

Big data, Open data, Statistical data
Dalla qualità dei dati alla qualità dell'informazione statistica

Il progetto Archimede: i nuovi prodotti per le statistiche comunali

Seconda sessione: Le nuove frontiere della statistica ufficiale

Trieste, 26 settembre 2014

Il progetto ARCHIMEDE: ARCHivio Integrato di Microdati Economici e Demo-sociali

Sperimentazione di percorsi metodologici per ampliare l'offerta informativa mediante produzione di collezioni di dati elementari di tipo longitudinale e *crossection*, utili alla ricerca sociale ed economica, alla programmazione territoriale e settoriale, alla valutazione delle politiche pubbliche a livello nazionale, regionale e locale.

Valorizzazione dei contenuti informativi di fonti amministrative «integrate» presenti nel Sistema Informativo di Microdati

Inversione del flusso delle informazioni elementari (esclusivamente per finalità statistiche) dall'Istat ad altri enti pubblici

Riutilizzo di metodi, modelli e «sintassi» già sviluppati dentro e all'esterno dell'Istat

Sotto
Progetto 1
P&P

*Sistema
informatico
Persons and
Places*

- **Analisi della mobilità sul territorio**
 - Matrici origine/destinazione per ambiti territoriali
 - Identificazione delle tipologie di «*city users*»
 - Analisi dei bacini universitari

Sotto
Progetto 2

PL
*Precarietà
Lavorativa*

- **Analisi delle caratteristiche economico-lavorative e degli aspetti socio-anagrafici di un universo di individui definiti «*lavoratori precari*»**
 - Concetti di atipicità e professionalità (elementi oggettivi)
 - Tipologie contrattuali utilizzate in forma impropria (P.IVA monocommittenti)
 - Analisi Longitudinale

Sotto
Progetto 3

CF
*Condizione
socio-
economica
delle
famiglie*

- **Analisi della condizione socio-economica delle famiglie**
 - Diverse dimensioni (tipologia familiare, reddito, istruzione, ecc.)
 - Reddito (integrazione dei redditi a tass. ordinaria con redditi esenti)
 - Classificazione delle famiglie mediante l'utilizzo di tecniche tipo Data Mining

Progetto ARCHIMEDE – i tracciati delle basi dati

Var. nuove (TR)	Descrizione
cap_df	CAP domicilio fiscale
cap_lav1	cap lavoro 1
cap_lav2	cap lavoro 2
cap_lav3	cap lavoro 3
cap_res	CAP di residenza
cap_scu	CAP Scuola
cap_uni	CAP facoltà/università
codice_individuo	Codice identificativo individuo SIM
ind_df	indirizzo di domicilio fiscale
ind_lav1	indirizzo lavoro 1
ind_lav2	indirizzo lavoro 2
ind_lav3	indirizzo lavoro 3
ind_res	indirizzo di residenza
ind_scu	indirizzo scuola

Variabile	Descrizione	Modalità (eventuale)
anno_n	Anno	Ottobre 2010 / Ottobre 2011
cfiscritto	Codice fiscale lavoratore	
cfsedeservizi_o	Codice fiscale datore	
sesto	Sesso	M / F
eta	Età	
provenienza_n	Fonte di provenienza	1. ASIA OCCUPAZIONE 2. POSIZIONI ASSICURATIVE EX INPDAP 3. AUTONOMI AGRICOLTURA 4. LAVORATORI DOMESTICI
OUT2_CON D	Numero di attività lavorative svolte	
OUT2_SEDI	Numero di datori di lavoro	
OUT1_FONTI	Numero di fonti di provenienza	
pos_ott_out1	Posizione lavorativa mensile	0 ÷ 1
pos_ott_out2	Posizione lavorativa mensile complessiva	0 ÷ 1
cig_out2	Evento di CIG e/o Contratto di solidarietà	0 / 1
condiz	Tipologia lavorativa 1	1. TI FULL TIME 2. TI PART TIME 3. AUTONOMO 4. ALTRO DIPENDENTE 5. TD 6. COLLABORATORE 7. DOMESTICO
condiz_bis	Tipologia lavorativa 2	1. AUTONOMO E TI NON CIG 2. TI IN CIG 3. ALTRO DIPENDENTE 4. TD 5. COLLABORATORE 6. DOMESTICO
condpri	Condizione di precarietà	1. STABILE 2. ALTRO DIPENDENTE 3. PRECARIO
RC005DIP	Reddito da lavoro dipendente e assimilato al lordo degli straordinari, rideterminato dal MEF - Anno 2011	
REDD_COM PL	Reddito complessivo lordo, rideterminato dal MEF - Anno 2011	

Variabile	Descrizione
COD_PRO	Codice Istat Provincia
COD_COM	Codice Istat Comune
CODICE_FAM	Codice identificativo famiglia
NCOMP	Numero componenti da LAC
N_SEZ_RIC	Numero sezione di censimento
TIP_FAM_2	Tipologia familiare
NCOMP_CLETA1	Numero di componenti tra 0 e 14 anni
NCOMP_CLETA2	Numero di componenti tra 15 e 34 anni
NCOMP_CLETA3	Numero di componenti tra 35 e 64 anni
NCOMP_CLETA4	Numero di componenti con 65 e più anni
NCOMP_DIS	Numero di componenti con disabilità
NCOMP_DIS_CLETA1	Numero di componenti tra 0 e 14 anni con disabilità
NCOMP_DIS_CLETA2	Numero di componenti tra 15 e 34 anni con disabilità
NCOMP_DIS_CLETA3	Numero di componenti tra 35 e 64 anni con disabilità
NCOMP_DIS_CLETA4	Numero di componenti con 65 e più anni con disabilità
CITTAD_FAM	Classificazione famiglia in base alla cittadinanza dei componenti
NPENS_REDD_MIN	Numero di pensionati con reddito minimo presenti in famiglia (numero di percettori di pensione/assegno sociale)
NCOMP_CLETA1_M	Numero di componenti tra 0 e 14 anni di sesso maschile
NCOMP_CLETA2_M	Numero di componenti tra 15 e 34 anni di sesso maschile
NCOMP_CLETA3_M	Numero di componenti tra 35 e 64 anni di sesso maschile
NCOMP_CLETA4_M	Numero di componenti con 65 e più anni di sesso maschile
NCOMP_CLETA1_F	Numero di componenti tra 0 e 14 anni di sesso femminile
NCOMP_CLETA2_F	Numero di componenti tra 15 e 34 anni di sesso femminile
NCOMP_CLETA3_F	Numero di componenti tra 35 e 64 anni di sesso femminile
NCOMP_CLETA4_F	Numero di componenti con 65 e più anni di sesso femminile
REDD_FAM	Somma del reddito (REDD_CALCLO4) dei componenti
CAPRE_FAM	Ammontare del reddito da capitale reale posseduto dalla famiglia
CAPFIN_FAM	Ammontare del reddito da capitale finanziario posseduto dalla famiglia
TRASFPRIV_FAM	Ammontare del reddito da trasferimenti privati posseduto dalla famiglia
TRASFPUUBL_FAM	Ammontare del reddito da trasferimenti pubblici posseduto dalla famiglia
PENS_FAM	Ammontare del reddito da pensione posseduto dalla famiglia
LAVDIP_FAM	Ammontare del reddito da lavoro dipendente posseduto dalla famiglia
LAVAUT_FAM	Ammontare del reddito da lavoro autonomo posseduto dalla famiglia
N_CONTRIB_FAMIGLIA	Numero di contribuenti BDR
N_PERC_FAMIGLIA	Numero di percettori di reddito
N_STUD PRIM	Numero di studenti della scuola primaria
N_STUD_sec1	Numero di studenti della scuola secondaria di I grado
N_STUD_sec2	Numero di studenti della scuola secondaria di II grado
N_UNI_1	Numero di studenti universitari di età inferiore ai 35 anni
N_UNI_2	Numero di studenti universitari di almeno 35 anni
N_OCC_CLETA1	Numero di occupati nel mese di ottobre tra 0 e 14 anni
N_OCC_CLETA2	Numero di occupati nel mese di ottobre tra 15 e 34 anni
N_OCC_CLETA3	Numero di occupati nel mese di ottobre tra 35 e 64 anni
N_OCC_CLETA4	Numero di occupati nel mese di ottobre con 65 e più anni
N_PREC_CLETA1	Numero di occupati precari nel mese di ottobre tra 0 e 14 anni
N_PREC_CLETA2	Numero di occupati precari nel mese di ottobre tra 15 e 34 anni
N_PREC_CLETA3	Numero di occupati precari nel mese di ottobre tra 35 e 64 anni
N_PREC_CLETA4	Numero di occupati precari nel mese di ottobre con 65 e più anni
N_NOSOR_F	Numero di persone in età attiva (15-64) che non studiano, non lavorano senza reddito di sesso femminile
N_NOSOR_M	Numero di persone in età attiva (15-64) che non studiano, non lavorano senza reddito di sesso maschile
TITSTU_CAPOFAMIGLIA	Titolo di studio del capofamiglia. Il capofamiglia è il percettore di reddito più alto all'interno della famiglia; nel caso in cui non ci siano percettori di reddito coincide con l'istatario.
POS_ANNO	Somma delle posizioni medie annue dei componenti occupati
INT_LAV	Intensità lavorativa: rapporto fra il numero totale di mesi lavorati dai componenti della famiglia di età compresa tra i 18 e i 59 anni (esclusi gli studenti 18-24enni) e il numero totale di mesi teoricamente disponibili
N_AMMORTIZZ	Numero di componenti che hanno beneficiato di almeno un assegno di disoccupazione o di mobilità
Sesso_int	Sesso dell'istatario (1.maschio; 2.femmina)
occupazione_int	Flag 0-1 che indica se l'istatario ha lavorato almeno un giorno nell'anno
app_redd_int	Apporto dell'istatario al reddito familiare (percentuale)
N_neet_m	Numero di persone di età compresa tra i 15 e i 29 anni che non studiano e non lavorano di sesso maschile
N_neet_f	Numero di persone di età compresa tra i 15 e i 29 anni che non studiano e non lavorano di sesso femminile
N_cleta_neet	Numero di persone di età compresa tra i 15 e i 29 anni

- 2013** - Realizzazione sperimentale di tre basi di dati:
P&P – Nazionale
Precarietà Lavorativa (PL) – Nazionale
Condizione socio-economica delle famiglie (CF) – comuni Bologna /Firenze /Modena

/Trieste/Brescia

- 2014** - Progetto Archimede inserito nel PSN 2014-2016:
sperimentazione delle tre basi realizzate da parte degli Uffici di statistica *compartecipanti* :
- Comuni di Firenze, Bologna, Brescia, Trieste, Modena, Milano
- Regioni e Prov. Autonome
Invio lettera informativa ai compartecipanti
Trasmissione dei microdati successiva alla pubblicazione del PSN
Risultati della fase di test (feedback) entro sei mesi
Elaborazione di un cruscotto di indicatori

- 2015** - Aggiornamento PSN 2015-2016:
Ingegnerizzazione della produzione delle basi dati
Diffusione micro/macro a tutti gli utenti Sistan delle tre basi di dati
Nuove sperimentazioni su:
- Percorsi di istruzione, formazione, abbandono scolastico e inserimento lavorativo degli individui tra i 16-29 anni.

Supporto
allo sviluppo
del progetto
GdL 5

Fase di test: valutazione delle completezza, congruità e utilità delle basi di microdati

UFF. di Stat	P&P	PL	CF
Com. Firenze	X (mobilità)	X (struttura occupazionale)	X (marginalità sociale e povertà)
Com. Milano	X (mobilità inter/intra comunale)	X (precaricato)	
Com. Trieste	X (diff. inform. rispetto all'anagrafe)		X (analisi redditi)
Com. Brescia			X (autonomia familiare)
Com. Modena		X (dinamiche occupazionali)	X (analisi redditi)

Obiettivo: identificare un sistema (*più sistemi*) di indicatori – per tipologia di ente e/o ambito tematico – a partire dalle collezioni di dati elementari

- Supporto ad uffici Sistan con limitata capacità elaborativa
- Omogeneizzazione degli strumenti per la valutazione

Opportunità: valutazione degli indicatori prodotti attraverso dati di benchmark (in particolare BES o URBES)

A che punto siamo?

Elaborazione di tre «cruscotti», uno per ogni sotto-progetto, che verranno discussi all'inizio di ottobre.

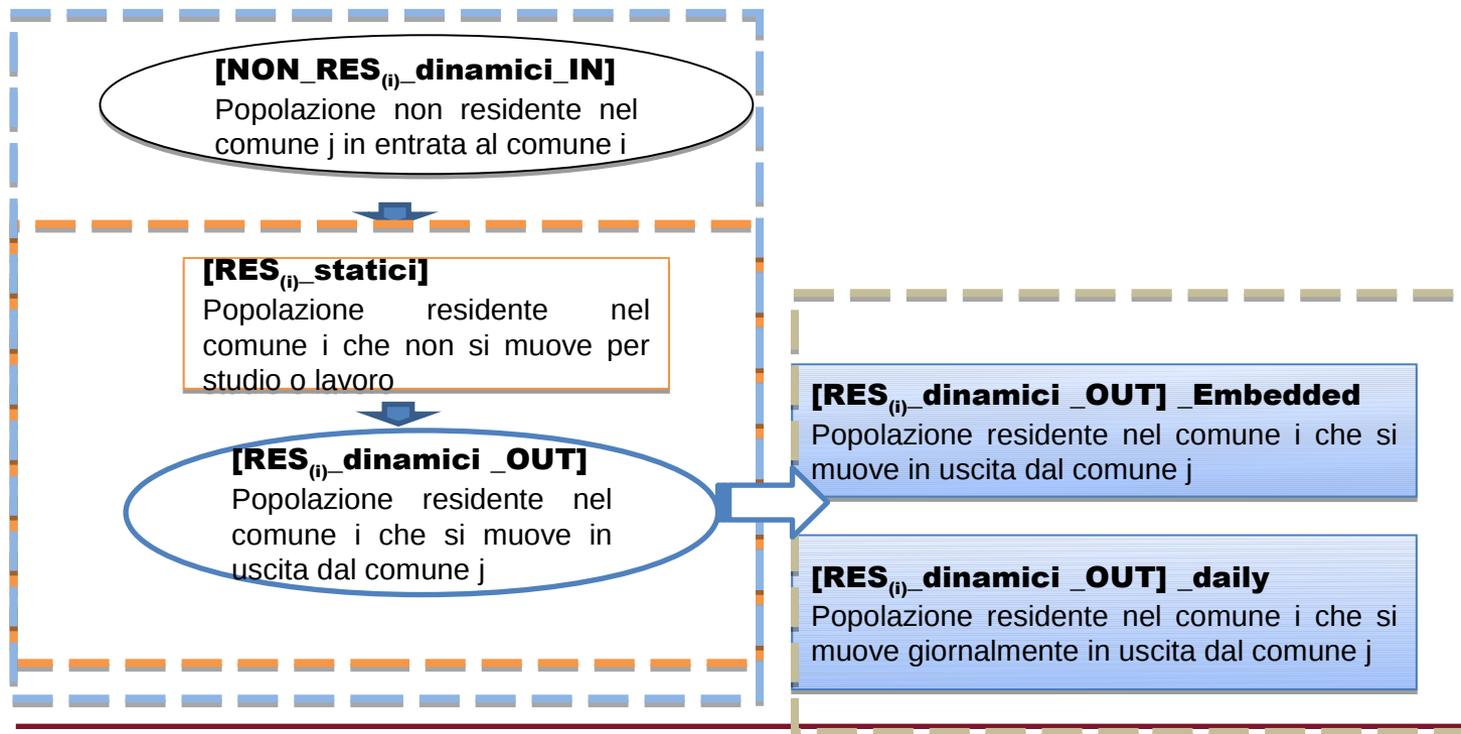
Un assaggio...

Indicatori su P&P (1/2)

Popolazione Residente: è la quota di popolazione che è iscritta in anagrafe nel comune i
[RES_{(i)_statici}] + [RES_{(i)_dinamici_OUT}]

Popolazione afferente_(i): è la quota di popolazione iscritta in anagrafe nel comune i più la quota di persone che lavorano o studiano nel comune i senza risiedervi
[RES_{(i)_statici}] + [RES_{(i)_dinamici_OUT}] + [NON_RES_{(i)_dinamici_IN}]

Popolazione Residente Dinamica: è la quota di popolazione che risiede nel comune i ma che studia o lavora in un altro comune
[RES_{(i)_dinamici_OUT}]_{daily} + [RES_{(i)_dinamici_OUT}]_{Embedded}



Indice di esternalità

$$\bullet \text{NON_RES}_{(t)_d} / \text{RES}_{(t)_s}$$

- Specifica la tendenza di un comune a richiamare domanda di lavoro o studio dall'esterno.
- E' il rapporto tra i flussi in entrata e i residenti statici.
- Più è basso maggiore sono gli statici, più è alto maggiore è il flusso in entrata.

	Firenze	Trieste
indice di esternalità	1,1	0,4
--per lavoro	1,1	0,4
--per università	3,9	2,4

Tasso di attrazione

$$\bullet \text{NON_RES}_{(t)_d} / (\text{Popolazione Afferente})_{(t)}$$

- Indica la capacità di un'area di attirare la domanda di lavoro o studio dall'esterno.
- La quantità dei flussi in ingresso è rapportata ai flussi generati dall'area (flussi che si esauriscono all'interno dell'area e flussi che escono dall'area stessa).
- Varia tra 0-100. Più è alta la percentuale maggiore è l'incidenza della componente dinamica in entrata.

	Firenze	Trieste
tasso di attrazione	44,8	24,8
--per lavoro	43,0	21,7
--per università	76,5	66,6

Indicatori su Precarietà lavorativa (1/2)

- **Lavoro precario:** (a) lavoro dipendente a termine (con scadenza) e (b) rapporti di collaborazione o lavoro autonomo consistenti in attività aventi caratteristiche di subordinazione
- **Lavoratore precario:** individuo che, in un determinato periodo di tempo, svolge esclusivamente lavori definiti «precari»

Tipologia	Categoria contrattuale / Rapporto di lavoro
Lavoro dipendente	Lavoro a tempo determinato - Full-time e Part-time (inclusi lavoratori domestici e contratti di somministrazione)
	Contratto di Formazione e Lavoro - CFL
	Apprendistato
	Contratto d'inserimento
Lavoro parasubordinato (Fonte Gestione separata Inps)	Collaboratore di giornali
	Dottorato di ricerca, assegno, borsa di studio
	Co.co.co. (con contratto a progetto)
	Venditore porta a porta
	Collaboratori occasionali (mini Co.co.co.)
	Rapporti occasionali autonomi
	Co.co.co. presso la PA
	Rapporti Co.co.co. prorogarti
	Associati in partecipazione
	Medici in formazione specialistica
Lavoro occasionale di tipo accessorio – <i>Voucher</i>	

Tasso di precarietà

- Esprime la quota di occupati precari sul totale degli occupati nel periodo di riferimento. Ottobre 2011.
- Comune di lavoro.

Tasso di precarietà	Firenze	Trieste
Totale	14,8	12,6
<i>Maschi</i>	10,9	10,3
<i>Femmine</i>	19,6	15,3

Tasso di precarietà tra i nuovi occupati

- Restituisce informazioni sulla frequenza di utilizzo di nuovi contratti precari. Ottobre 2011.
- E' il rapporto tra i nuovi lavoratori precari (non occupati nel periodo precedente) e la totalità dei nuovi occupati.
- Comune di lavoro.

Tasso di precarietà nuovi occupati	Firenze	Trieste
Totale	18,6	17,6

Caratteristiche dei residenti nei comuni di Firenze e Trieste. Anno 2011

	Firenze	Trieste
Numero Totale famiglie	185.023	107.033
Unipersonali	77.559	49.868
Coppie senza figli	27.488	19.167
Coppie con figli	32.934	18.545
Monogenitore	18.269	9.701
Altro	28.773	9.752
Ammontare reddito (in milioni di euro)	6.962,6	3.556,7
<i>N famiglie (con reddito > 0)</i>	<i>172.461</i>	<i>101.041</i>
Reddito equivalente mediano	20.726	21.244

Percentuale di famiglie con intensità di lavoro molto bassa

- L'intensità è calcolata per ogni famiglia residente rapportando il numero di mesi lavorati dai componenti di età compresa tra 18 e 59 anni (con esclusione degli studenti nella fascia di età tra i 18 e i 24 anni) e il numero totale dei mesi dell'anno.
- Per bassa intensità lavorativa si intende un valore inferiore a 0,2

Famiglie con bassa intensità di lavoro	Firenze	Trieste
% famiglie	3,29	3,27
N famiglie	3.441	1.841

Percentuale di famiglie in cui è presente almeno un giovane tra i 15 e i 29 anni che non studia né lavora

- Numero famiglie in cui almeno un giovane (15-29 anni) non risulta né occupato né iscritto a scuola o università/Totale famiglie in cui è presente almeno un giovane tra i 15 e i 29 anni *100

Presenza di neet in famiglia	Firenze	Trieste
% famiglie	27,9	27,7
N famiglie	36.612	19.762

Concetti chiave

Compartecipazione -- Lavorare nell'ottica della convergenza, anche nella realizzazione di nuovi percorsi di sperimentazione (istruzione e formazione)

Diffusione -- Far circolare le basi di microdati per verificarne la congruità rispetto alle esigenze conoscitive degli enti pubblici (Sistan)

Semplificazione -- Dare la possibilità a tutti i comuni, anche quelli che non hanno la sufficiente autonomia per elaborare i dati elementari, di utilizzare dati macro di sintesi

Grazie per
l'attenzione