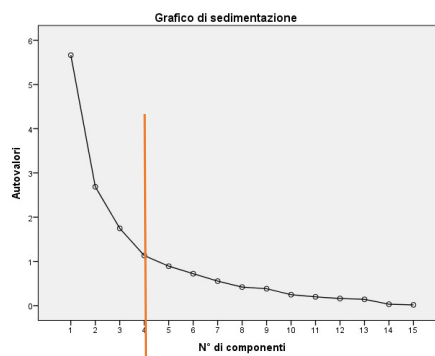
- Primo obiettivo: ridurre il numero di variabili al fine di facilitare il controllo intellettuale e la comprensibilità delle soluzioni che verranno proposte.
- 4 componenti con autovalori maggiore di 1 sono stati.
- Vengono analizzati sia la tabella dei risultati che il grafico di sedimentazione. La capacità esplicativa delle prime tre componenti riproduce poco più dei due terzi della varianza dell'intera matrice di correlazione e il quarto componente spiega solo il 7% della varianza.
- Si è deciso di continuare l'analisi dei primi 3 componenti.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

Componenti	Autovalori iniziale		
	Total	% di variazioni	% total accumulata
1	5,428	38,775	38,775
2	2,458	17,558	56,333
3	1,746	12,475	68,808
4	1,048	7,489	76,297



6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Si traccia lo spazio fattoriale, in cui le prime due componenti estratte (da cui risulterà la prima dimensione) vengono proiettate su un piano cartesiano, che spiega le percentuali di varianza più elevate (56,33%) delle variabili nello spazio.
- Quando si traccia la traversata della prima con la terza componente (dove è prevista la seconda dimensione) o la seconda con la terza (con l'aspettativa che emerga una terza dimensione), avremo figure diverse, come proiezioni su piani diversi di uno spazio di 13 (cioè,  $v-1$ ) dimensioni.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- La dispersione delle 14 variabili rende l'interpretazione difficile.
- Prossimo passo: selezionare le variabili in base a 2 criteri:
  - una semantica (derivata dal suo riferimento empirico)
  - una spaziale (dalla posizione nel grafico e basata sul peso che avevano nelle due dimensioni che costituiscono gli assi della rappresentazione).
- In questa fase le decisioni sono prese fundamentalmente dal criterio semantico, prendendo in considerazione possibili sovrapposizioni.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

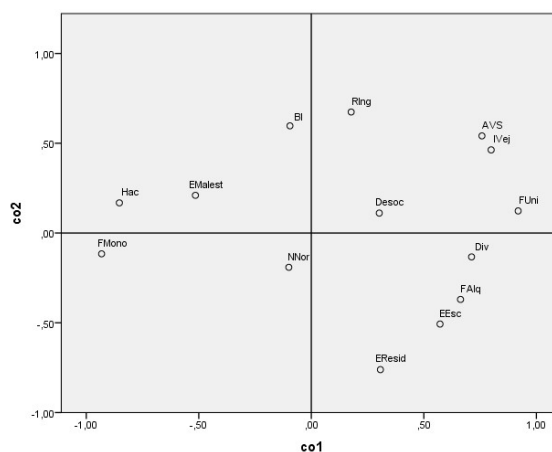
## Deprivazione in Argentina Analisi

- Questa vicinanza semantica tra le variabili viene giudicata dagli interessi cognitivi, analizzando ciascuna variabile e le sue relazioni.
- Il processo continua con lo scarto di variabili non correlate per l'analisi.
- Come sottolinea Marradi, in questa fase iniziale è opportuno essere "generosi" nella selezione delle variabili.
- Significar includere variabili che potrebbero essere eliminate in fasi successive, perché sono scarsamente collegate al componente o ridondanti.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi



6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

	Matrice di componenti			
	Componente			
	1	2	3	4
Bassa Istruzione	-0,095	0,597	-0,244	0,501
Reciproco del reddito * 1000	0,177	0,674	-0,085	-0,497
Nati nel Nordest	-0,1	-0,191	-0,803	0,356
Sovraffollamento	-0,853	0,168	0,162	0,097
Indice di vecchiaia	0,799	0,463	0,159	-0,083
Anziani vivono da soli	0,758	0,541	0,141	-0,04
Stranieri residenti	0,307	-0,761	0,086	-0,081
Stranieri nelle scuole	0,572	-0,507	0,282	-0,142
Famiglie in affitto	0,663	-0,37	0,2	0,184
Divorziati	0,712	-0,133	-0,042	0,305
Disoccupati	0,302	0,111	0,633	0,397
Edifici in stato mediocre	-0,515	0,21	0,599	0,286
Famiglie monoparentali	-0,932	-0,116	0,176	-0,077
Famiglie unipersonali	0,919	0,123	-0,209	0,08

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Le variabili BI (istruzione bassa), anello (reciproco di reddito), NNor (Nati nel Nordest) e Desoc (disoccupati) hanno un peso ridotto nella componente 1.
- Le variabili FUn (famiglie monoparentali), AVS (Anziani che vivono da soli) e Ivej (Indice di vecchiaia), sebbene siano spazialmente separati da uno dei cluster trovati - quello composto da FAlq (Famiglie che affittano), EEsc (Stranieri nelle scuole) e Div (Divorziati) -, hanno un peso considerevole nella Componente 1, quindi possono essere assegnati all'alone del cluster, quindi vengono mantenuti.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

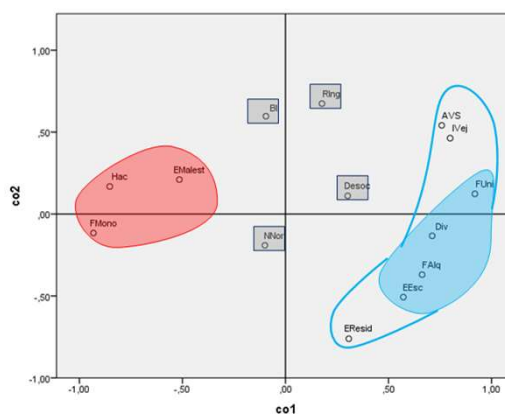
## Deprivazione in Argentina Analisi

- Sebbene la variabile Ersid (Residenti Stranieri) mostri un peso basso nella Componente 1, può essere considerata semanticamente vicina alla variabile EEsc (Stranieri nelle scuole), quindi è anche considerato appartenente al suddetto cluster.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi



6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

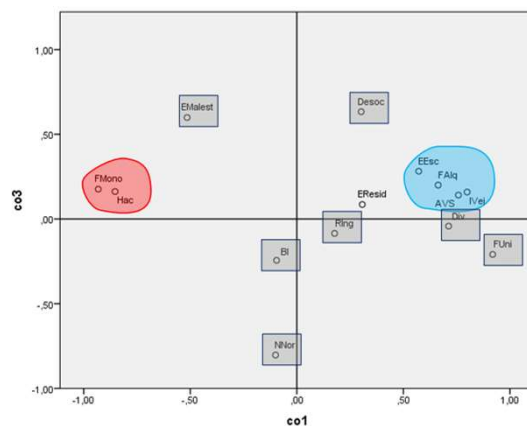
- Come vedremo, da una prospettiva spaziale, le variabili da scartare (ombreggiate in grigio) sono Emalest (Edificio in stato mediocre), Desoc (disoccupati), Ring (Reciproco del reddito), NNor (Nati nel Nordest), BI (Basso reddito), Div (Divorziati) e FUn (famiglie monoparentali).
- Sebbene si possano osservare 2 gruppi, uno nel primo quadrante e l'altro nel secondo quadrante, è difficile trovare una dimensione che li unisca e dia loro un significato semantico.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

■ Traversata dei componenti 1 e 3

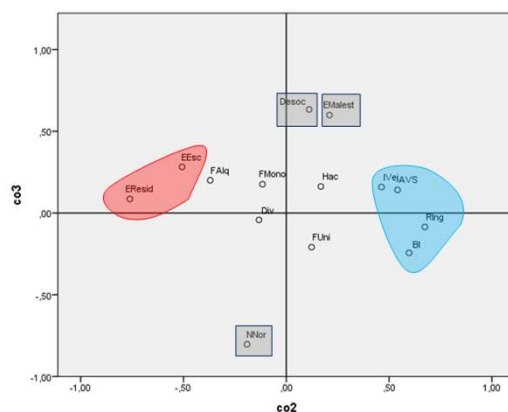


6 maggio 2019

**UNIREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

■ Traversata dei componenti 2 e 3



6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Dopo avere analizzato a fondo le diverse coppie di componenti emerse dal primo ACP, e considerando che i risultati ottenuti incrociando il Componente 1 con il Componente 2 sono i più appropriati per continuare, abbiamo proceduto a scartare le suddette variabili: BI (Bassa Istruzione), RIng (reciproco di reddito), NNor (Nati nel Nordest), Desoc (disoccupati).
- Questo insieme di variabili viene nuovamente analizzato e il risultato è rappresentato graficamente.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Varianza totale spiegata

Componente	Autovalori iniziali		
	Totale	% varianza	% cumulativo
1	5,315	53,153	53,153
2	1,818	18,179	<b>71,332</b>

- Le due componenti estratte spiegano il 71,3% della varianza.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO



## Deprivazione in Argentina Analisi

- I pesi delle variabili finali.
- Ad eccezione di EEst (gli stranieri frequentano la scuola), che ha un peso di .319, tutte le variabili pesano più di 0,57 o -5,5 (lato negativo).

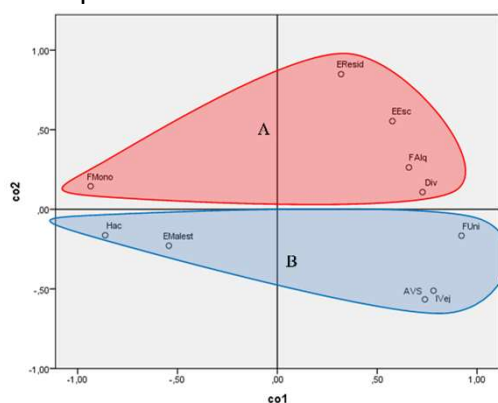
Matrice di componenti	
Famiglie in affitto	0,922
Indice di vecchiaia	0,782
Anziani vivono da soli	0,739
Divorziati	0,727
Famiglie in affitto	0,659
Stranieri nelle scuole	0,576
Stranieri residenti	0,319
Edifici in stato mediocre	-0,543
Sovraffollamento	-0,862
Famiglie monoparentali	-0,935

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- In ciascun cluster viene osservata la coerenza interna ed è facilmente interpretabile.



6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Definita questa dimensione, è il momento di passare da un vettore di pesi componenziali a un vettore di coefficienti componenziali per calcolare il contributo netto di una variabile a un componente.
- I coefficienti positivi variano tra 0,256 e 0,277; i negativi variano tra -0,164 e -0,221.
- Significa che ogni variabile scelta per formare l'indice dà un contributo che non si distingue dagli altri e non è nemmeno evidenziato da essi.
- L'indice è come un lago in cui dieci affluenti forniscono praticamente la stessa quantità d'acqua.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Coefficienti componenziali per il calcolo dei punteggi nella dimensione PRIVMAT-SOCIOECO

Matrice del coefficiente del punteggio del componente	
Famiglie unipersonali	,227
Indice di vecchiaia	,265
Anziani vivono da soli	,256
Edifici in stato mediocre	-,164
Sovraffollamento	-,221

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- I coefficienti elencati della tabella servono a costruire la formula in base alla quale il punteggio di ogni conglomerato viene calcolato nella dimensione PRIVMAT-SOCIOECO.
- $\text{PRIVMAT-SOCIOECO} = (\text{Sov} * -0,221) + (\text{EstaoM} * -0,164) + (\text{AVS} * 0,256) + (\text{IVech} * 0,265) + (\text{FUni} * 0,277)$

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Dall'interpretazione dei fattori principali e dalla raffinatezza delle dimensioni (in questo caso, solo una), è possibile attribuire a ciascun conglomerato un punteggio (component score) in ciascuna dimensione.
- Se i pesi dei componenti fossero usati per questo scopo, gli elementi di ridondanza sarebbero incorporati nel punteggio del componente.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- I *factor loadings* sono equivalenti ai coefficienti di correlazione tra una variabile e un fattore dedotto e non tiene conto delle correlazioni di tale variabile con altre variabili correlate anche con lo stesso fattore, il che implicherebbe che più a volte lo stesso materiale semantico.
- Sarebbe come usare i coefficienti di correlazione invece dei pesi beta in una regressione multipla.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Per questo, per ogni dimensione, i coefficienti del punteggio dei componenti che sono mostrati nella Tabella sono stati stimati con l'SPSS.
- Con questi coefficienti, il contributo netto (come non condiviso con altre variabili presenti nel paniere) di ciascuna variabile nel punteggio del cluster sulla dimensione è ponderato.
- Questi sono ottenuti con una regressione multipla in cui il fattore viene preso come variabile dipendente e le variabili del paniere come variabili indipendenti.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Analisi

- Il coefficiente componenziale di una variabile è sempre inferiore al suo peso perché elimina la porzione della sua varianza che ha in comune con le altre variabili.
- I coefficienti componenziali sono quindi analoghi, dal punto di vista tecnico, ai coefficienti parziali standardizzati (beta weights) di una regressione lineare multipla e rappresentano il contributo netto delle variabili alla costruzione del componente, tenendo sotto controllo la sua relazione con le altre variabili.

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

## Deprivazione in Argentina Conclusioni

Conglomerato	Indice
Ciudad de Bs As	-0,072
Gran Paraná	-0,246
Gran Rosario	-0,276
Bahía Blanca - Cerri	-0,271
.....	.....
Gran Catamarca	-0,603
Salta	-0,620
S. del Estero - La Banda	-0,699

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO

# Grazie!!

6 maggio 2019

**UNTREF** UNIVERSIDAD NACIONAL  
DE TRES DE FEBRERO