

## ESERCIZIO FACOLTATIVO – OPEN DATA LOMBARDIA

Accedere al portale Open Data della Regione Lombardia ed individuare i datasets di ambito sanitario (o utilizzare quelli già linkati in queste slides).

Analizzare tali datasets ed individuare in ognuno di essi (o in qualche unione di essi opportunamente descritta – anche utilizzando schemi ER) alcune delle metriche di efficacia, efficienza o entrambe che possano essere calcolate a partire dai dati presenti (il più possibile fedeli ai dati reali!)

Spiegare tali metriche, definirle e calcolare sui dati reali.

### EFFICIENZA

Il dataset utilizzato per la prima metrica è *Dati regionali ricoveri per DRG*, dove per DRG si intende la *diagnosis-related group*, un sistema che considera i pazienti dimessi e le risorse impiegate grazie al quale si possono fare anche delle comparazioni fra diversi ospedali.

Nel dataset proposto sono preseti delle variabili interessanti per questa analisi:

- CASI01G: numero di casi gestiti in giornata
- CASI.1G: numero di casi gestiti in più giorni
- GG: numero di giorni totali di degenza

Un processo è tanto più efficiente quanto più è ottimizzata la quantità di risorse utilizzate in base all'output ottenuto. Prendendo come riferimento le tre variabili citate, un'intuitiva metrica potrebbe essere il calcolo dei giorni medi di degenza, ottenibile come:

$$degMedia = \frac{GG}{(CASI01G + CASI.1G)}$$

Per poter fare un'analisi più approfondita si è deciso di analizzare le diverse diagnosi differenziate per anno. Per semplificare le analisi si è deciso di prendere solamente le malattie che avevano una somma annua di giorni di degenza superiore a 10000 per anno.

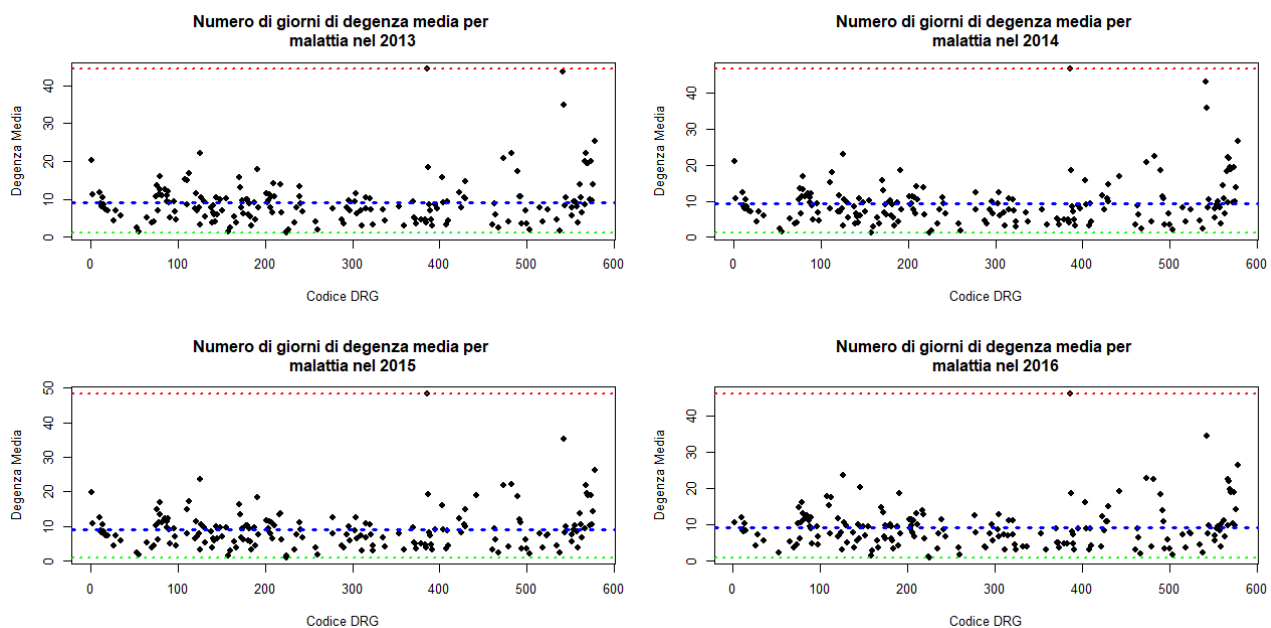
I risultati delle analisi sono stati i seguenti:

Anno	Degenza Minima (giorni)	DRG	Degenza Media (giorni)	Degenza Massima (giorni)	DRG
2013	1.08	INT. SUL PIEDE	9.10	44.29	NEONATI GRAVEMENTE IMMATURI O SINDR.DISTRESS RESP.
2014	1.16	INT. SUL PIEDE	9.23	46.59	NEONATI GRAVEMENTE IMMATURI O CON SINDROME DA DIST
2015	0.91	INT. SUL PIEDE	8.98	48.36	NEONATI GRAVEMENTE IMMATURI O CON SINDROME DA DIST
2016	0.93	INT. SUL PIEDE	9.18	46.30	NEONATI GRAVEMENTE IMMATURI O CON SINDROME DA DIST

Si nota un trend relativamente costante nella durata media dei giorni di degenza con un picco minimo nel 2015 con 8.98 giorni e un massimo di 9.23 registrato l'anno precedente. Gli interventi al piede sono quelli più comuni con i giorni medi di degenza più bassi, inferiori all'unità negli ultimi due anni, 2015 e 2016, mentre per quanto riguarda la diagnosi che richiede la degenza maggiore è quella relativa ai neonati nati prematuramente. I risultati dal punto di vista logico paiono sensati.

Da notare anche una piccola imperfezione nel dataset in cui per il 2013 la diagnosi viene chiamata *NEONATI GRAVEMENTE IMMaturi O SINDR.DISTRESS RESP.*, nome leggermente cambiato negli anni successivi. Attraverso il codice numerico della diagnosi si è potuto capire che effettivamente c'è stato un cambio nel nome e sono la stessa diagnosi, dato che il codice negli anni non è cambiato.

L'immagine successiva rappresenta i giorni medi di degenza per ogni malattia nei diversi anni, con la linea **rossa** in corrispondenza del valore massimo, quella **blu** per la media e quella **verde** per il valore minimo:

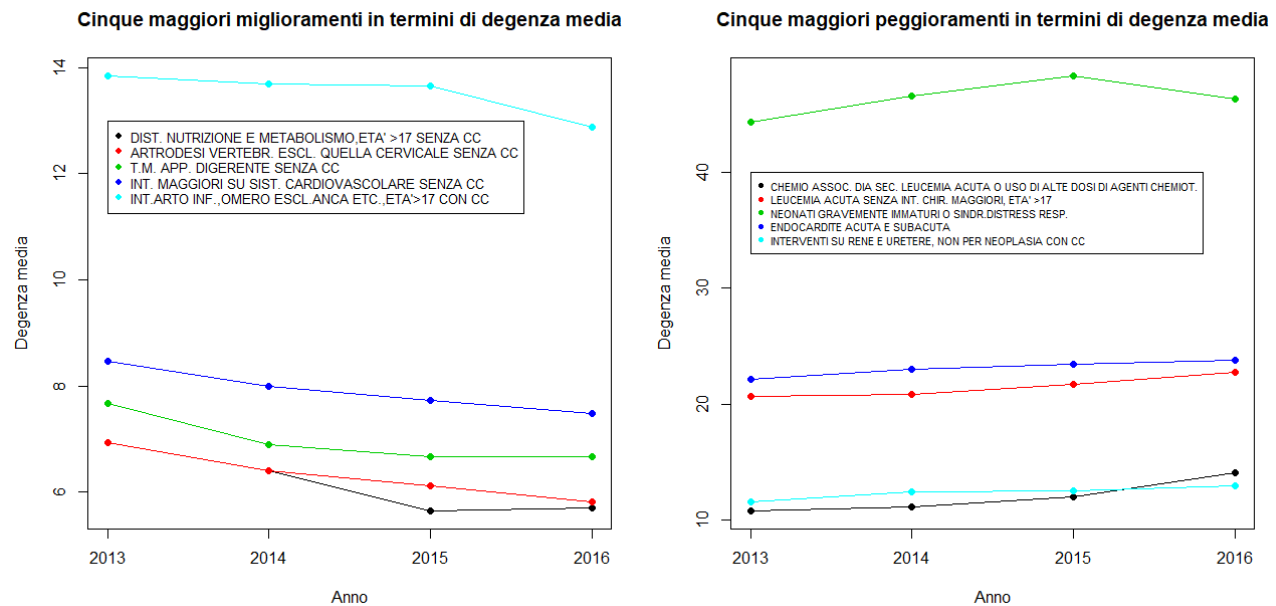


Entrando più nello specifico per l'efficienza, possiamo individuare le cinque diagnosi che hanno visto il miglioramento e il peggioramento più evidente, ovvero quelle che hanno ridotto ad aumentato il numero di giorni richiesti di degenza media. La seguente tabella riporta a sinistra le diagnosi che hanno riscontrato un miglioramento dell'efficienza fra il 2013 e il 2016 con la relativa decrescita dei giorni media di degenza, mentre a destra l'esatto opposto con i cinque peggioramenti più evidenti.

Successivamente è proposta una rappresentazione grafica dei valori in tabella.

DRG	2013	2016	Diff
<i>DIST. NUTRIZIONE E METABOLISMO, ETA' &gt;17 SENZA CC</i>	6.93	5.70	<b>-1.23</b>
<i>ARTRODESI VERTEBR. ESCL. QUELLA CERVICALE SENZA CC</i>	6.94	5.82	<b>-1.12</b>
<i>T.M. APP. DIGERENTE SENZA CC</i>	7.67	6.66	<b>-1.01</b>
<i>INT. MAGGIORI SU SIST. CARDIOVASCOLARE SENZA CC</i>	8.46	7.48	<b>-0.98</b>
<i>INT.ARTO INF.,OMERO ESCL.ANCA, ETA'&gt;17 CON CC</i>	13.85	12.87	<b>-0.98</b>

DRG	2013	2016	Diff
<i>CHEMIO ASSOC. DIA SEC. LEUCEMIA ACUTA O USO DI ALTE DOSI DI AGENTI CHEMIOT.</i>	10.75	14.06	<b>3.31</b>
<i>LEUCEMIA ACUTA SENZA INT. CHIR. MAGGIORI, ETA' &gt;17</i>	20.67	22.75	<b>2.08</b>
<i>NEONATI GRAVEMENTE IMMaturi O SINDR.DISTRESS RESP.</i>	44.29	46.30	<b>2.01</b>
<i>ENDOCARDITE ACUTA E SUBACUTA</i>	22.15	23.79	<b>1.64</b>
<i>INTERVENTI SU RENE E URETERE, NON PER NEOPLASIA CON CC</i>	11.55	12.95	<b>1.40</b>



Si nota purtroppo come i peggioramenti più evidenti nell'efficienza si sono verificati in corrispondenza di degenze che richiedono già un elevato numero di giorni in ospedale e che la diminuzione globale dei giorni di degenza è dovuto soprattutto ad un miglioramento per malattie/interventi di bassa o media serietà.

## EFFICACIA

Usando il dataset relativo alle performance degli ospedali, è possibile calcolare come metrica di efficacia il numero di pazienti che non hanno dovuto subire un ulteriore ricovero nel corso dello stesso anno nella stessa struttura ospedaliera in quanto per efficacia si intende la capacità del sistema di fornire all'utente ciò che l'utente stesso chiede al sistema. È quindi logico pensare che un paziente che sia trattato in ospedale abbia forti aspettative relative al fatto che una volta curata non debba essere ricoverato successivamente per lo stesso problema.

In questo dataset i dati sono strutturati in maniera simile a quella del primo dataset, con la differenza che in questo caso non ci sono dati relativi alle malattie/operazioni ma sulle strutture ospedaliere.

I dati sono stati filtrati in modo tale da avere solamente gli ospedali in cui si sono registrati almeno 1000 ricoveri all'anno. Successivamente si è creata una metrica di efficacia ottenuta nel seguente modo:

$$eff = 100 * \left( 1 - \frac{NUM\_RICOVERI\_RIPETUTI}{NUM\_RICOVERI\_TOTALI} \right)$$

Successivamente si sono ordinati i risultati per ottenere le 10 migliori e peggiori strutture in un determinato anno. Si è deciso di non raggruppare le strutture ospedaliere per anno, ma si è tenuto questo campo separato per confrontare eventualmente delle differenze fra ospedali nel corso del tempo.

I dati sono riportati nelle tabelle seguenti:

Struttura	Anno	Ricoveri	Ripetuti	Efficacia
C.O.F. LANZO HOSPITAL S.P.A.	2016	1096	0	100%
IST.CLINICO S.SIRO SPA-MILANO	2016	1527	0	100%
OSPEDALE DEI BAMBINI UMBERTO I - BRESCIA	2016	3218	0	100%
PRESIDIO OSPED. V. BUZZI - MILANO	2010	16775	7	99.96%
PRESIDIO OSPED. V. BUZZI - MILANO	2011	16165	7	99.96%
PRESIDIO OSPED. V. BUZZI - MILANO	2014	15418	8	99.95%
PRESIDIO OSPED. V. BUZZI - MILANO	2012	16141	9	99.94%
PRESIDIO OSPED. V. BUZZI - MILANO	2015	14343	9	99.94%
PRESIDIO OSPED. V. BUZZI - MILANO	2013	15643	13	99.92%
PRESIDIO OSPED. MACEDONIO MELLONI-MILANO	2010	10693	12	99.89%
PRESIDIO OSPED. MACEDONIO MELLONI-MILANO	2010	10693	12	99.89%

Struttura	Anno	Ricoveri	Ripetuti	Efficacia
PRESIDIO OSPEDALIERO GARDONE V.T.	2013	4619	508	89.00%
OSPEDALE CIVICO ROSSI - CASALPUSTERLENGO	2013	1451	159	89.04%
OSPEDALE DI CIRC. C.CANTU' - ABBiateGRASSO	2013	3339	361	89.19%
OSPEDALE SS.ANNUNZIATA - VARZI	2013	1300	140	89.23%
OSPEDALE S.MARTINO - MEDE I.R.C.C.S.	2013	1943	188	90.32%
POLICLINICO SAN DONATO I.R.C.C.S.	2013	14089	1350	90.42%
POLICLINICO SAN DONATO I.R.C.C.S.	2014	13891	1238	91.09%
OSPEDALE CIVICO ROSSI - CASALPUSTERLENGO	2014	1474	130	91.18%
OSP. CIRCOLO A.BELLINI - SOMMA LOMBARDO	2013	1829	149	91.85%
PRESIDIO OSPEDALIERO GARDONE V.T.	2014	4459	363	91.86%

Queste informazioni tuttavia possono presentare delle limitazioni: in primo luogo vediamo che il presidio ospedaliero di via Buzzi a Milano compare diverse volte in classifica con un'efficacia che sfiora il 100% il che significa che o ha una qualità eccelsa del servizio oppure ha dei reparti che si occupano solamente di operazioni/malattie che difficilmente poi portano ad un secondo ricovero, oppure è una struttura abilitata a casi molto gravi e quindi sfortunatamente avvengono molti decessi.

Curioso notare come nei casi peggiori le osservazioni sono tutte comprese fra il 2013 e il 2014.