

Statistica Applicata

Corso di Laurea in Informatica
Università Ca' Foscari Venezia

2014-2015

cristiano.varin@unive.it

Predictive Analytics

Corso di Laurea in Informatica
Università Ca' Foscari Venezia

2014-2015

cristiano.varin@unive.it

Organizzazione

Orario

lezione martedì 15:45-17:15 aula A

laboratorio con **R** giovedì 15:45-17:15 laboratorio 3

Ricevimento (a partire dalla prossima settimana)

martedì 14:30-15:30 ufficio ospiti

Contatti cristiano.varin@unive.it

This copy is for your personal, noncommercial use only. You can order presentation-ready copies for distribution to your colleagues, clients or customers [here](#) or use the "Reprints" tool that appears next to any article. Visit [www.nytreprints.com](#) for samples and additional information. Order a reprint of this article now.



February 11, 2012

The Age of Big Data

By STEVE LOHR

GOOD with numbers? Fascinated by data? The sound you hear is opportunity knocking.

Welcome to the Age of Big Data. The new megarich of Silicon Valley, first at Google and now Facebook, are masters at harnessing the data of the Web — online searches, posts and messages — with Internet advertising. At the World Economic Forum last month in Davos, Switzerland, Big Data was a marquee topic. A report by the forum, “**Big Data, Big Impact**,” declared data a new class of economic asset, like currency or gold.

“McKinsey Global Institute projected that the United States needs 140,000 to 190,000 more workers with “deep analytical” expertise and 1.5 million more data-literate managers, whether retrained or hired.”

Google's prediction: What will be the "sexy" job in the next ten years?



Here's a strange prediction from Google's Chief Economist: "I keep saying that the sexy job in the next 10 years will be statisticians. And I'm not kidding."

That quote came from a [New York Times article](#) about the rapidly increasing demand for statisticians.



SAPVoice

FOLLOW

RUN BETTER
full bio →

Comment Now

+ Follow Comments

ForbesBrandVoice Connecting marketers to the Forbes audience. [What is this?](#)

BUSINESS | 1/21/2014 @ 8:29AM | 6,835 views

Data Scientist: Sexiest Job Of The Century?

SAP Guest, SAP

[+ Comment Now](#) [+ Follow Comments](#)

By Judith Magyar, Strategic Communications Director, SAP Platform Solutions



Photo: Shutterstock



S&OP
can be
simple and
outstandin
profitable

Find out how

If there's one thing that gives a job an indefinable allure, it's that no one really knows exactly what you do, or how you do it; they just know you do something important that generates impressive results.

A recent article in the [Harvard Business Review](#), for instance, claims the data scientist is the sexiest job of the century. It's not a job you'll find advertised, nor can you get a degree in it, but as you might have guessed, [data science](#) is all about transforming data into business value using math and algorithms.

▲ CONFERENCES AND MORE

What is a data scientist

About data scientists

Rising alongside the relatively new technology of [big data](#) is the new job title data scientist. While not tied exclusively to [big data](#) projects, the data scientist role does complement them because of the increased breadth and depth of data being examined, as compared to traditional roles.



Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data

→ [Download the ebook](#)

So what does a data scientist do?

A data scientist represents an evolution from the business or data analyst role. The formal training is similar, with a solid foundation typically in computer science and applications, modeling, statistics, analytics and math. What sets the data scientist apart is strong business acumen, coupled with the ability to communicate findings to both business and IT leaders in a way that can influence how an organization approaches a business challenge. Good data scientists will not just address business problems, they will pick the right problems that have the most value to the organization.

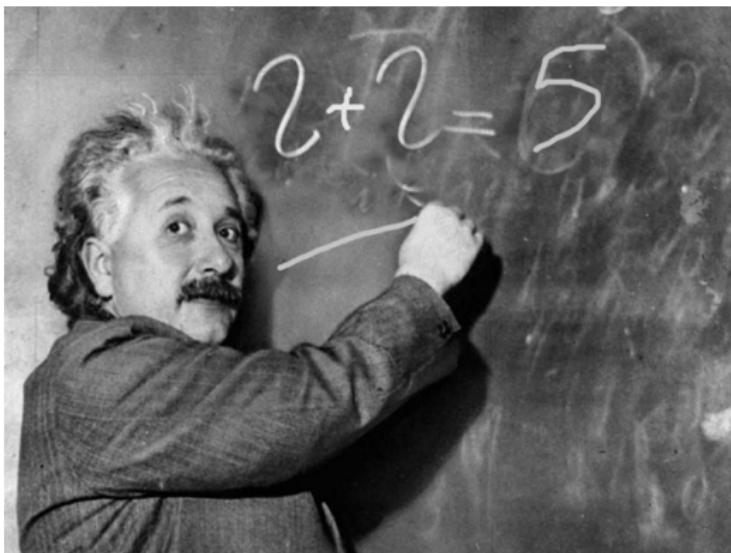
Un data scientist al lavoro...



1. Mathematician / \$101,360
2. Tenured University Professor / \$68,970
3. Statistician / \$75,560
4. Actuary / \$93,680
5. Audiologist / \$69,720
6. Dental Hygienist / \$70,210
7. Software Engineer / \$93,350
8. Computer Systems Analyst / \$79,680
9. Occupational Therapist / \$75,400
10. Speech Pathologist / \$69,870



Statistics + Computer Science = Data Science



Applied Statistics

- Health care
- Epidemiology
- Biology
- Ecology
- Chemistry
- Agriculture
- Engineering
- Technology
- Government
- Demography
- Economics
- Finance
- Business
- ...

Business Analytics



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox

Print/export

Languages
Deutsch

Create account Log in

Article [Talk](#)

Read [Edit](#) [View history](#)

Search

Business analytics

From Wikipedia, the free encyclopedia



This article **needs additional citations for verification**. Please help [improve this article](#) by adding citations to reliable sources. Unsourced material may be challenged and removed. (October 2010)

Not to be confused with [Business analysis](#).

Business analytics (BA) refers to the skills, technologies, applications and practices for continuous iterative exploration and investigation of past business performance to gain insight and drive business planning.^[1] Business analytics focuses on developing new insights and understanding of business performance based on [data](#) and [statistical methods](#). In contrast, [business intelligence](#) traditionally focuses on using a consistent set of metrics to both measure past performance and guide business planning, which is also based on data and statistical methods.

Business [analytics](#) makes extensive use of data, [statistical](#) and [quantitative](#) analysis, explanatory and [predictive modeling](#),^[2] and fact-based management to drive [decision making](#). Analytics may be used as input for human decisions or may drive fully automated decisions. Business intelligence is [querying](#), [reporting](#), [OLAP](#), and "alerts".

In other words, querying, reporting, OLAP, and alert tools can answer questions such as what happened, how many, how often, where the problem is, and what actions are needed. Business analytics can answer questions like why is this happening, what if these trends continue, what will happen next (that is, predict), what is the best that can happen (that is, optimize).^[3]

IBM Business Analytics

[Soluzioni](#)[Servizi](#)[Prodotti](#)[Supporto & Download](#)[La mia IBM](#)[cerca](#)

Software IBM > Prodotti >

Business Analytics

Ottimizzazione dei processi decisionali, per migliorare i risultati e gestire i rischi.

[Cos'è](#)[Prodotti](#)

Performance superiori con strumenti di Analytics e Business Intelligence

La Business Analytics consente di trasformare i dati in informazioni utili a definire i trend, monitorare l'andamento del mercato, controllare i costi, identificare in anticipo i rischi da evitare e le opportunità da cogliere. Aiuta così il management a prendere decisioni più intelligenti in modo più rapido e tutta l'azienda a diventare più reattiva. Oltre a promuovere la crescita e l'efficienza, la Business Analytics permette, ad esempio, di effettuare previsioni più efficaci e di modificare in tempo reale una pianificazione.

I risultati sono dimostrati: recenti studi hanno evidenziato che le organizzazioni che conoscono e utilizzano abitualmente questi strumenti ottengono performance tre volte migliori rispetto alle altre.

Il software IBM di Business Analytics offre soluzioni esclusive per applicare le analisi ai processi decisionali, in qualsiasi momento e a prescindere dalla propria localizzazione. E' possibile:

- accedere a qualsiasi tipo e fonte di informazione, da desktop come da dispositivi mobile;

Siamo qui per aiutarti

 [Email](#) [Richiedi un preventivo](#) [Telefona: +39 0259629449
Priorità codice: 101KR29W](#)

Valuta il tuo Analytics Quotient

 [Inizia il test \(US\)](#)

Qual è il tuo AQ? Quanto sei in grado di adattare strategie e processi alle indicazioni che emergono dai dati? Con questo breve questionario online, scoprirai come si posiziona la tua azienda rispetto a quelle che ottengono il massimo dagli analytics.

Google Analytics

Sito web ufficiale di Google Analytics – Analisi dei dati web e rapporti - Google Analytics

www.google.com/analytics/ Reader

Dopo il jail...tua Apple TV Pagamenti e I...ore (Italia) Apple www.bering...gorithm.pdf Using R | CA...sis | US EPA Import to Mendeley Associate e... RSSeNews > +

Google Analytics

Cerca nel sito

Home Funzioni Impara Partner Guida Accedi o Crea un account

Analisi dei dati web di livello aziendale

Fornita sulla piattaforma di livello mondiale di Google. Ulteriori informazioni

Home page Rapporti standard Rapporti personalizzati

Dashboard personale

Visite giornaliere

Tipi di traffico

Paese/zona	Visite	Tempo medio sul sito
Stati Uniti	67.445	00:01:54
Regno Unito	18.948	00:01:37
India	8.882	00:00:58
Canada	6.371	00:01:02
Germania	5.845	00:00:32
Francia	5.243	00:00:38

14 / 45

Crea un'utenza Entrà



WIKIPEDIA
L'enciclopedia libera

Voce Discussione

Leggi Modifica Visualizza cronologia

Ricerca



Google Analytics

Da Wikipedia, l'enciclopedia libera.



Questa voce sull'argomento informatica è solo un abbozzo.

Contribuisci a migliorarla secondo le convenzioni di Wikipedia. Segui i suggerimenti del progetto di riferimento.

Google Analytics è un servizio gratuito di Google che consente di analizzare delle dettagliate statistiche sui visitatori di un sito web. Il sito si rivolge principalmente al marketing di internet e ai webmaster. Google Analytics è il servizio di statistiche più usato nel web^[1], attualmente in uso presso circa il 57% dei 10.000 siti web più popolari. Un'altra analisi delle quote di mercato afferma che Google Analytics è utilizzato da circa il 49,95% del primo milione di siti web (per come vengono attualmente classificate da Alexa).

Sul sito è possibile monitorare i visitatori provenienti da tutte le fonti, siano essi motori di ricerca, siti referer o campagne a pagamento, il Display Advertising, le reti di Pay per click e l'email marketing, ma anche i collegamenti all'interno dei documenti PDF. Integrato con AdWords, gli utenti possono analizzare le campagne online, monitorando la qualità delle pagine di destinazione e le conversioni (obiettivi). Gli obiettivi possono includere le vendite, la lead generation, la visualizzazione di una specifica pagina o il download di un particolare file e possono anche essere monetizzati. Utilizzando Google Analytics, il marketing è in grado di determinare quale sia il rendimento degli annunci, fornendo le informazioni per ottimizzare le campagne o abbattere i costi. L'approccio del servizio è quello di mostrare ad alto livello vari tipi di dashboard, in grado di soddisfare velocemente l'utente occasionale, ma anche report più approfonditi che rispondono alle esigenze degli utenti più esperti e dei responsabili di Marketing. Attraverso l'utilizzo di GA, è possibile individuare quali siano le pagine più visualizzate dai visitatori di un sito, la loro provenienza (referrer), per quanto tempo sono rimasti all'interno del sito e la loro posizione geografica. Gli utenti possono impostare fino a 50 profili. Ogni profilo può corrispondere ad un sito web, ma anche ad una particolare sezione o sottosezione del sito. Il servizio di tracciamento dei dati è gratuito per i siti che si limitano ad un traffico di meno di 5 milioni di pagine visualizzate al mese (circa 2 pagine visualizzate al secondo); questo limite viene sfondato nel caso in cui l'account Google Analytics venga associato ad un account AdWords (e quindi alla realizzazione di campagne a pagamento).

Google Analytics

Google Analytics

URL <http://www.google.com/analytics/>

Tipo di sito Statistiche di siti web

Registrazione Sì

Proprietario Google

Creato da Google

Stato Attivo

corrente del
sito

« □ □ »

Predictive analytics alla Target Corporation

The New York Times

STOKER
COMING SOON

February 16, 2012

How Companies Learn Your Secrets

By CHARLES DUHIGG

Andrew Pole had just started working as a statistician for Target in 2002, when two colleagues from the marketing department stopped by his desk to ask an odd question: "If we wanted to figure out if a customer is pregnant, even if she didn't want us to know, can you do that?"

Pole has a master's degree in statistics and another in economics, and has been obsessed with the intersection of data and human behavior most of his life. His parents were teachers in North Dakota, and while other kids were going to 4-H, Pole was doing algebra and writing computer programs. "The stereotype of a math nerd is true," he told me when I spoke with him last year. "I kind of like going out and evangelizing analytics."

As the marketers explained to Pole — and as Pole later explained to me, back when we were still speaking and before Target told him to stop — new parents are a retailer's holy grail. Most shoppers don't buy everything they need at one store. Instead, they buy groceries at the grocery store and toys at the toy store, and they visit Target only when they need certain items they associate with Target — cleaning supplies, say, or new socks or a six-month supply of toilet paper. But Target sells everything from milk to stuffed animals to lawn furniture to electronics, so one of the company's primary goals is convincing customers that the only store they need is Target. But it's a tough message to get across, even with the most ingenious ad campaigns, because once consumers' shopping habits are ingrained, it's incredibly difficult to change them.

There are, however, some brief periods in a person's life when old routines fall apart and buying habits are suddenly in flux. One of those moments — *the* moment, really — is right around the birth of a child, when parents are exhausted and overwhelmed and their shopping patterns and brand loyalties are up for grabs. But as Target's marketers explained to Pole, timing is everything. Because birth records are usually public, the moment a couple have a new baby, they are almost instantaneously barraged with offers and incentives and advertisements from all sorts of companies. Which means that the key is to reach them earlier, before any other retailers know a baby is on the way. Specifically, the marketers said they wanted to send specially designed ads to women in their second trimester, which is when most expectant mothers begin buying all sorts of new things, like prenatal vitamins and maternity clothing. "Can you give us a list?" the marketers asked.

The Colbert Report Video:



Il progetto R



[About R](#)
[What is R?](#)
[Contributors](#)
[Screenshots](#)
[What's new?](#)

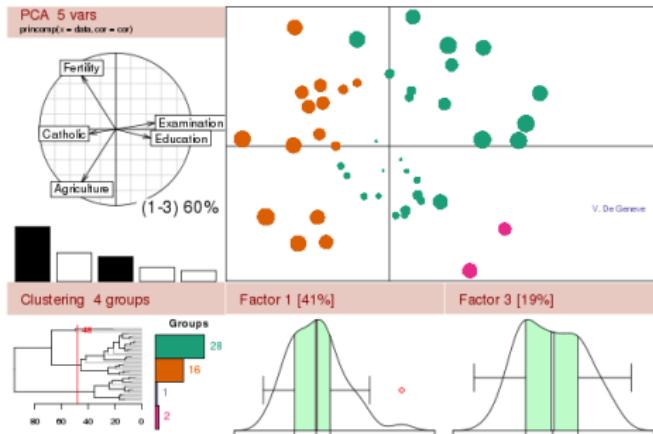
[Download, Packages](#)
[CRAN](#)

[R Project](#)
[Foundation](#)
[Members & Donors](#)
[Mailing Lists](#)
[Bug Tracking](#)
[Developer Page](#)
[Conferences](#)
[Search](#)

[Documentation](#)
[Manuals](#)
[FAQs](#)
[The R Journal](#)
[Wiki](#)
[Books](#)
[Certification](#)
[Other](#)

[Misc](#)
[Bioconductor](#)
[Related Projects](#)
[User Groups](#)
[Links](#)

The R Project for Statistical Computing



Getting Started:

- R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).
- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

News:

- R version 2.15.2 (Trick or Treat) has been released on 2012-10-26.
- [The R Journal Vol.4/1](#) is available.
- [useR! 2012](#), took place at Vanderbilt University, Nashville Tennessee, USA, June 12-15, 2012.
- [useR! 2013](#), will take place at the University of Castilla-La Mancha, Albacete, Spain, July 10-12 2013..

January 6, 2009

Data Analysts Captivated by R's Power

By ASHLEE VANCE

To some people R is just the 18th letter of the alphabet. To others, it's the rating on racy movies, a measure of an attic's insulation or what pirates in movies say.

R is also the name of a popular programming language used by a growing number of data analysts inside corporations and academia. It is becoming their lingua franca partly because data mining has entered a golden age, whether being used to set ad prices, find new drugs more quickly or fine-tune financial models. Companies as diverse as Google, Pfizer, Merck, Bank of America, the InterContinental Hotels Group and Shell use it.

But R has also quickly found a following because statisticians, engineers and scientists without computer programming skills find it easy to use.

"R is really important to the point that it's hard to overvalue it," said Daryl Pregibon, a research scientist at Google, which uses the software widely. "It allows statisticians to do very intricate and complicated analyses without knowing the blood and guts of computing systems."

It is also free. R is an open-source program, and its popularity reflects a shift in the type of software used inside corporations. Open-source software is free for anyone to use and modify. I.B.M., Hewlett-Packard and Dell make billions of dollars a year selling servers that run the open-source Linux operating system, which competes with Windows from Microsoft. Most Web sites are displayed using an open-source application called Apache, and companies increasingly rely on the open-source MySQL database to store their critical information. Many people view the end results of all this technology via the Firefox Web browser, also open-source software.

R is similar to other programming languages, like C, Java and Perl, in that it helps people perform a wide variety of computing tasks by giving them access to various commands. For statisticians, however, R is particularly useful because it contains a number of built-in mechanisms for organizing data, running calculations on the information and creating graphical representations of data sets.

- To some people R is just the 18th letter of the alphabet. To others, it's also the name of a popular **programming language** used by a growing number of **data analysts** inside corporations and academia.
- Companies as diverse as Google, Pfizer, Merck, Bank of America, the InterContinental Hotels Group and Shell use it.
- R contains a number of built-in mechanisms for organizing data, running calculations on the information and creating graphical representations of data sets.
- “The great beauty of R is that you can modify it to do all sorts of things,” said Hal Varian, chief economist at Google. “And you have a lot of prepackaged stuff that’s already available, so you’re standing on the shoulders of giants.”



Left, Stuart Isett for The New York Times; right, Kieran Scott for The New York Times

R first appeared in 1996, when the statistics professors Robert Gentleman, left, and Ross Ihaka released the code as a free software package.

Data analytics con R?



[Read the Revolutions Blog](#) | [Follow us on Twitter](#) | [Get our Newsletter](#)

[Buy Now](#)

[Google Custom Search](#)

[Search](#)

[What is R?](#)

[Products](#)

[Services](#)

[R Downloads](#)

[Why Revolution R?](#)

[Support](#)

[News & Events](#)

[About Us](#)

High Performance R Analytics for the Enterprise

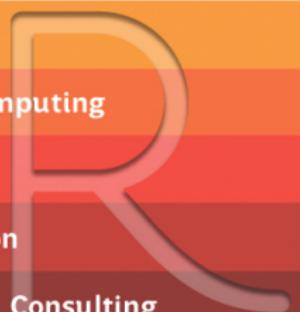
→ Big Data Analytics

→ High Performance Computing

→ Open Source R

→ Analytics in Production

→ R Services: Training & Consulting



WEBINARS

Using R with Hadoop
Jan 24th, 2013

Introduction to R for Data Mining
Feb 14, 2013

Revolution R Enterprise - 100% R and More
Mar 14, 2013

LATEST NEWS & EVENTS

Revolution Analytics Brings Big Data Decision Trees and New Hadoop Support to Predictive Analytics

UpStream Software's Big Data Analytics Platform for Marketing Optimization Helps Clients Understand Buying Behavior and Improve Customer Targeting

Revolution R Enterprise Boosts Big Data Analytics Capabilities

[Download
Revolution R
FREE](#)

[Free Enterprise
Software for
Academics](#)

[Revolution R
Enterprise for
Production-
Grade Analytics](#)

[Visit our R
Community Site
inside-R.org](#)

Google e l'influenza

L'influenza...

- Ogni anno causa da 250 000 a 500 000 vittime nel mondo (!)
- Oltre alla “solita” influenza stagionale, negli ultimi anni si sono osservate nuove forme influenzali a cui l'uomo non è stato esposto precedentemente.
- Queste nuove forme influenzali sono considerate molto pericolose quando fanno il “salto di qualità” che permette il passaggio da uomo a uomo.
- Rischio di drammatiche pandemie con milioni di morti...

google.org Trend influenzali

Lingua: ▼

[Home page di Google.org
\(inglese\)](#)

[Diffusione della dengue](#)

[Trend influenzali](#)

[Home page](#)

[Selezione Paese/regione](#) ▼

[Come funziona?](#)

[Domande frequenti
\(FAQ\)](#)

Attività influenzale

Intensa

Alta

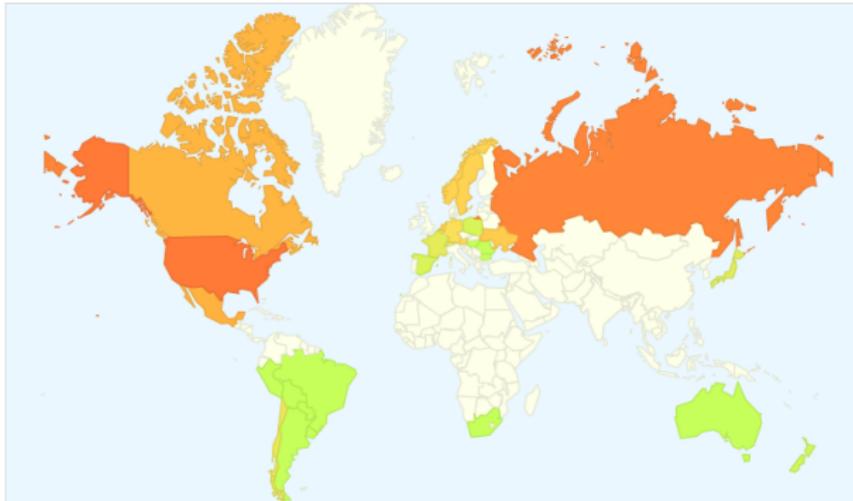
Moderata

Bassa

Minima

Scopri i trend influenzali nel mondo

Abbiamo scoperto che determinati termini di ricerca sono validi indicatori dell'attività influenzale. Google Trend influenzali utilizza dati di ricerca aggregati di Google per stimare l'attività influenzale. [Ulteriori informazioni »](#)



[Scarica i dati sulle attività influenzali nel mondo](#)

Google flu trends

Come funziona?

Abbiamo scoperto che determinati termini di ricerca sono validi indicatori dell'attività influenzale. Google Trend influenzali utilizza dati di ricerca aggregati di Google per stimare l'attività influenzale attuale nel mondo quasi in tempo reale.

Ogni settimana, milioni di utenti nel mondo cercano informazioni sanitarie online. Com'era prevedibile, vengono effettuate più ricerche relative all'influenza durante la stagione influenzale, più ricerche relative alle allergie durante la stagione delle allergie e più ricerche relative alle scottature solari in estate. Puoi esaminare tutti questi fenomeni utilizzando [Google Statistiche di ricerca](#). Ma è possibile partire dalle query di ricerca per arrivare a delineare un modello attendibile dei fenomeni che si verificano nella realtà?

Abbiamo scoperto che esiste una stretta correlazione tra il numero di persone che cercano argomenti relativi all'influenza e il numero di persone che manifestano effettivamente sintomi influenzali. Naturalmente non tutti gli utenti che effettuano ricerche digitando la parola "influenza" sono realmente malati, ma quando tutte le query di ricerca relative all'influenza vengono riunite, emerge un modello. Abbiamo confrontato i nostri conteggi delle query con i sistemi di sorveglianza tradizionali dei dati influenzali e abbiamo rilevato che molte query di ricerca tendono ad acquisire popolarità nel periodo esatto in cui si manifesta la stagione influenzale. Calcolando la frequenza di queste query di ricerca, possiamo stimare il livello di diffusione dell'influenza in vari Paesi e regioni del mondo. I nostri [risultati](#) sono stati [pubblicati](#) sulla rivista Nature.

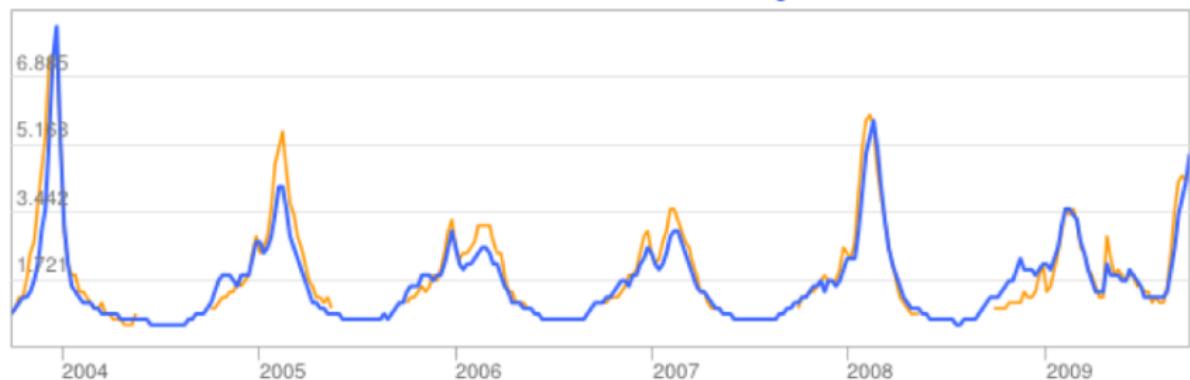
Stime storiche

Visualizza dati per: Stati Uniti

Attività influenzale Stati Uniti

Stima sull'influenza

● Stima di Google Trend influenzali ● Dati Stati Uniti



Stati Uniti: dati ILI (Influenza-Like Illness) forniti pubblicamente dagli [U.S. Centers for Disease Control](#).



Journal home > Archive > Letter > Full Text

Search go Advanced search

Journal content

- [+ Journal home](#)
- [+ Advance online publication](#)
- [+ Current issue](#)
- [+ Nature News](#)
- [+ Archive](#)
- [+ Supplements](#)
- [+ Web focuses](#)
- [+ Podcasts](#)
- [+ Videos](#)
- [+ News Specials](#)

Journal information

- [+ About the journal](#)
- [+ For authors](#)
- [+ Online submission](#)
- [+ Nature Awards](#)

Letter

Nature 457, 1012-1014 (19 February 2009) | doi:10.1038/nature07634; Received 14 August 2008;
Accepted 13 November 2008; Published online 19 November 2008; [Corrected](#) 19 February 2009

Detecting influenza epidemics using search engine query data

Jeremy Ginsberg¹, Matthew H. Mohebbi¹, Rajan S. Patel¹, Lynnette Brammer²,
Mark S. Smolinski¹ & Larry Brilliant¹

1. Google Inc., 1600 Amphitheatre Parkway, Mountain View, California 94043, USA
2. Centers for Disease Control and Prevention, 1600 Clifton Road, NE, Atlanta, Georgia 30333, USA

Correspondence to: Matthew H. Mohebbi¹. Correspondence and requests for materials should be addressed to J.G. or M.H.M. (Email: flutrends-support@google.com).

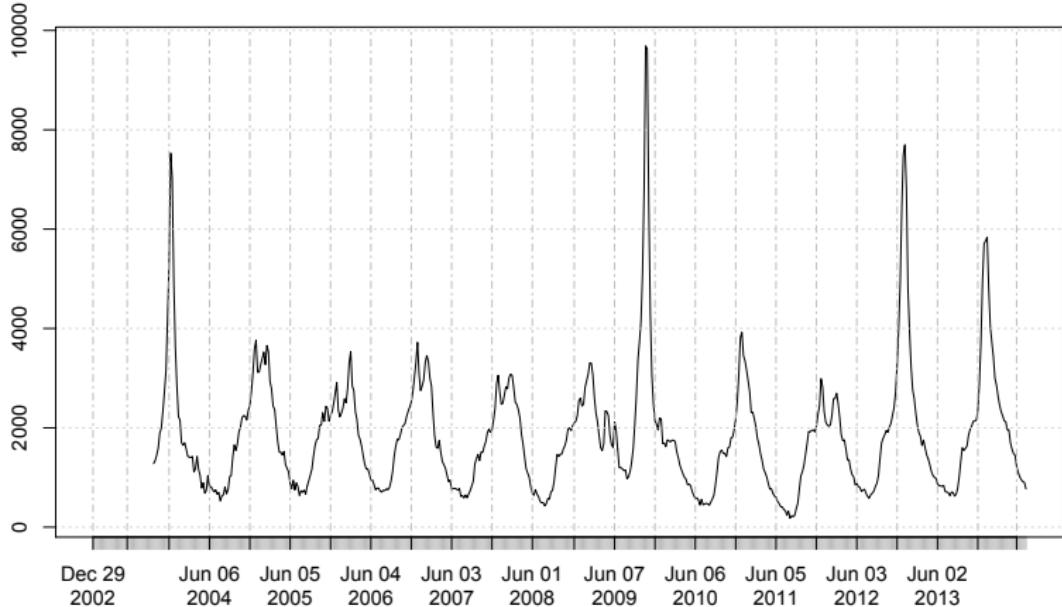
Seasonal influenza epidemics are a major public health concern, causing tens of millions of respiratory illnesses and 250,000 to 500,000 deaths worldwide each year¹. In addition to seasonal influenza, a new strain of influenza virus against which no previous immunity exists and that demonstrates human-to-human transmission could result in a pandemic with millions of fatalities². Early detection of disease activity, when followed by a rapid response, can reduce the impact of both seasonal and pandemic influenza^{3,4}. One way to improve early detection is to monitor health-seeking behaviour in the form of queries to online search engines, which are submitted by millions of users around the world each day. Here we present a method of analysing large numbers of Google search queries to track influenza-like illness in a population.

FULL TEXT

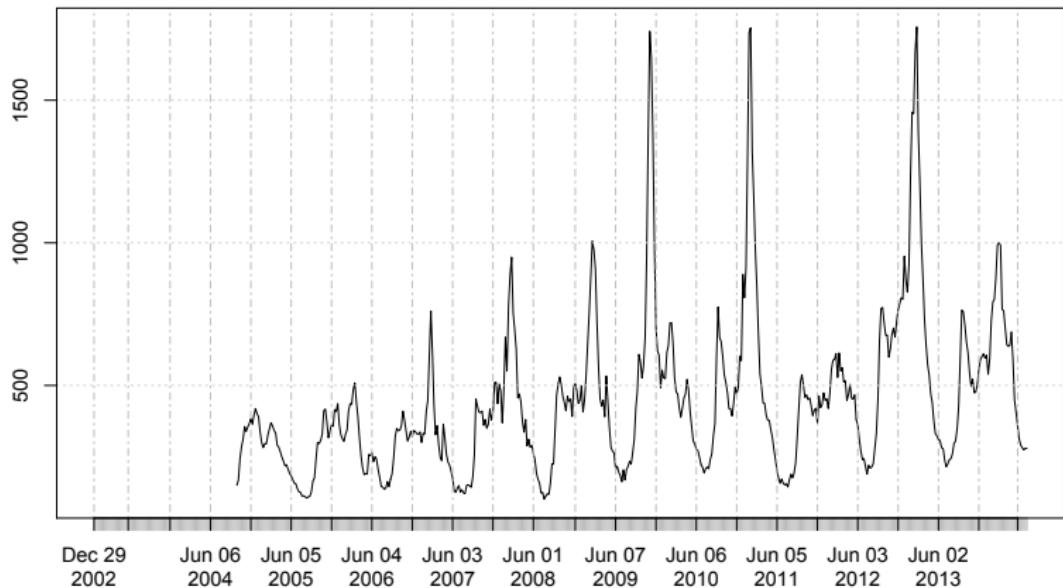
- [+ Previous | Next +](#)
- [+ Table of contents](#)
- [Download PDF](#)
- [Send to a friend](#)
- [CrossRef lists 157 articles citing this article](#)
- [Scopus lists 285 articles citing this article](#)
- [Export citation](#)
- [Export references](#)
- [Rights and permissions](#)
- [Order commercial reprints](#)
- [Bookmark in Connotea](#)

subscribe to
nature

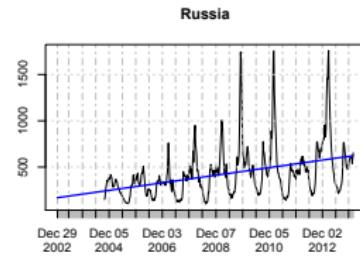
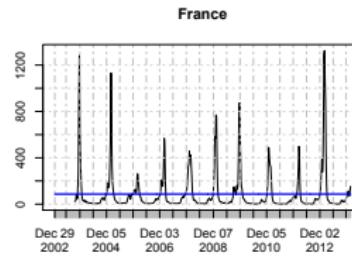
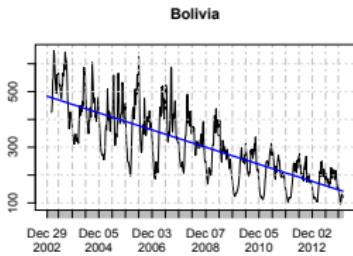
Canada



Russia



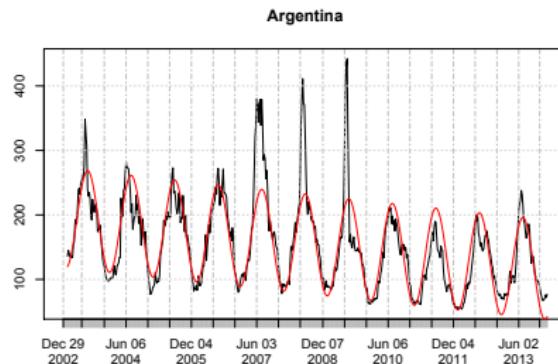
Trend



Trend lineare decrescente, costante, crescente:

$$\text{trend} = \beta_0 + \beta_1 \text{tempo}$$

Stagionalità



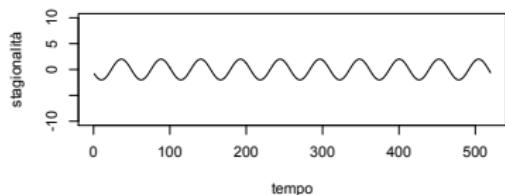
La stagionalità può essere espressa tramite funzioni trigonometriche

$$\text{stagionalità} = A \cos(2\pi\omega \text{tempo} + \phi)$$

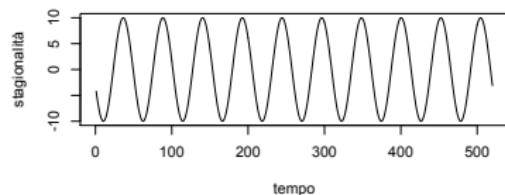
dove:

- A è l'**ampiezza**, cioè l'altezza della componente stagionale
- ω è la **frequenza di oscillazione**
- ϕ è lo **spostamento di fase**

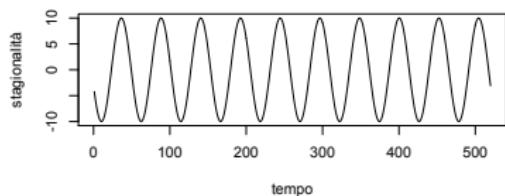
$A = 2$



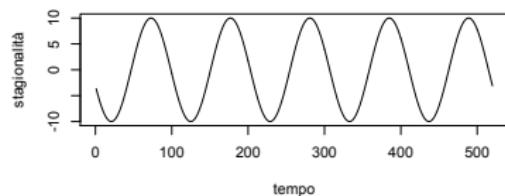
$A = 10$



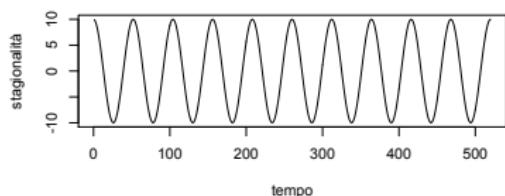
$\omega = 1/52$



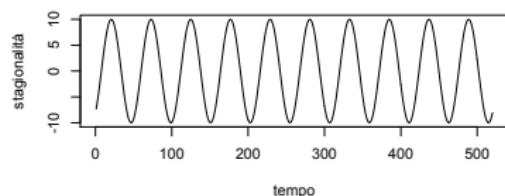
$\omega = 1/104$



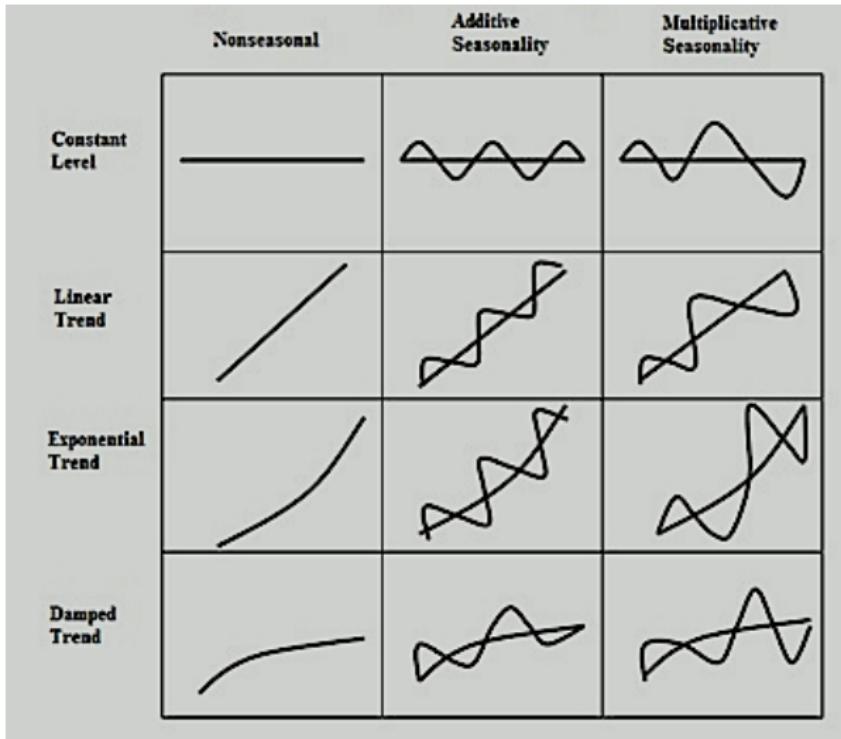
$\phi = 0$



$\phi = 1.2\pi$



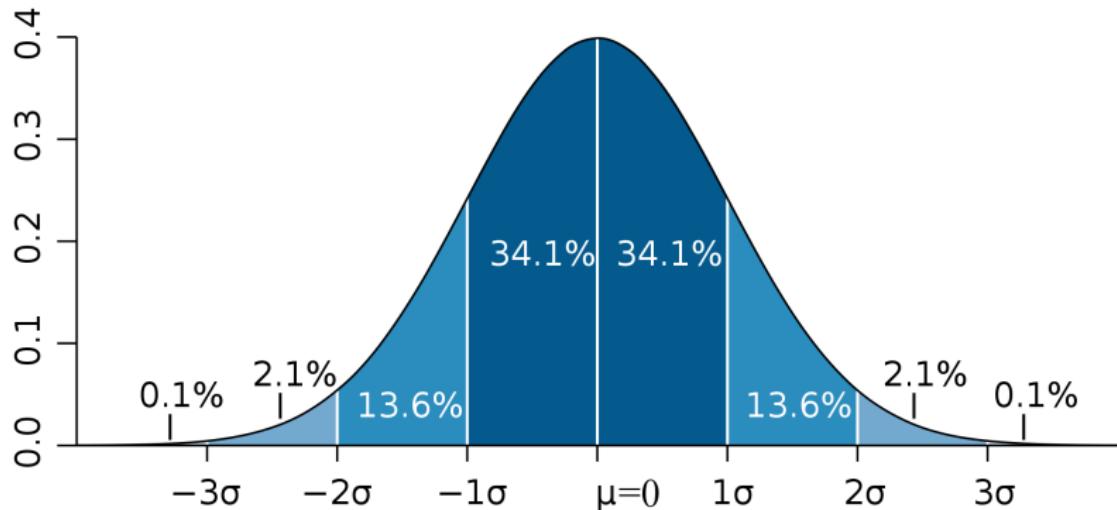
Trend e stagionalità

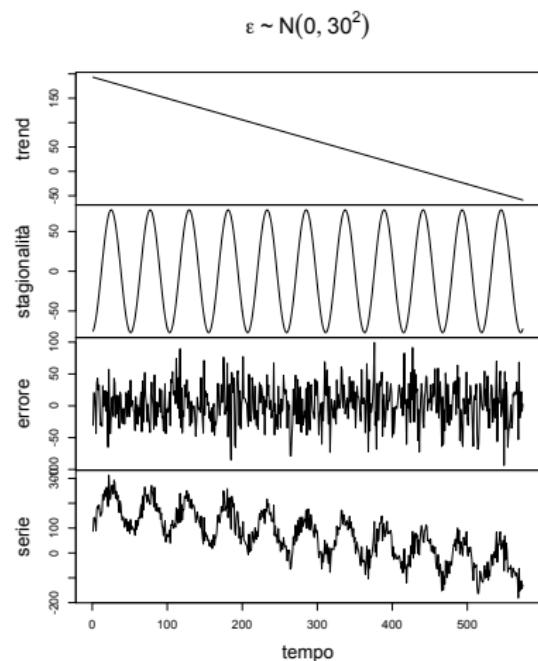
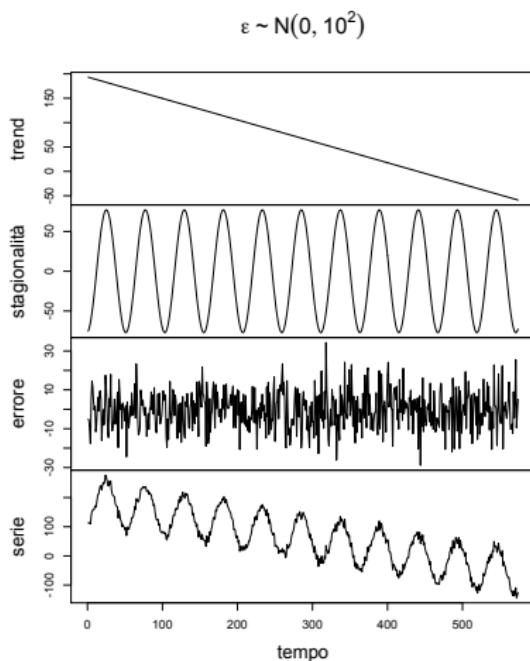


Errori “normali”

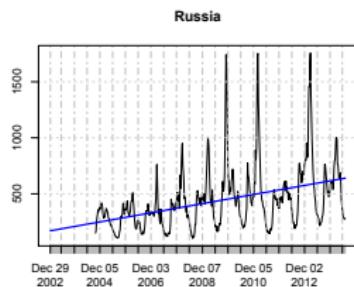
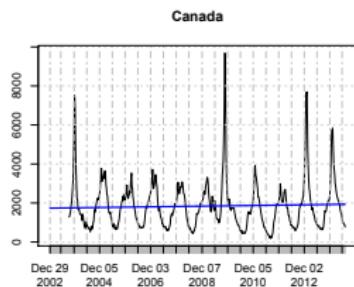
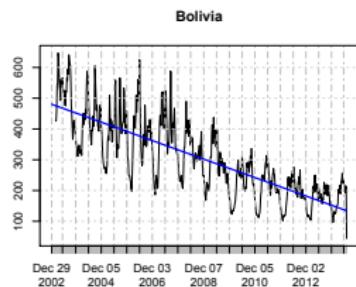
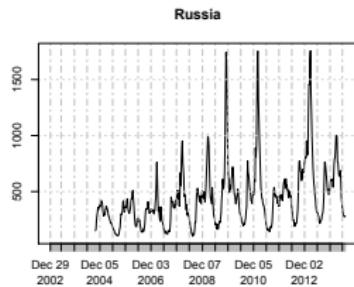
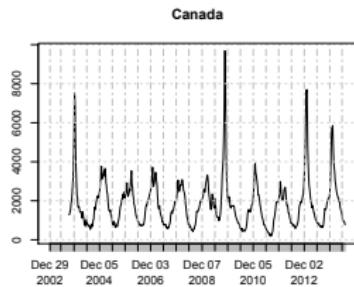
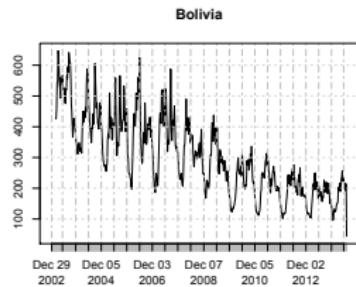
Mettendo assieme tutti i pezzi

$$ILI = \text{trend} + \text{stagionalità} + \text{errore}$$

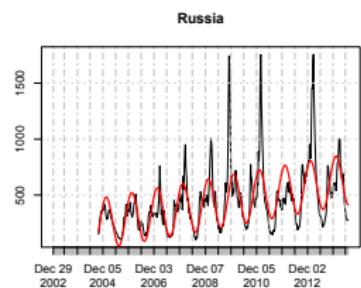
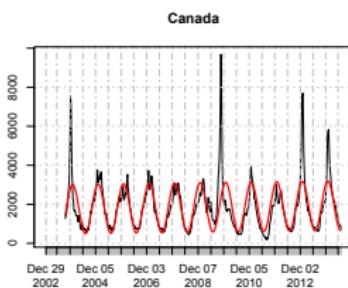
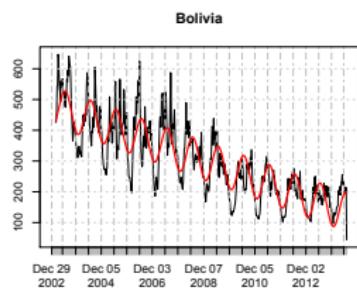
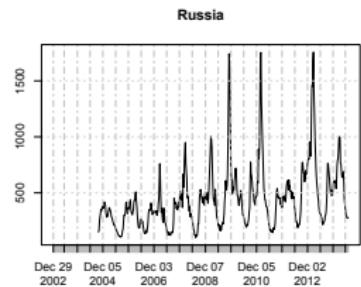
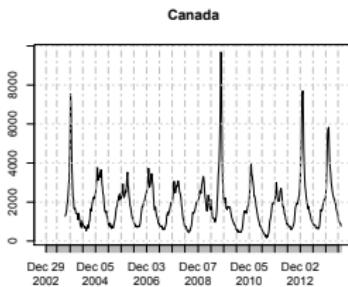
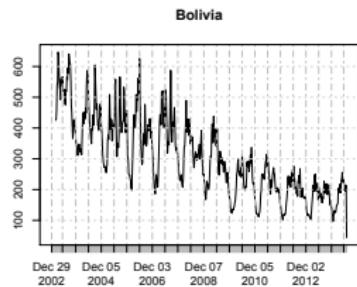




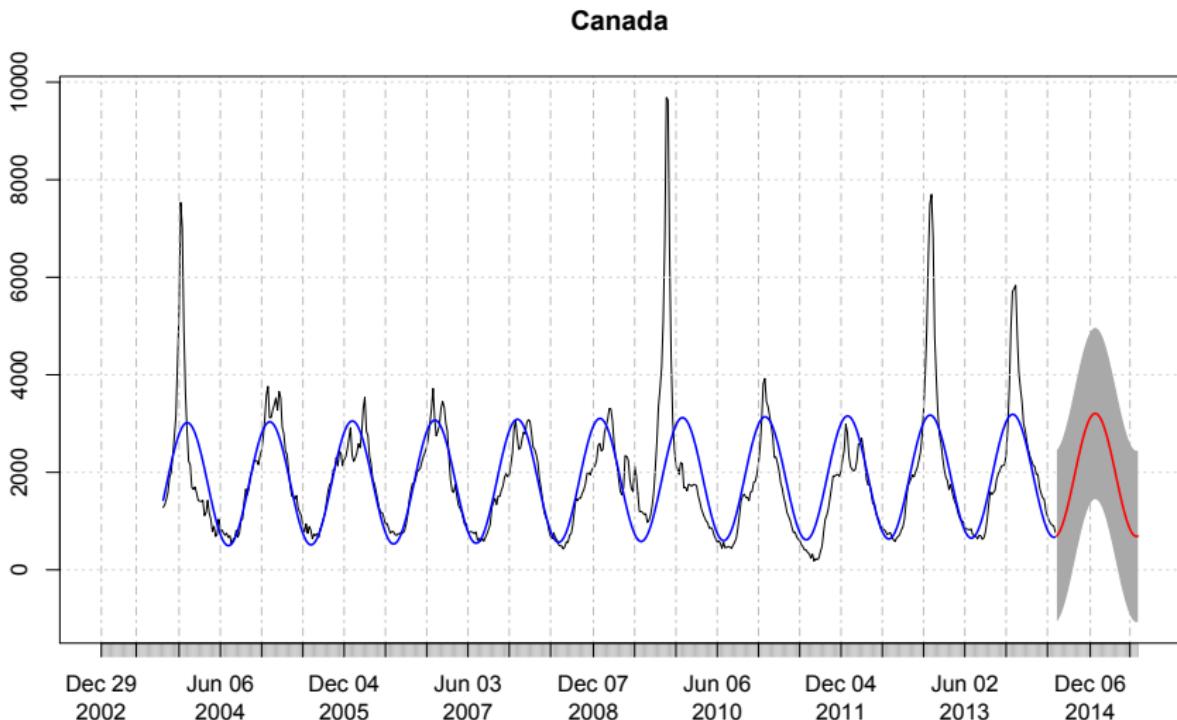
Trend



Stagionalità



Previsioni





WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact Wikipedia

Toolbox
Print/export
Languages
Español
Français
Suomi
中文

Article Talk

Read Edit View history

Search



Create account Log in

2009 flu pandemic in Canada

From Wikipedia, the free encyclopedia

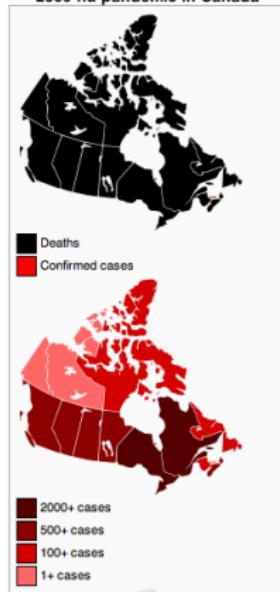
Main articles: 2009 flu pandemic and 2009 flu pandemic in North America

The **2009 flu pandemic in Canada** is part of an epidemic in 2009 of a new strain of **influenza A virus subtype H1N1** causing what has been commonly called **swine flu**. In Canada, roughly 10% of the populace (or 3.5 million) has been infected with the virus.^{[2][3]} with 426 confirmed deaths (as of 20 February 2010);^[1] non-fatal individual cases are for the most part no longer being recorded. About 40% of Canadians have been immunized against H1N1 since a national vaccination campaign began in October,^[4] with Canada among the countries in the world leading in the percentage of the population that has been vaccinated.^{[5][6]} The widespread effect of H1N1 in Canada has raised concerns during the months leading to the **XXI Olympic Winter Games**, which took place in Vancouver on February 2010.^[6]

Contents [hide]

- 1 Human cases
 - 1.1 Alberta
 - 1.2 British Columbia
 - 1.3 Manitoba
 - 1.4 Newfoundland and Labrador
 - 1.5 New Brunswick
 - 1.6 Northwest Territories
 - 1.7 Nova Scotia
 - 1.8 Nunavut
 - 1.9 Ontario
 - 1.10 Prince Edward Island
 - 1.11 Quebec
 - 1.12 Saskatchewan
 - 1.13 Yukon
 - 1.14 Totals
 - 1.15 Progression Chart

2009 flu pandemic in Canada



Big data: non sempre funzionano...

NEWS IN FOCUS

GENETIC DIAGNOSIS Data barriers hamper search for meaning in mutations p.156

FUNDING US science agencies gird themselves for the budget axe p.158

MALARIA Plant source of key drug faces lab-made competition p.160

BIOMEDICINE A Texas-style showdown over stem-cell therapy p.166



JOHN ANGELILLO/UNIVERSITY OF TORONTO



The latest US influenza season is more severe and has caused more deaths than usual.

EPIDEMIOLOGY

When Google got flu wrong

US outbreak foxes a leading web-based method for tracking seasonal flu.

virulent of the three main seasonal flu strains.

Traditional flu monitoring depends in part on national networks of physicians who report cases of patients with influenza-like illness (ILI) — a diffuse set of symptoms, including high fever, that is used as a proxy for flu. That estimate is then refined by testing a subset of people with these symptoms to determine how many have flu and not some other infection.

With its creation of the Sentinelles network in 1984, France was the first country to computerize its surveillance. Many countries have since developed similar networks — the US system, overseen by the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, includes some 2,700 health-care centres that record about 30 million patient visits annually.

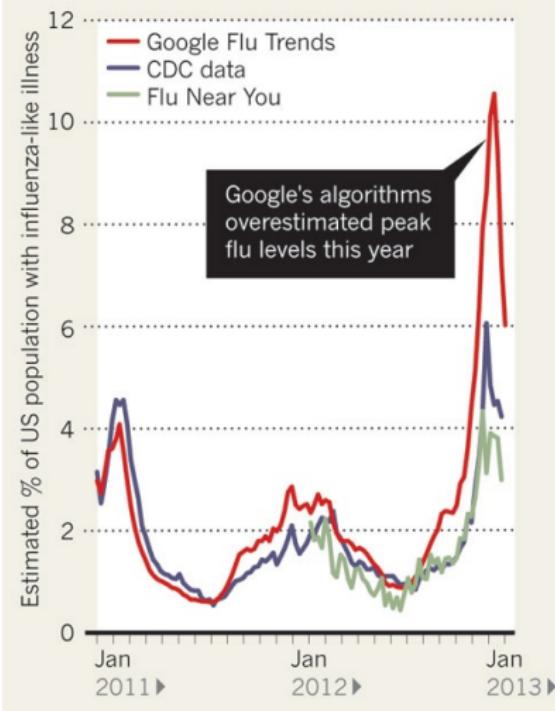
But the near-global coverage of the Internet and burgeoning social-media platforms such as Twitter have raised hopes that these technologies could open the way to easier, faster estimates of ILI, spanning larger populations.

The mother of these new systems is Google's, launched in 2008. Based on research by Google and the CDC, it relies on data mining records of flu-related search terms entered in Google's search engine, combined with computer modelling. Its estimates have almost exactly matched the CDC's own surveillance data over time — and it delivers them several days faster than the CDC can. The system has since been rolled out to 29 countries worldwide, and has been extended to include surveillance for a second disease, dengue.

Google Flu Trends has continued to perform remarkably well, and researchers in many countries have confirmed that its ILI estimates

FEVER PEAKS

A comparison of three different methods of measuring the proportion of the US population with an influenza-like illness.



Questo corso...

- È un'introduzione a metodi di **data analytics** e più in particolare di **predictive analytics**
- È centrato sull'intuizione e sulle idee piuttosto che sui dettagli tecnici...
- ... che comunque non vengono trascurati!
- Ha un taglio pratico: casi studio (dati reali) vengono utilizzati per illustrare e insegnare i metodi

L'esame

- Ad ogni studente viene assegnato un caso studio e i relativi dati
- Viene chiesto di analizzare i dati (con quello che si è imparato in questo corso, in altri corsi, con la propria fantasia, intuizione, . . .)
- Due fasi:
 - ▶ Breve report (max. 15 pagine) con la descrizione e l'interpretazione delle analisi svolte
 - ▶ Presentazione delle analisi svolte
- Domande sulla presentazione e sul report
- Sintesi e capacità analitiche vengono premiate!

Materiali didattici

- Appunti delle lezioni!
- Materiale distribuito durante il corso:
 - Esercitazioni svolte in laboratorio
 - Letture integrative
- Per chi (*non*) segue il corso, si consiglia

