

# Le innovazioni metodologiche nella rilevazione dei prezzi al consumo

Federico Polidoro  
(polidoro@istat.it)

Palermo, 17 settembre 2015

# Indice della presentazione

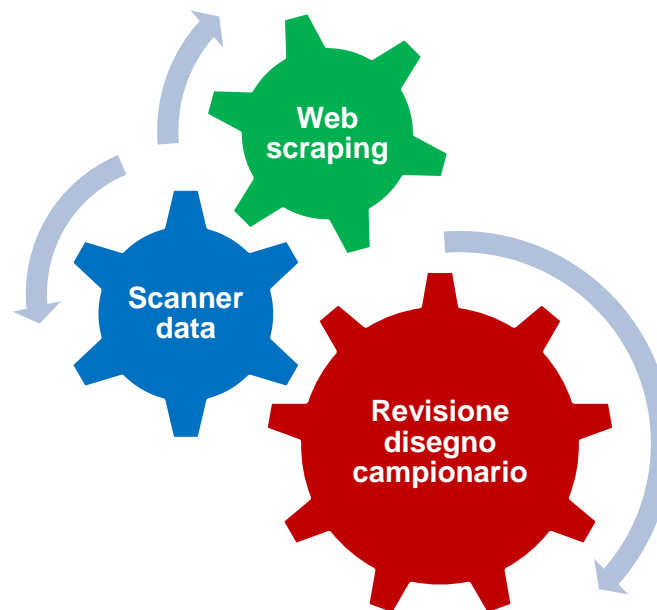
- Le linee guida del nuovo progetto di innovazione dell'indagine sui prezzi al consumo
- Lo stato di avanzamento delle attività sugli «scanner data»
- Lo stato di avanzamento delle attività sulle tecniche di «web scraping»
- Il progetto di innovazione e le prospettive di riorganizzazione della rete comunale e dell'informazione territoriale
- Gli obiettivi strategici

# 1. Le linee guida del nuovo progetto di innovazione dell'indagine sui prezzi al consumo

- Il contesto europeo e il progetto denominato “Multipurpose Price Statistics” (MPS)
- MPS finalizzato a una modernizzazione complessiva dell'impianto delle indagini sulla base delle quali sono prodotte le statistiche sui prezzi al consumo sia degli Stati membri sia europee
- Uno degli aspetti peculiari del progetto MPS è quello della modernizzazione delle tecniche di acquisizione dati con particolare riferimento agli scanner data e all'utilizzo del web scraping, ossia di procedure automatiche di scarico dei dati dal Web

# 1. Le linee guida del nuovo progetto di innovazione dell'indagine sui prezzi al consumo

- Il tema della modernizzazione delle tecniche di acquisizione dati, è stato legato al progetto di revisione della strategia campionaria dell'indagine sui prezzi al consumo con l'intento di introdurre un approccio probabilistico che, almeno in parte, possa sostituire quello ragionato attualmente utilizzato



# 1. Le linee guida del nuovo progetto di innovazione dell'indagine sui prezzi al consumo

- Il disegno campionario con riferimento ai prodotti grocery sarà a tre stadi:

1. PRIMO STADIO: Unità primarie i comuni

2. SECONDO STADIO: Unità campionarie i punti vendita, selezionati con probabilità proporzionali al fatturato, tenendo conto della presenza/assenza di un determinato gruppo di prodotti con schema di selezione coordinata campioni distinti, uno per ciascuna classe di prodotto , in modo da ottenere un alto livello di sovrapposizione dei campioni selezionati per ciascuna classe di prodotto, riducendo la dimensione del campione complessivo di punti vendita

TERZO STADIO: Unità finali le referenze. Si stanno testando diversi schemi di campionamento

# 1. Le linee guida del nuovo progetto di innovazione dell'indagine sui prezzi al consumo

- Gli scanner data e il nuovo disegno di campionamento:
  1. PRIMO STADIO: i comuni e le possibilità di ampliamento delle coperture territoriali del campione con gli scanner data
  2. SECONDO STADIO: i punti vendita ASIA – PV ormai regolarmente aggiornato sulla base di una metodologia che utilizza una pluralità di fonti per l'aggiornamento delle informazioni necessarie al disegno
  3. TERZO STADIO: le referenze e le possibilità inedite proposte dall'utilizzo degli scanner data anche di valutazione della bontà dei diversi schemi di selezione campionaria adottabili

## 2. Lo stato di avanzamento delle attività sugli «scanner data»: il tavolo con la GDO

- È stato avviato a partire dalla fine del 2013, un tavolo informale di lavoro con la Grande Distribuzione Organizzata (GDO), rappresentata dall'Associazione della Distribuzione Moderna (ADM);
- Con ADM è stata concordata una triangolazione per l'acquisizione degli scanner data (Istat richiede alla GDO la fornitura dei dati in formato elettronico per il calcolo degli indici dei prezzi al consumo e la GDO autorizza Nielsen a fornire a Istat i dati trasmessi a Nielsen dalla GDO);
- Attenzione concentrata su sei catene della GDO (Coop, Conad, Selex, Esselunga, Auchan, Carrefour) che coprono circa il 57% del valore delle vendite con riferimento ai prodotti grocery;

## 2. Lo stato di avanzamento delle attività sugli «scanner data»: il tavolo con la GDO

- Il flusso dei dati è stato avviato a fine 2014 e riguarda finora sei province (Torino, Ancona, Palermo, Piacenza, Cagliari, Roma);
- Delle sei province e con riferimento alle sei catene prima elencate, sono stati forniti i dati di valore delle vendite e quantità settimanali, a partire da dicembre 2012 (per Roma da febbraio 2013), per tutti gli EAN code dei prodotti grocery commercializzati, riferiti a tutti gli ipermercati e i supermercati. Questi flussi vengono aggiornati mensilmente;
- Entro l'anno le province, scelte in modo tale da coprire il 55% della popolazione italiana, per le quali riceveremo i dati, saranno 37;



## 2. Lo stato di avanzamento delle attività sugli «scanner data»: il flusso dei dati

Fornitura scanner data 2015 – copertura 55% popolazione italiana

### Da dicembre 2012

1. Torino
2. Ancona
3. Palermo
4. Piacenza
5. Cagliari
6. Roma

### Da dicembre 2013

7. L'Aquila
8. Potenza
9. Reggio-Calabria
10. Napoli
11. Bologna
12. Trieste
13. Genova
14. Milano
15. Campobasso
16. Bari
17. Firenze
18. Trento
19. Venezia

### Da dicembre 2014

20. Pescara
21. Catanzaro
22. Ravenna
23. Modena
24. Udine
25. Bergamo
26. Varese
27. Pesaro-Urbino
28. Foggia
29. Sassari
30. Catania
31. Messina
32. Lucca
33. Pisa
34. Perugia
35. Aosta
36. Verona
37. Padova

## 2. Lo stato di avanzamento delle attività sugli «scanner data»: i test

- Elaborazione di indici sperimentali paralleli a quelli elaborati sulla base della rilevazione attuale (per il Comune di Torino) per:
  - a. riprodurre esattamente l'attuale campione e metodologia di aggregazione e confrontare i risultati ottenuti con gli indici attualmente elaborati, al fine di valutare l'impatto dell'introduzione degli scanner data nella stima dell'inflazione;
  - b. testare diversi schemi di aggregazione degli indici elementari, comparare gli indici e valutare la possibilità di superare gli attuali schemi di aggregazione basati su medie geometriche semplici (per esempio utilizzando gli indici di Lowe)**

## 2. Lo stato di avanzamento delle attività sugli «scanner data»: i test

- c. **testare diversi schemi di campionamento e aggregazione dei dati elementari per verificare quali siano in grado di ridurre l'errore campionario in funzione del parametro stimato sull'universo grazie agli scanner data**
- d. Il parametro dell'universo sarà calcolato «localmente» e cioè per i prodotti grocery, per la provincia di Torino, per la GDO e se possibile per catena, estendendo poi la sperimentazione ad altre province.

## 2. Lo stato di avanzamento delle attività sugli «scanner data»: la sperimentazione 2016

- da ottobre sarà avviata la sperimentazione del nuovo disegno campionario nei diversi stadi (a partire quindi dalla selezione dei punti vendita) e del calcolo di indici sulla base degli scanner data; in particolare:
  - i. a fine settembre sarà rilasciata una versione definitiva dell'archivio ASIA-PV, tenendo conto del disegno campionario bilanciato con la sola indicazione, per ciascun punto vendita, che siano venduti o meno i prodotti di un determinato raggruppamento COICOP
  - ii. entro ottobre sarà estratto il campione dei punti vendita per i Comuni per i quali è in corso la fornitura degli scanner data che aderiranno alla sperimentazione

## 2. Lo stato di avanzamento delle attività sugli «scanner data»: la sperimentazione 2016

- iii. i Comuni che parteciperanno alla sperimentazione effettueranno quindi la rilevazione negli stessi punti vendita per i quali abbiamo a disposizione gli scanner data, seguendo le indicazioni tradizionali per la selezione delle referenze (la/e più venduta/e)
- iv. entro novembre sarà selezionato il campione di referenze per la simulazione del calcolo di indici provinciali sulla base degli scanner data per i raggruppamenti COICOP dei prodotti grocery utilizzando uno o più schemi di campionamento
- v. indagine ed elaborazione parallele saranno condotte per tutto il 2016 per valutare se e con che modalità portare in produzione la nuova fonte di dati da gennaio 2017

## 2. Lo stato di avanzamento delle attività sugli «scanner data»: il workshop europeo

- è in corso la preparazione del workshop europeo dell'1 e 2 ottobre sugli scanner data, che l'Istat ha l'onore di ospitare nell'aula magna di Via Balbo
- nel corso del workshop i ricercatori impegnati nel progetto svolgeranno numerose presentazioni (quattro) e ADM interverrà con una relazione che tratterà il quadro della situazione del commercio al dettaglio, della GDO e delle prospettive future del mercato
- nel corso del workshop sarà discussa una prima bozza delle raccomandazioni Eurostat sull'ottenimento degli scanner data

### 3. Lo stato di avanzamento delle attività sulle tecniche di «web scraping»

- le sperimentazioni finora condotte hanno riguardato i prodotti dell'elettronica di consumo e le tariffe aeree ed è stato utilizzato iMacros, software licenziato, acquistato dall'Istat.
- Il focus dell'attività è stato indirizzato alla riproduzione della rilevazione svolta "manualmente" mediante tecniche di «copy and paste»
- Nel corso delle sperimentazioni si sono proposti problemi di duplice natura:
  - a) problemi di «lista» (individuazione dell'universo) e di disegno campionario
  - b) accesso ai siti (paratie alzate contro i robot)
  - c) si sono evidenziati i limitati guadagni di efficienza con riferimento alle tariffe aeree

### 3. Lo stato di avanzamento delle attività sulle tecniche di «web scraping»

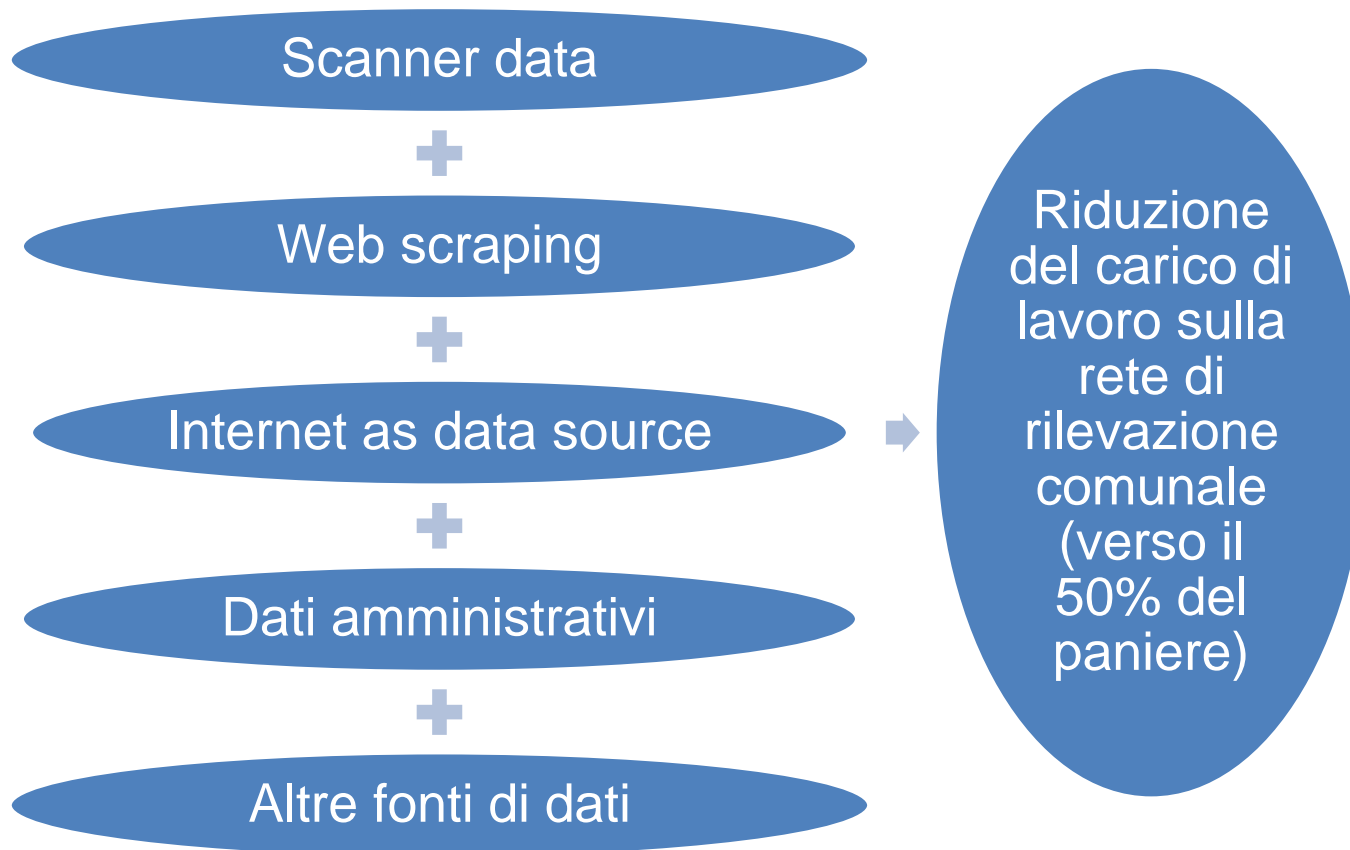
- Alla luce delle evidenze emerse i problemi che saranno affrontati nel prossimo periodo possono essere così riassunti:
  - A. Valutare compiutamente il peso dell'e-commerce e valutarne le conseguenze nella definizione dell'archivio ASIA-PV e della revisione della strategia campionaria
  - B. consolidare i risultati ottenuti
  - C. esplorare il passaggio all'utilizzo di tecniche di web scraping in una logica maggiormente orientata ai big data;
  - D. passare alla sperimentazione del web scraping (abbigliamento e calzature) e più in generale di Internet come fonte di dati (esercizi ricettivi) per merceologie rilevate a livello territoriale



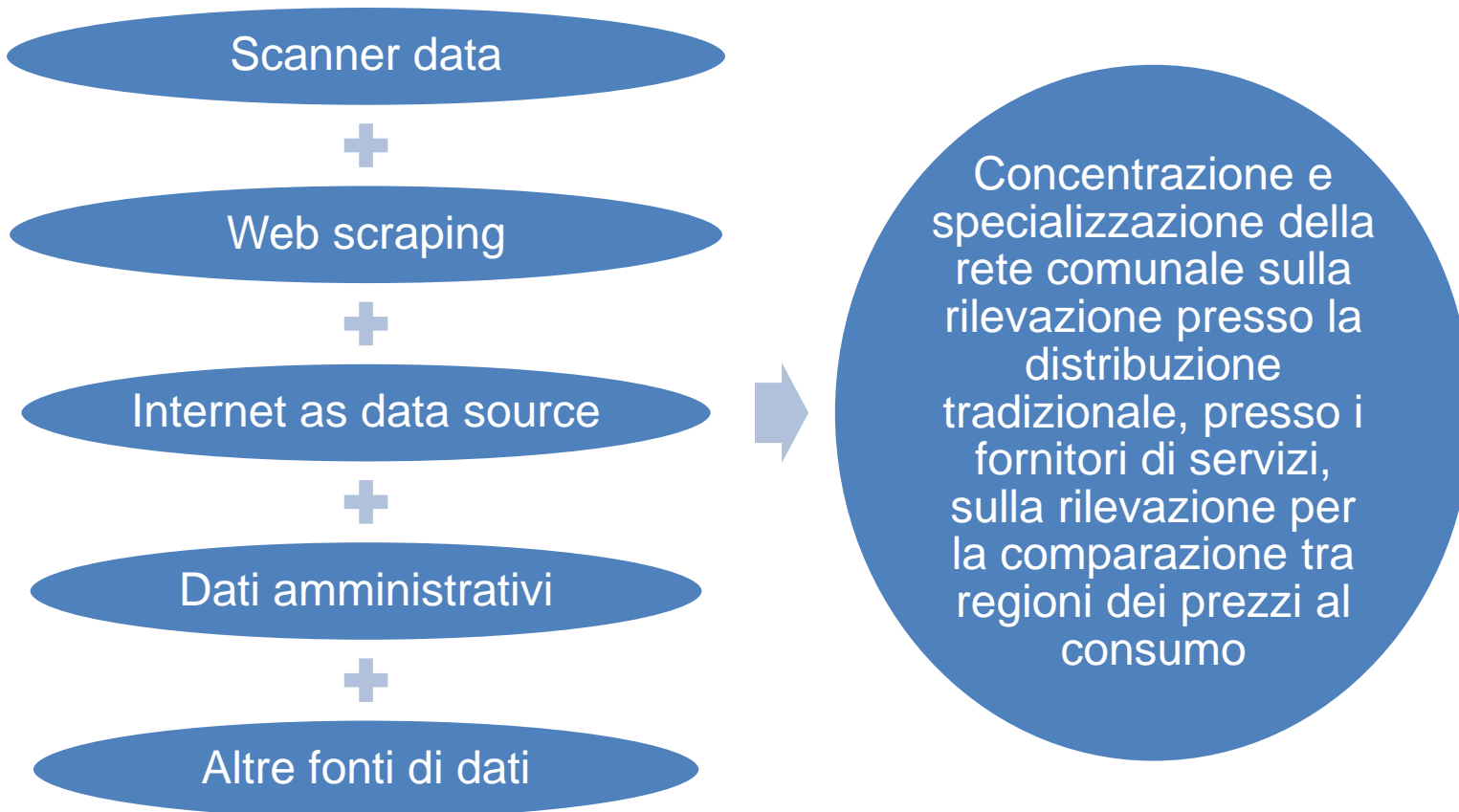
## 4. Il progetto di innovazione e le prospettive di riorganizzazione della rete comunale e dell'informazione territoriale

- L'attività in corso con il Ministero dello sviluppo economico per valutare l'usabilità ai fini del calcolo dell'inflazione della base dati delle quotazioni dei carburanti rilevate dal Ministero nell'ambito dei propri compiti di vigilanza
- La possibilità di un primo utilizzo delle informazioni di questa base dati per i Comuni che partecipano all'indagine per una parte del paniere
- Le esplorazioni in corso con Quattroruote per accedere alle informazioni dei listini sui pezzi di ricambio e aprire la strada a una revisione della rilevazione dei prezzi dei diversi servizi di manutenzione dell'auto (costo contrattuale manodopera + costo pezzo di ricambio)

## 4. Il progetto di innovazione e le prospettive di riorganizzazione della rete comunale e dell'informazione territoriale



## 4. Il progetto di innovazione e le prospettive di riorganizzazione della rete comunale e dell'informazione territoriale



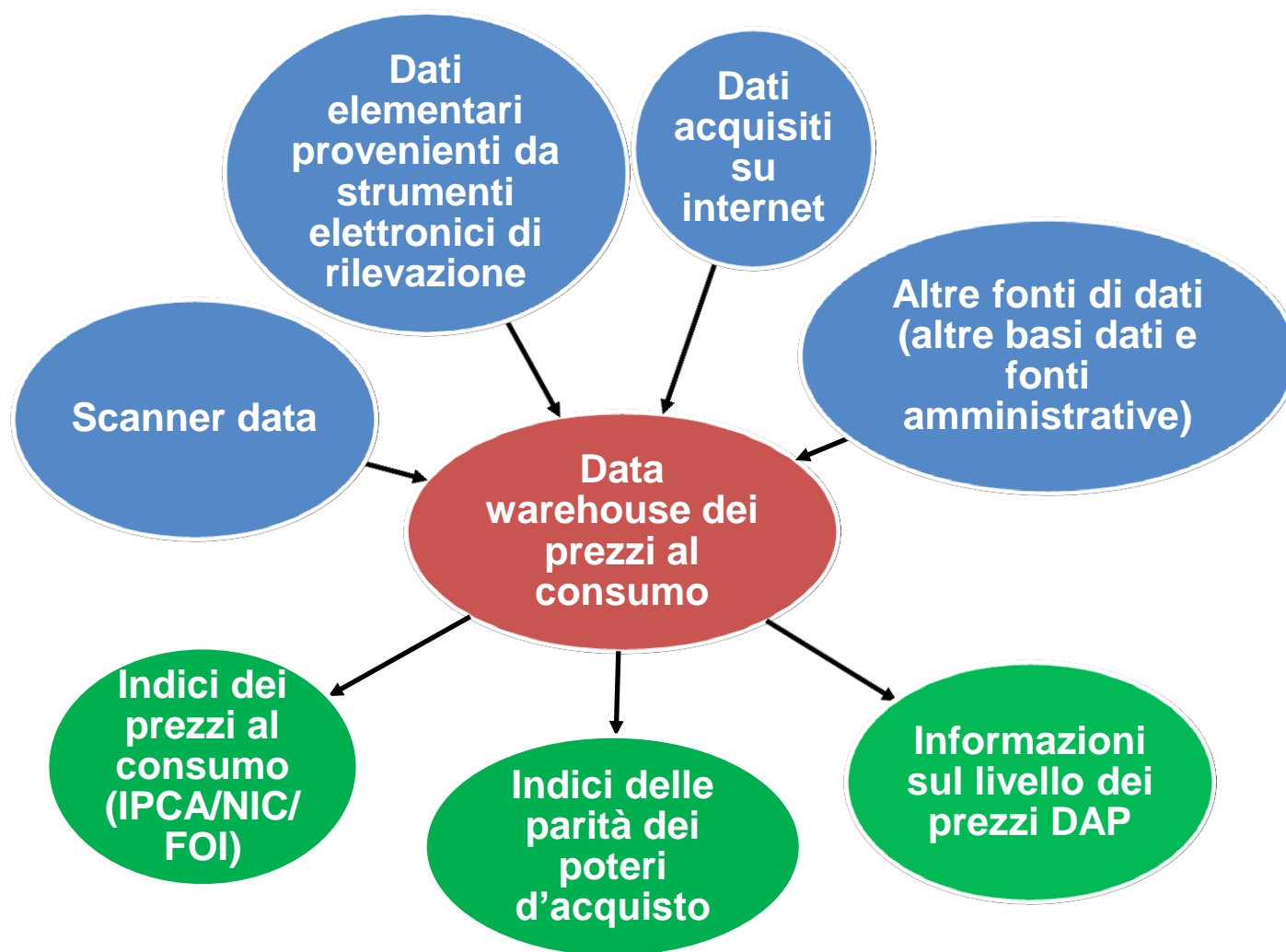
## 4. Il progetto di innovazione e le prospettive di riorganizzazione della rete comunale e dell'informazione territoriale

- La stella polare del progetto di innovazione dell'indagine sui prezzi al consumo: **il miglioramento della qualità**
- L'accuratezza in termini sia di controllo dell'errore non campionario sia di controllo dell'errore campionario è l'elemento che stiamo aggredendo con determinazione
- La forte territorialità di alcune delle nuove fonti di dati preludono a una proiezione sull'informazione territoriale di alcune novità che ci stiamo prefiggendo sul scala nazionale (la diffusione degli indici per tipologia distributiva)
- Le potenzialità che si stanno aprendo per una stabile produzione di indicatori di comparazione spaziale infra nazionale dei prezzi al consumo

## 5. Gli obiettivi strategici

- Nel 2017-18 portare a compimento la modernizzazione dell'indagine sui prezzi al consumo sia sotto il profilo campionario sia sotto il profilo della raccolta dei dati elementari
- Conseguire un ulteriore miglioramento della qualità e dell'efficienza
- Colmare il gap informativo e rispondere all'ulteriore e crescente articolazione della domanda di informazione statistica sui prezzi al consumo
- Concretizzare l'auspicata centralità dell'IPCA
- Ridefinire dentro il contesto della modernizzazione dell'Istat la macchina organizzativa delle statistiche sui prezzi al consumo a partire dal Servizio prezzi al consumo, passando per gli uffici territoriali dell'Istat fino agli Uffici Comunali di statistica

## 5. Gli obiettivi strategici



**Grazie per l'attenzione**