

ANALISI DEI DATI DI SALUTE PUBBLICA CON TECNICHE GIS E DI GEOSTATISTICA

Vittorio Casella^a; Marica Franzini^a; Barbara Padova^a; Riccardo Bellazzi^b; Simona Dalle Carbonare^{b,c}; Carlo Cerra^c

^aDipartimento di Ingegneria Edile e del Territorio, Università di Pavia - (vittorio.casella, marica.franzini, barbara.padova)@unipv.it

^bDipartimento di Informatica e Sistemistica, Università di Pavia - riccardo.bellazzi@unipv.it, simona.dallecarbonare@gmail.com

^cAzienda Sanitaria Locale di Pavia - simona.dallecarbonare@gmail.com, Carlo_Cerra@asl.pavia.it

KEY WORDS: GIS, thematic map, geographic information, geostatistics.

RIASSUNTO

La nota illustra alcuni risultati della collaborazione fra la ASL di Pavia, il Laboratorio di Informatica Biomedica "Mario Stefanelli" del Dipartimento di Informatica e Sistemistica dell'Università di Pavia e il gruppo di Geoinformation del Dipartimento di Ingegneria Edile e del Territorio della stessa Università.

In una prima fase della collaborazione sono state prodotte mappe tematiche per un insieme di indicatori raccolti dall'ASL per tutta la Provincia di Pavia, relativi alla salute pubblica. La presente nota verte su alcune analisi, svolte con tecniche di geostatistica, riguardanti la auto-correlazione di alcuni indicatori significativi e la cross-correlazione fra coppie di indicatori che si influenzano.

ABSTRACT

The present note describes some initial results of the collaboration between the local healthcare agency (ASL) of Pavia, the Laboratory of Biomedical Informatics "Mario Stefanelli" of the Department of Computer Engineering and System Science of the University of Pavia and the Geoinformation group of the Department of Building and Territorial Engineering, at the same University. In the first step of the collaboration, several thematic maps were produced, concerning a set of public health indicators collected by ASL for all the Province of Pavia. The present notes concerns geostatistics analysis of the ASL public health indicators: self-correlation of some of them, particularly interesting, and cross-correlation between some significant couples of them.

RIASSUNTO ESTESO

Fra i compiti istituzionali delle ASL vi è la raccolta di una quantità di indicatori riguardanti la salute pubblica: dati anagrafici, incidenza e prevalenza di numerose patologie, spesa sanitaria, uso di varie tipologie di farmaci. Molti di questi dati hanno intrinsecamente una dipendenza spaziale: è ben noto ad esempio che l'età media della popolazione, per fare un solo esempio, manifesta un *andamento*: ci sono zone più "vecchie" e zone più "giovani". I dati raccolti dalla ASL sono organizzati in modo da avere una referenza spaziale, in quanto sono organizzati per Comuni. La Figura 1 mostra l'andamento dell'età media della popolazione assistibile della Provincia di Pavia, che evidenzia un chiaro trend: la zona collinare e montuosa a Sud presenta un'età media più elevata, mentre la zona che confina con la Provincia di Milano presenta valori più bassi.

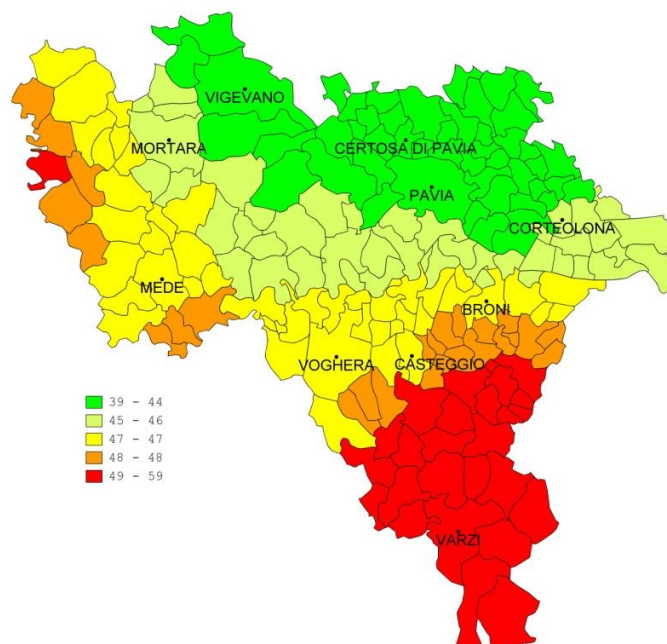


Figura 1: Età media della popolazione assistibile

I dati ASL sono stati inseriti in un ambiente ESRI ArcGIS 9.3. Ad ogni Comune sono stati associati il poligono dei suoi limiti amministrativi e un punto, localizzato nel centro geometrico del poligono.

Al punto è stata associata una tabella di database avente la struttura mostrata dalla Figura 2 e contenente 90 parametri fra quelli contenuti nel data Warehouse della ASL. Ad ogni Comune è associato un record i cui campi sono: codice ISTAT del Comune, nome in chiaro e successivamente i 90 parametri prescelti.

Come detto, i punti rappresentativi sono stati per il momento collocati nel centro geometrico dei poligoni; ciò equivale a dire che le quantità studiate si concentrano in un punto e costituisce una inevitabile approssimazione. In una fase successiva i punti rappresentativi verranno collocati nel centro del capoluogo, dove dovrebbe risiedere la maggioranza della popolazione, almeno in generale; in una ulteriore fase, la distribuzione della popolazione e delle quantità studiate verranno modellizzate come un continuo.

Nella prima fase della collaborazione sono state prodotte mappe tematiche per ciascuno degli indicatori considerati, aventi cinque classi, i cui estremi sono stati definiti secondo il criterio della uguale numerosità. Nella Figura 3 è mostrata ad esempio quella relativa alla prevalenza del diabete.

Shape	istat	nome	F5	F6	F7
Point	018050	CILAVEGNA	12.11	8	45
Point	018051	CODEVILLA	14.35	9	48
Point	018052	CONFENZA	13.81	7	47
Point	018053	COPIANO	9.18	11	43
Point	018054	CORANA	14.52	8	47

Figura 2: Estratto della tabella di database

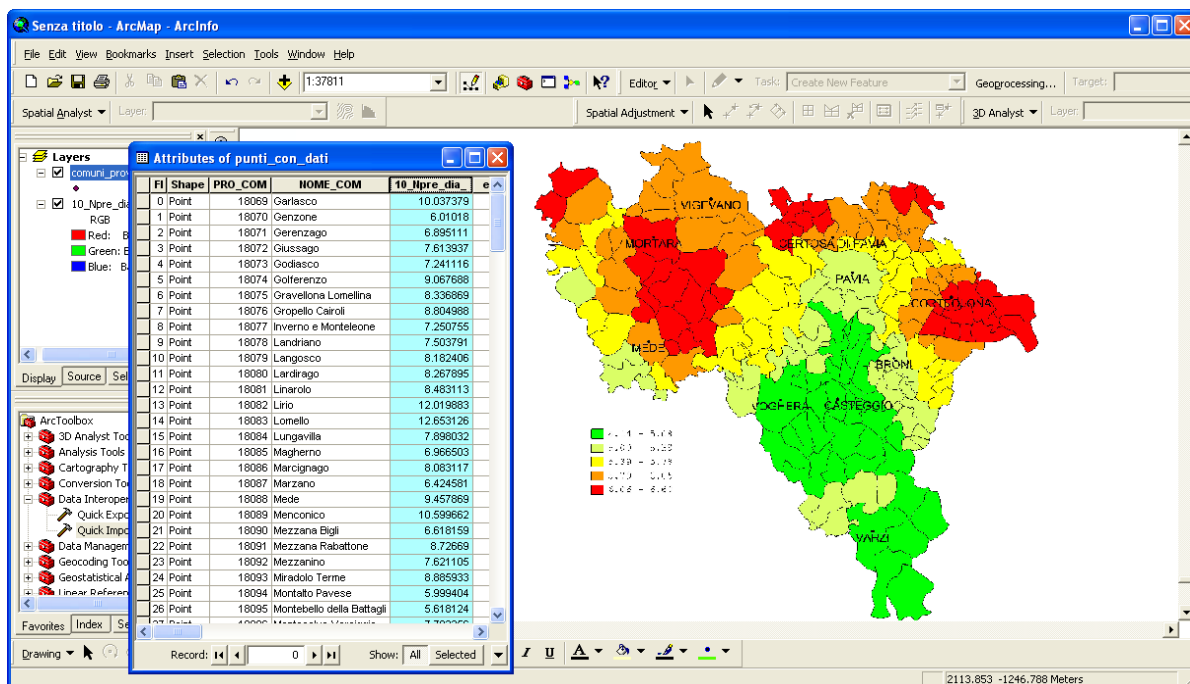


Figura 3: Mappa tematica relativa all'indicatore di prevalenza del diabete

Nella versione completa della presente nota, che verrà pubblicata dopo il Convegno su rivista ancora da designare, vengono svolte analisi con tecniche di geostatistica, per alcuni indicatori particolarmente significativi; le analisi riguardano la auto-correlazione e la cross-correlazione. Nel primo caso si vuole determinare in che misura, e fino a che distanza, la situazione in un Comune influenzi i Comuni vicini; nel secondo caso si cerca di capire se la distribuzione spaziale di un indicatore influenzi e spieghi il comportamento di un altro.