

Introduction –DataScience

Bigdata Vs Data Science

- Lot of Questionnaires
 - What is “BigData” Anyway?
 - What does “ data Science” mean?
 - What is the relationship between Bigdata and Data Science?
 - Is Data Science only the stuff going on in companies like Google and Facebook and tech Companies.
 - Why do many people refer to Big Data as Crossing discipline and to data science



Analytics tools like
Hadoop and Pig

Knowledge
in

Mathematics and
Statistics

Basic Statistics

Uses

Advanced Statistics

SQL querying, data
visualization tools like
Tableau, Hadoop

Proficiency
in

Python and R languages

To look at the problem
from a business
perspective

Must have
ability

To find insights
in the data

60,000\$ per annum

Earns

130,000\$ per annum

Banking

- Risk Analytics
- Operations Optimization
- Enhance customer experience using Customer Analytics
- Fraud Detection

Retail

- Predicting trends for future demand
- Customizing services
- Prize Optimization
- Inventory Management
- Identifying loyal customers
- Provide smarter shopping experience

Gaming

- Analyzing gamer data to create engaging games
- Changing game mechanics to keep customers engaged
- Tracking player statistics
- Analyzing winners & creating virtual rewards dynamically

Media

- Predicting customer interests using past data
- Optimizing the scheduling
- Promotional campaigns for customer retention
- Target Advertising

Health Care

- Analyze disease patterns & Track disease outbreaks
- Improve clinical trial design through statistical tools & algorithms
- Predict patients at risk for disease
- Preventive care using Profile Analytics

Insurance

- Predictive Analytics for fraud detection
- Predict potentially expensive claims
- Analyze driving style to cut premiums for safe drivers in vehicle insurance

Data Analysis Has Been Around for a While...

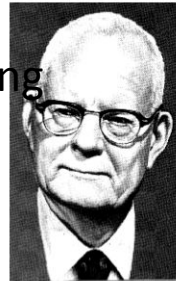
1935: "The Design of Experiments"

R.A. Fisher



1939: "Quality Control"

W.E. Deming



1958: "A Business Intelligence System"

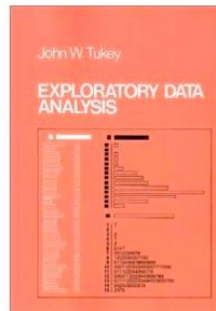


Peter Luhn

1997: "Machine Learning"



1977: "Exploratory Data Analysis"



1989: "Business Intelligence"

Howard
Dresner

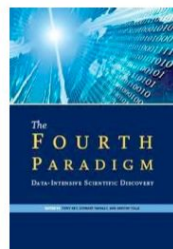


2010: "The Data Deluge"

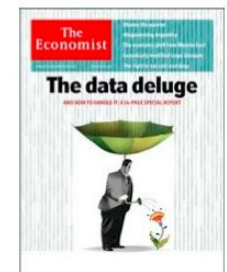
1996: Google



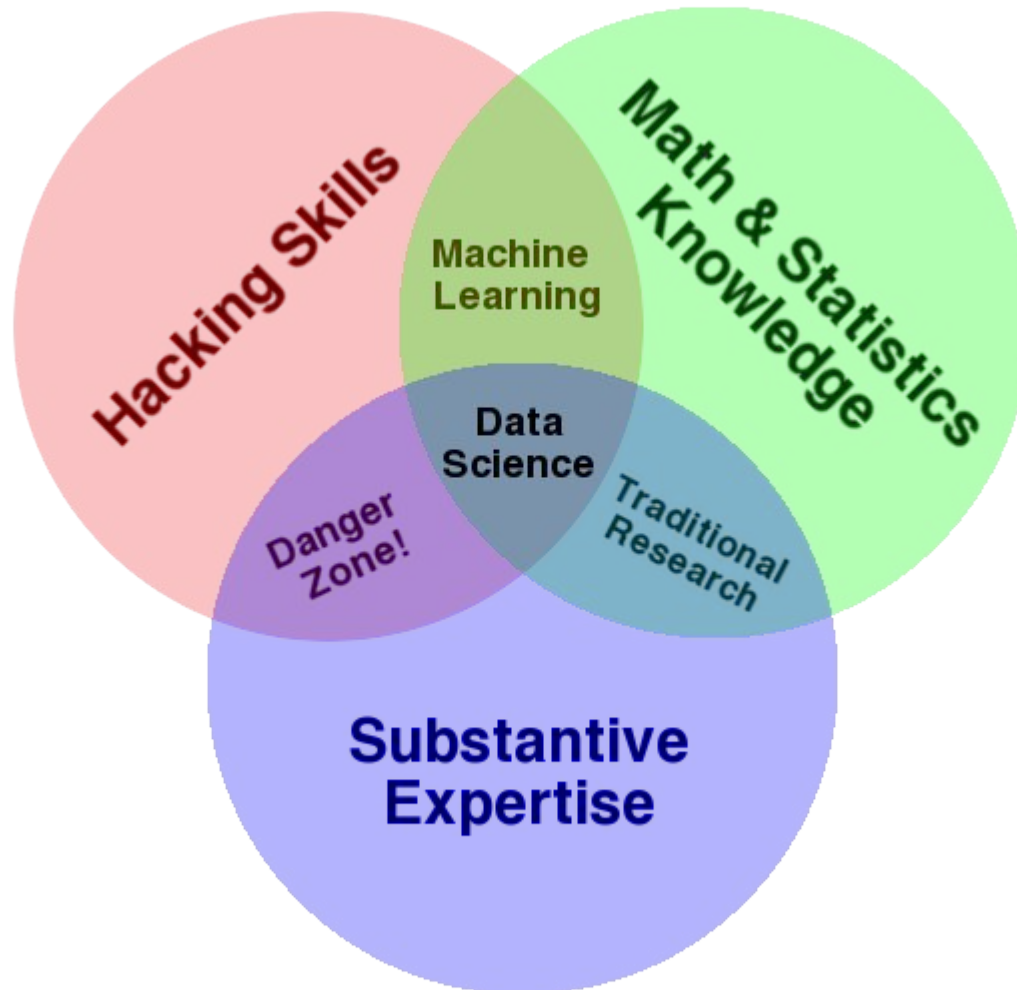
2007: "The Fourth Paradigm"



2009: "The Unreasonable Effectiveness of Data"



Data Science – A Visual Definition



Data Science

1. It is not new , the researchers in academia and industry labs were doing this (Analysis) from decades.(Centuries)

- Statisticians
- Computer Scientists
- Mathematician
- Engineers and Scientists
-

Data Science

2. Statisticians already feel that they are studying and working on the “Science of Data”
3. Data Science is not a rebranding of statistics or machine learning but rather a field itself.

Data Science- Why now

- Massive amounts of data about many aspects of our lives
- Inexpensive computing power
- Data Available
 - Shopping
 - Communicating
 - Reading news
 - Listening to music
 - Expressing our opinions etc

Data Science: Why all the Excitement?



Exciting new effective applications of data analytics

e.g.,
Google Flu Trends:

Detecting outbreaks
two weeks ahead
of CDC data

New models are estimating
which cities are most at risk
for spread of the Ebola virus.

Prediction model is built on
Various data sources,
types and analysis.

Why the all the Excitement?

elections2012

Live results | [President](#) | [Senate](#) | [House](#) | [Governor](#) | [Choose your](#)

Numbers nerd Nate Silver's forecasts prove all right on election night

FiveThirtyEight blogger predicted the outcome in all 50 states, assuming Barack Obama's Florida victory is confirmed

Luke Harding

guardian.co.uk, Wednesday 7 November 2012 10.45 EST



Predicting political
champagne and election
Outcome:

*the signal and the
and the noise and
the noise and the
noise and the no
why most noise a
predictions fail to
but some don't n
and the noise and
the noise and the
nate silver noise
noise and the no*

Data and Election 2012 (cont.)

- ...that was just one of several ways that Mr. Obama's campaign operations, some unnoticed by Mr. Romney's aides in Boston, **helped save the president's candidacy**. In Chicago, the campaign recruited a team of behavioral scientists to build an **extraordinarily sophisticated database**
- ...that allowed the Obama campaign not only to alter the very nature of the electorate, making it younger and less white, but also to create a portrait of shifting voter allegiances. **The power of this operation stunned Mr. Romney's aides on election night**, as they saw voters they never even knew existed turn out in places like Osceola County, Fla.
 - New York Times, Wed Nov 7, 2012
- The White House Names Dr. DJ Patil as the First U.S. Chief Data Scientist, Feb. 18th 2015

A history of the (Business) Internet: 1997

BackRub Search: university

BackRub Query Results


BackRub's Highest Ranked Sites

University of Illinois at Urbana-Champaign

 <http://www.uiuc.edu/>


694.687 8460 backlinks 12k - 10/25/96 - 11/1/96

Stanford University Homepage

 <http://www.stanford.edu/>

609.303 8857 backlinks 4k - none - 11/1/96

Stanford University: Portfolio Collection

 <http://www.stanford.edu/home/administration/portfolio.html>

167.919 34 backlinks

Sponsored search

[Advanced Search](#)

Web [Show options...](#)

Results 1 - 10 of about 661,000 for [gatco towel bars](#). (0.33 seconds)

Gatco Towel Bars

www.eFaucets.com/Gatco

Low Price Guarantee, Free Shipping. 110% Price Match, No Tax, Shop Now!



Gatco Towel Bars

www.AmericanHomePlus.com/Gatco

110% Low Price Guarantee. Buy Now. Free Shipping On Gatco Accessories.

Gatco Towel Bars

www.PlumberSurplus.com

Free S/H Available. Large Selection Gatco Towel Bars, ready to ship!



Gatco Towel bars, Gatco toilet paper holders, Gatco Bathroom ...

Gatco Towel bars, Gatco toilet paper holders, Gatco Bathroom Accessories, Gatco robe hooks, Gatco bathroom shelves, Gatco hotel shelves, Gatco double towel ...

www.kitchensnbath.com/gatco-bath-accessories.html - [Cached](#) - [Similar](#) -

Discount Towel Bars, Towel Bars, Towel Racks

Discount Towel Bars offers only the highest quality bath hardware. We are factory direct distributors for Moen, Baldwin, Dynasty Hardware and Gatco. ...

www.discounttowelbars.com/ - [Cached](#) - [Similar](#) -

Gatco at Lowe's: 24" Franciscan Chrome Double Towel Bar

24" Franciscan Chrome Double Towel Bar - 69656 5286.

www.lowes.com/lowes/lkn?action=productDetail... - [Cached](#) - [Similar](#) -

Shopping results for gatco towel bars



[Gatco Bleu 18 in. Towel Bar - Polished Chrome](#)

\$39.99 new - [Sears](#)

[Gatco 4240 24-Inch Latitude II Towel Bar, Chrome](#)

\$30.53 new - [Amazon.com](#)

[4621 Camden Towel Bar 18 Gatco Inc](#)

Sponsored Links

Sponsored Links



[Gatco Bleu 18 in. Towel Bar - Polished Chrome](#)

\$39.99 - [Sears](#)



[Gatco Spa Towel Rack, 3 Tier - Satin Nickel](#)

\$94.99 - [Sears](#)



[Gatco Chenille 18 in. Towel Bar - Vintage ...](#)

\$44.99 - [Sears](#)

Gatco Bath Accessories

Lowest prices.

All Gatco collections

www.TheHomeDecor.net

Towel Bars on Sale

Save 20%-50% Off List- New Styles

Lowest Prices + Free Shipping!

www.FixtureUniverse.com



Sponsored search

- Google revenue around \$50 bn/year from marketing, 97% of the companies revenue.
- Sponsored search uses an auction – a pure competition for marketers trying to win access to consumers.
- In other words, a competition for **models** of consumers – their likelihood of responding to the ad – and of determining the right bid for the item.
- There are around 30 billion search requests a month. Perhaps a **trillion events** of history between search providers.
- Google Adwords and Adsense

TOP 20

Keyword Categories

Percentages correspond to the number of keywords in the top 10,000 keywords that belong to that category.



Other Data Science Applications

- Transaction Databases → Recommender systems (NetFlix), Fraud Detection (Security and Privacy)
- Wireless Sensor Data → Smart Home, Real-time Monitoring, Internet of Things
- Text Data, Social Media Data → Product Review and Consumer Satisfaction (Facebook, Twitter, LinkedIn), E-discovery
- Software Log Data → Automatic Trouble Shooting (Splunk)
- Genotype and Phenotype Data → Epic, 23andme, Patient-Centered Care, Personalized Medicine

Datafication

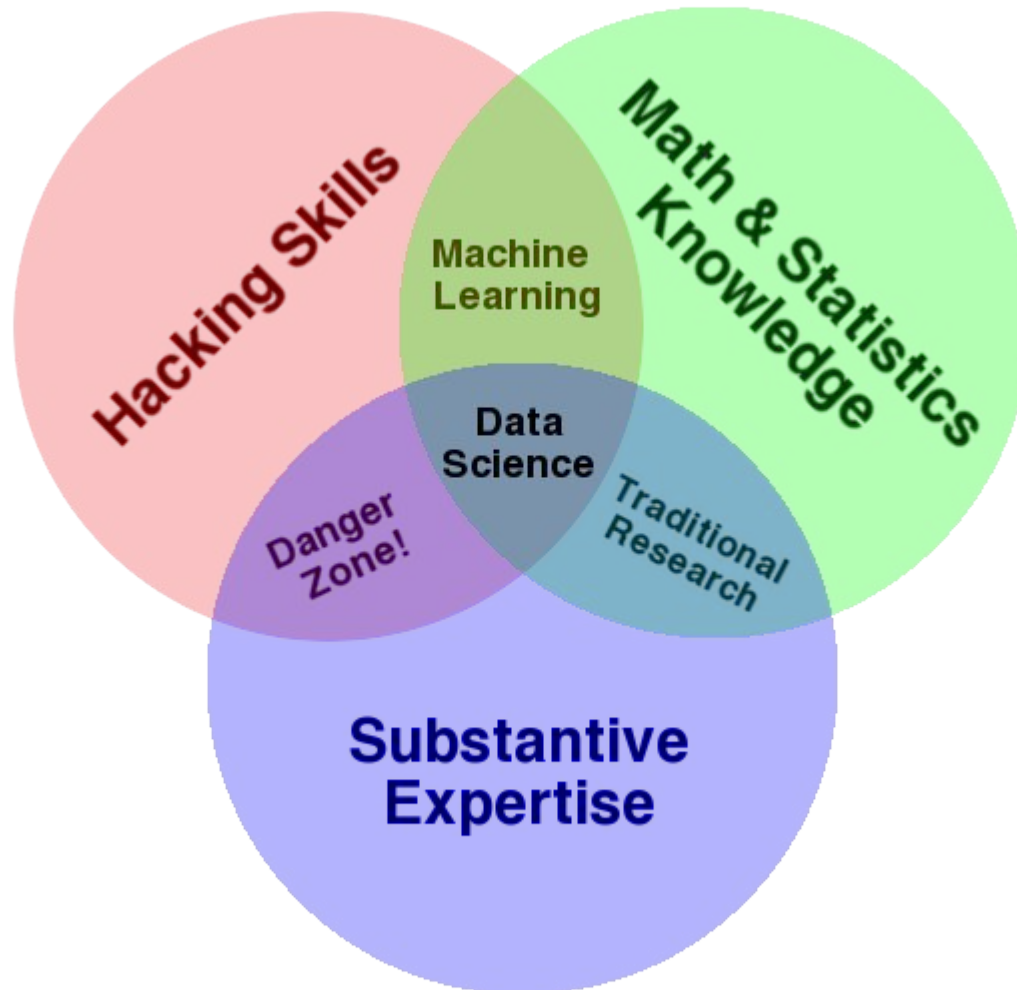
- Datafication is a process of “taking all aspects of life and turning them into data”
 - Example
 - Google’s Augmented-reality glasses datafy the gaze
 - Twitter datafies stray thoughts
 - Linkedin datafies professional networks

Datafication

- Datafication is an interesting Concept
 - It deals with
 - When we people's share intention , intentions are respected. i.e datafied.

Once we datafy things, we can transform their purpose and turn the information into new forms of value.

Data Science – A Visual Definition



- Rise of the Data Scientists, which includes
 - Statistics(Traditional analysis you're used to think about)
 - Data Munging(Parsing,scraping and formatting data)
 - Visualization(Graphs, Tools, etc)