Workshop on HRMT – Modernizing statistics: how to get there?

Geneva, 15-17 October 2014

A new job for statisticians: the data scientist. Which skills, how to build them

Antonio Ottaiano*, Italian National Institute of Statistics*

Istat

Explosion of data available

- millions of people on the web

    - data driven apps

    - mobile applications
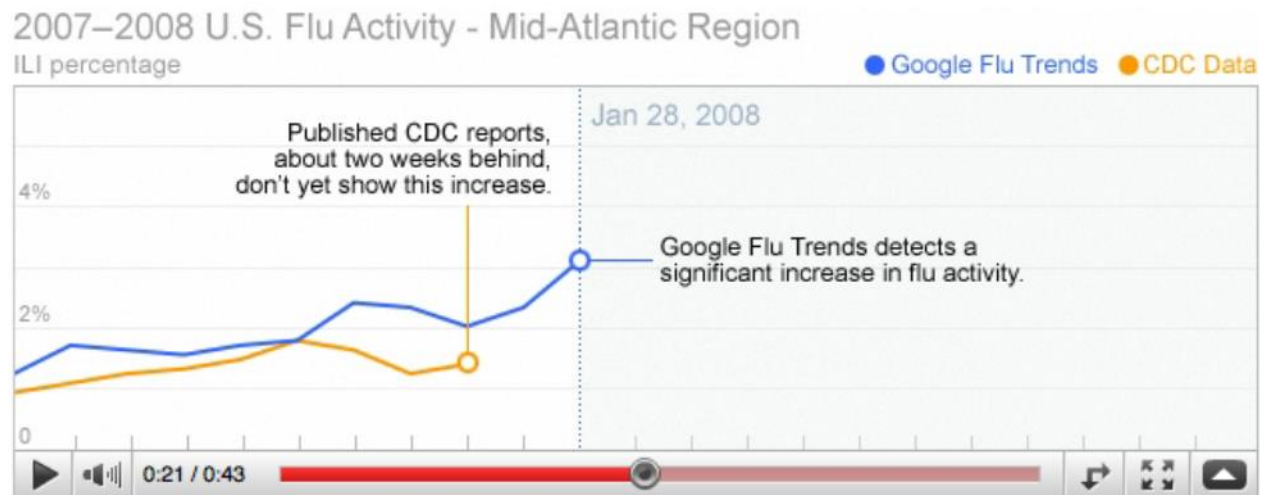
    - data trail

Istat

What do we mean by *Data science*?

■ "A data application acquires its value from the data itself, and creates more data as a result. It's not just an application with data; it's a data product. Data science enables the creation of data products" (Mike Loukides, What is Data Science?)

Where Data science begins

- PYMK - **Linked** in
- Page Rank - Google

2007–2008 U.S. Flu Activity - Mid-Atlantic Region

ILI percentage

● Google Flu Trends   ● CDC Data

Jan 28, 2008

Published CDC reports, about two weeks behind, don't yet show this increase.

Google Flu Trends detects a significant increase in flu activity.

A new job for statisticians: the data scientist. Which skills, how to build them – Antonio Ottaiano, Istat

Istat

How to use data effectively?

- Holistic approach

  - data gathering

  - data processing

  - turn data into a story

  - presenting

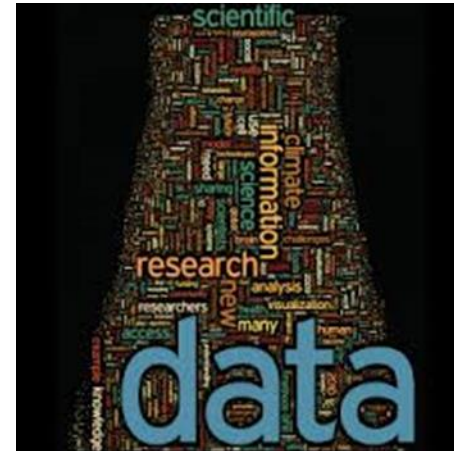## This is a job for *Data scientists*!

- people who use both data and science *"to create something new"* (D.J. Patil, J. Hammerbacher, 2008)

- explore data from multiple sources

- pick the right problem

- "make discoveries while swimming in data"

A new job for statisticians: the data scientist. Which skills, how to build them – Antonio Ottaiano, Istat

Activities

- looking for rich data sources

- working with large volumes of data

- cleaning the data

- crossing multiple datasets

- analyzing connections among data

- visualizing the data analysis

A new job for statisticians: the data scientist. Which skills, how to build them – Antonio Ottaiano, Istat

What makes a good data scientist

- Technical expertise

- Curiosity

- Storytelling

- Cleverness

Istat

An interdisciplinary profile

■ Are such profiles available?

■ Who shall we look for?

Istat

Let's grow our own data scientists – definition of a skill profile

SkillProfiles.eu

Profile sheet WSP-G3-024
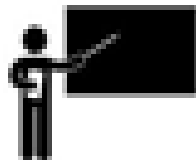"Data scientist"
G3 Web Skills Profiles - versione 2.0
Generation 3 European ICT Professional Profiles

■ IWA - *Web professional skills profile*

■ CEN   guidelines   (Generation   3
European ICT Profiles)

Istat

Let's grow our own data scientists – design of training activities

- Methodologies:

    - seminars

    - lectures

    - laboratory

        - Contents:

            - statistical methods

            - IT techniques

            - visualization tools

Istat