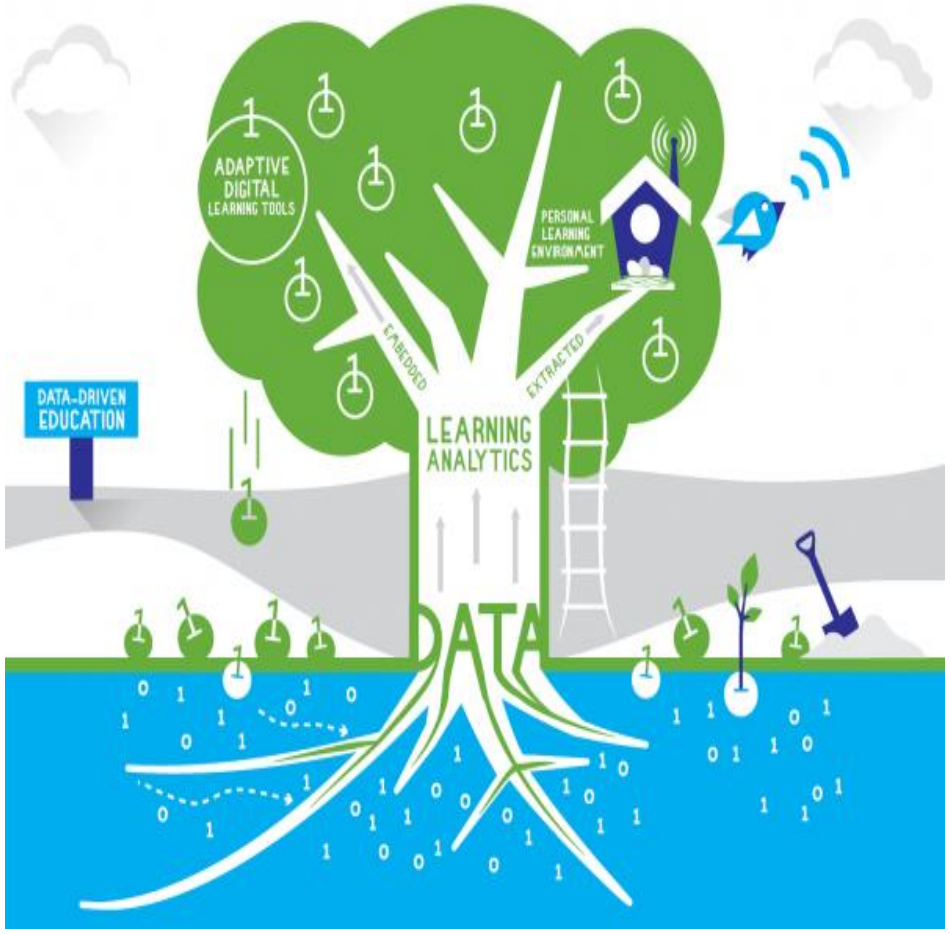
BUSINESS ANALYTICS :: AN INTRODUCTION

DEFOUR ANALYTICS

Welcomes you all!!

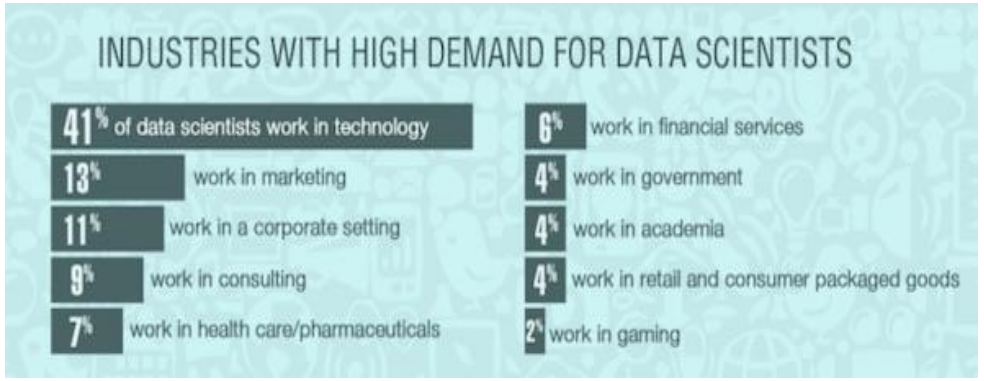
Welcome Note

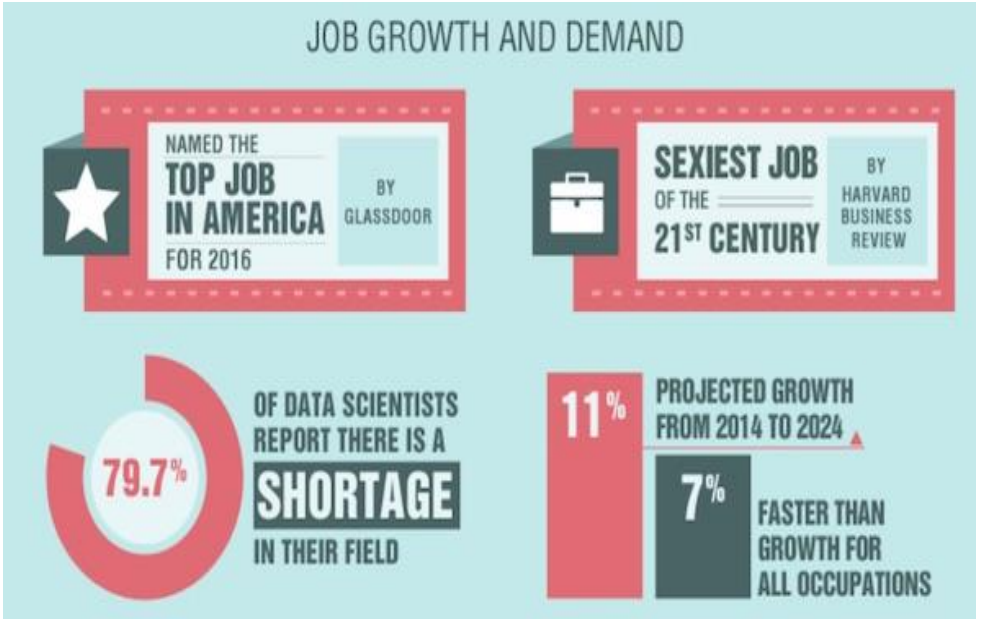


Lets drive together in the journey of “Learning Analytics” in the pursuit of “Digital Transformation”

The Future

- Data literacy will be critical survival element in the industry as “The price of light is less than the cost of darkness”
- Data Scientist need to be a feather in the cap of every engineer to become a decision maker
- Every failure is detectable and predictable if appropriate analytics principles are applied at right time
- Data Scientist are in high demand as they can predict the future

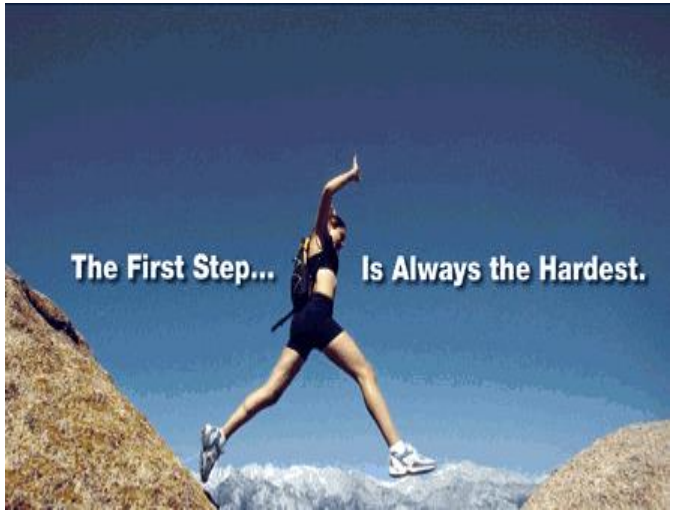




Perspective

✓Gains

- Capitalize on your business knowledge for taking pro-active decisions
- Improve Efficiencies & Effectiveness of processes
- Become a driver for Digital Transformation
- Enable Risk based management thru predictive analytics
- Empower, Enhance your stake in the organization decision-making
- Cross-functional analytics for uncover the hidden capacities



✓Pains

- Invest sometime on Analytics
- Practice with your own data and defined problems

"Do it yourself"

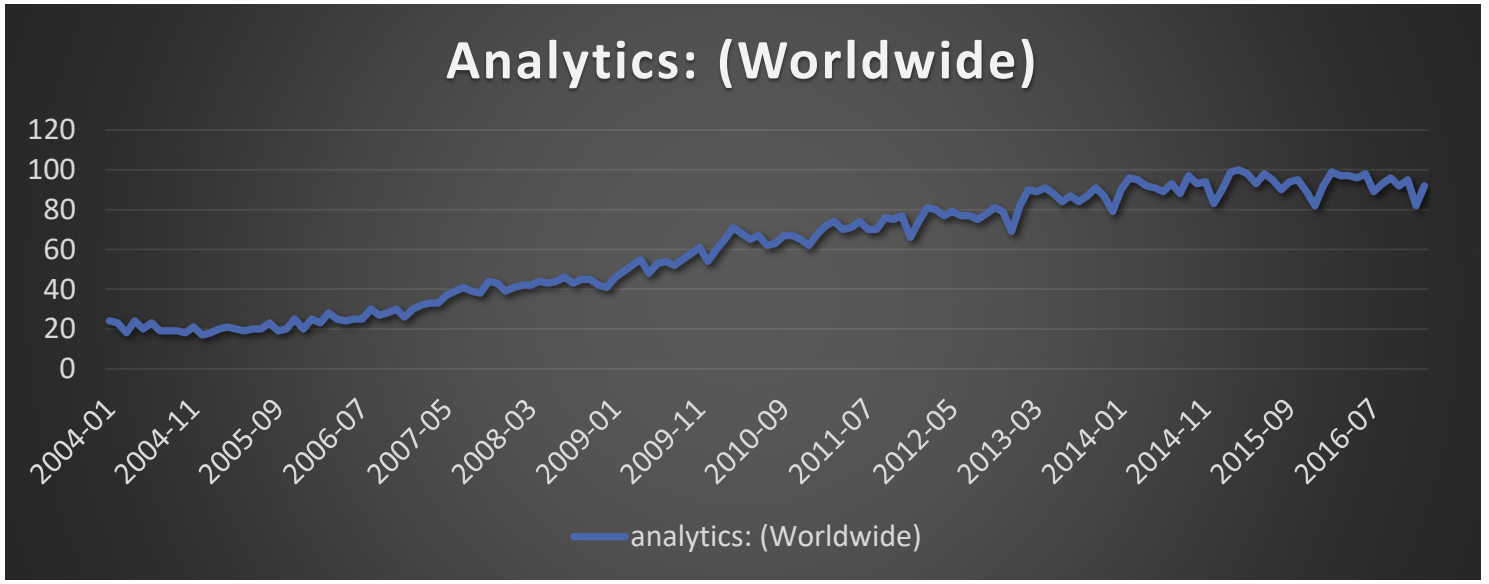
Famous Quotes on Analytics

defour analytics

"Numbers have an important story to tell. They rely on you to give them a voice."
– Stephen Few

"You can have data without information, but you cannot have information without data.

"Without big data analytics, companies are blind and deaf, wandering out onto the web like deer on a freeway."
– Geoffery Moore



defour analytics

BUSINESS ANALYTICS

Definition of BA

Business Analytics is the process of converting data into insights. It is "the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions."

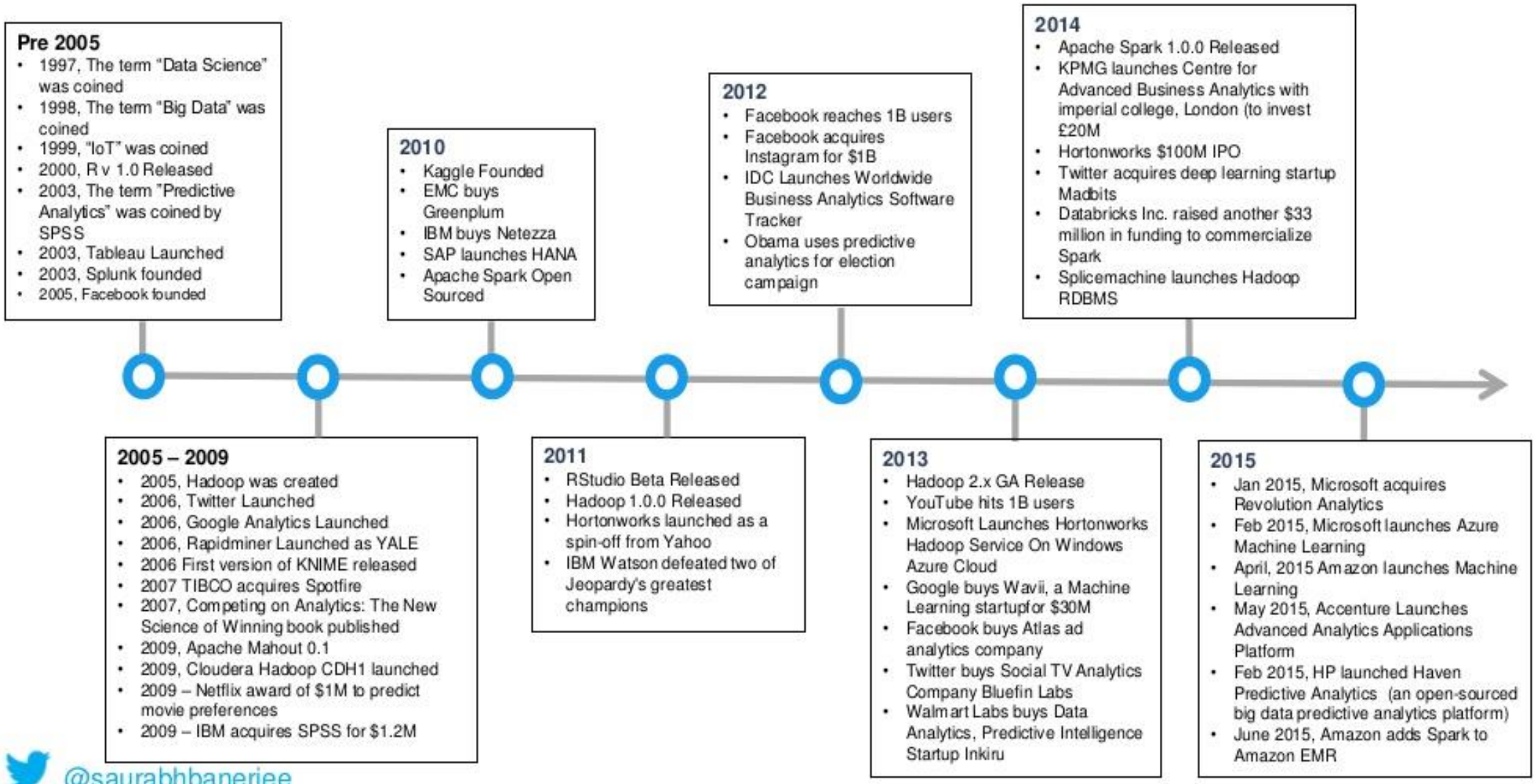
With the increase in the availability of data, Analytics has now become a major differentiator in both the top line and the bottom line of any organization.

It is hence not surprising that research has shown that data-driven companies perform 5%-6% better per annum.

BA is the use of:

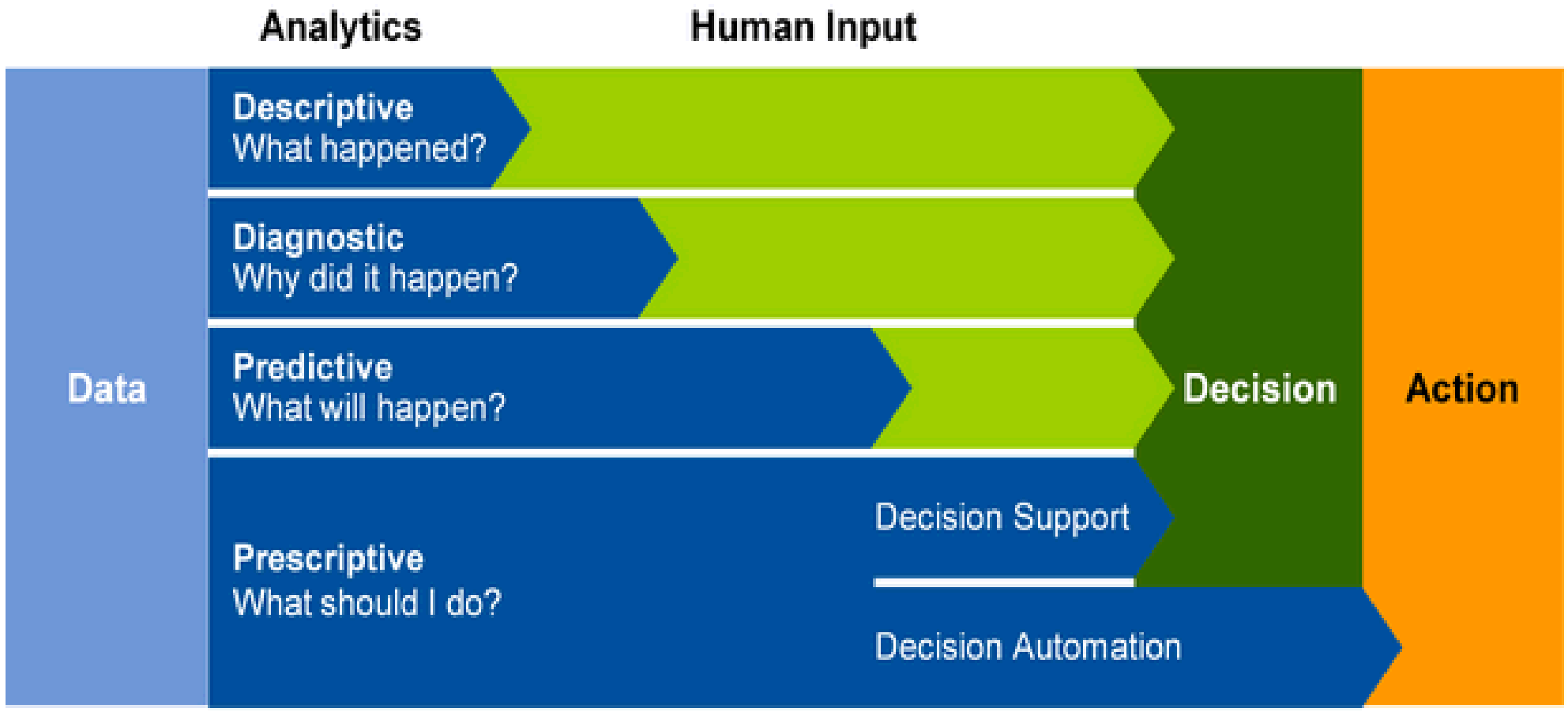
- data,
- Information technology,
- statistical analysis,
- quantitative methods, and
- mathematical or computer-based models

Evolution of Analytics – a timeline perspective



twitter @saurabhbanerjee

Business Analytics Maturity Stages



Descriptive analytics or data mining are at the bottom of the data value chain, but they can be valuable for uncovering patterns that offer insight. Descriptive analytics can be useful in the sales cycle, for example, to categorize customers by their likely product preferences and sales cycle.

Diagnostic analytics are used for discovery or to determine why something happened. For example, there can be thousands of online mentions that can be distilled into a single view to see what worked in your past campaigns and what didn't.

Predictive analytics use data to identify past patterns to predict the future. For example, some companies are using predictive analytics for sales lead scoring. Properly tuned predictive analytics can be used to support sales, marketing, or for other types of complex forecasts.

Prescriptive analytics is really valuable, but largely not used. According to Gartner, 13 percent of organizations are using predictive but only 3 percent are using prescriptive analytics. Prescriptive analytics gives you a laser-like focus to answer specific questions.

Uses of Business Analytics

defour analytics

- Exploring data to find new patterns and relationships (data mining)
- Explaining why a certain result occurred (statistical analysis, quantitative analysis)
- Experimenting to test previous decisions (A/B testing, multivariate testing)
- Forecasting future results (predictive modeling, predictive analytics)

Answers the questions:

Why did it happen?

Will it happen again?

What will happen if we change *x*?

What else does the data tell us that never thought to ask?

BUZZ WORDS!!

Artificial Intelligence

Machine learning

Big Data

Hadoop

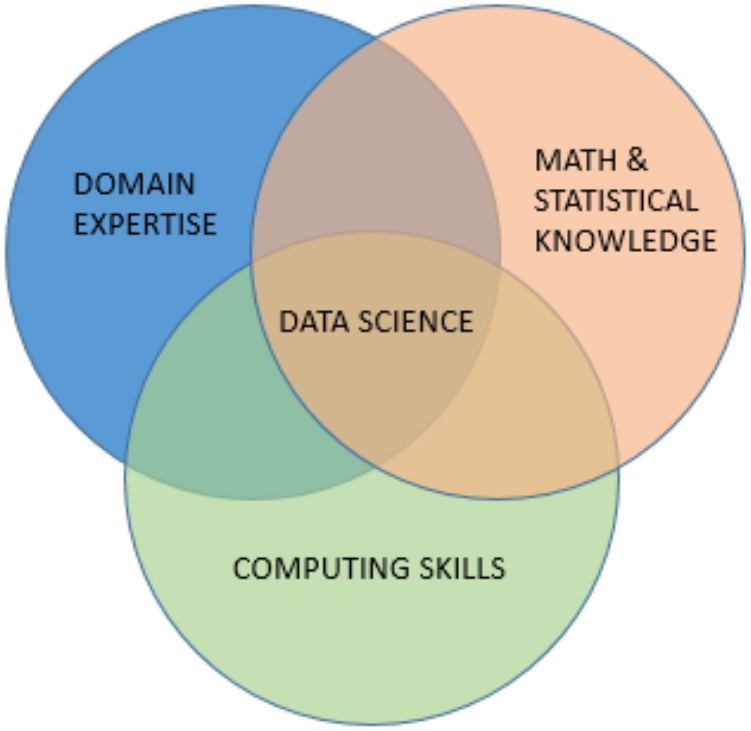
Data science

Internet Of Things


data
science
research
scientists
support
faculty
environment
students
activities
fellows
sciences
program
institute
metrics
budget
career
methodology
committee
social
across
scientific
funding
areas
one
university
physical
statistics
center
engineering
studio
reproducibility
tools
projects
space
working
researchers
section
project
education
uw
director
bridge
escience
incubator
activity
evaluation
graduate
participants
establish
established
campus
year
impact
time
including
training
participate
paths
software
number
development
professor
appendix
computer
theme
programs
new
work
funded
initiative
courses
postdocs
flagship
ethnography
steering
collaborations
effort
group
community
administrative
people
collaboration
fellow
environments
long-term
domain
described
open
alternative
see
use
goal
team
grant
successful
three
current
among
fields
success
events
online
libraries

Data science

Data science, also known as **data-driven science**, is an interdisciplinary field about scientific processes and systems to extract knowledge or insights from data in various forms, either structured or unstructured, which is a continuation of some of the data analysis fields such as statistics, machine learning, data mining, and predictive analytics.



5 Questions that Data Science answers?

defour analytics

1. *Is this A or B?*
   Classification Algorithm
2. *Is this weird?*
   Anomaly Detection Algorithm
3. *How Much / How Many?*
   Regression Algorithm
4. *How is this organized?*
   Clustering Algorithm
5. *What should I do next?*
   Reinforcement Learning Algorithm

5 Questions that Data Science answers?

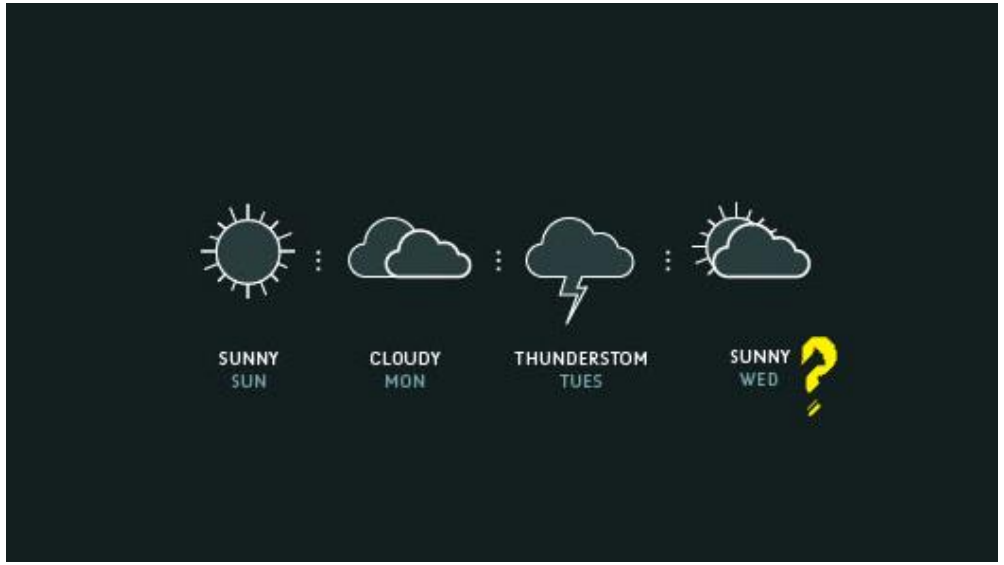
defour analytics

1. *Is this A or B?*
2. *Is this weird?*
3. *How Much / How Many?*
4. *How is this organized?*
5. *What should I do next?*

1. Classification Algorithm

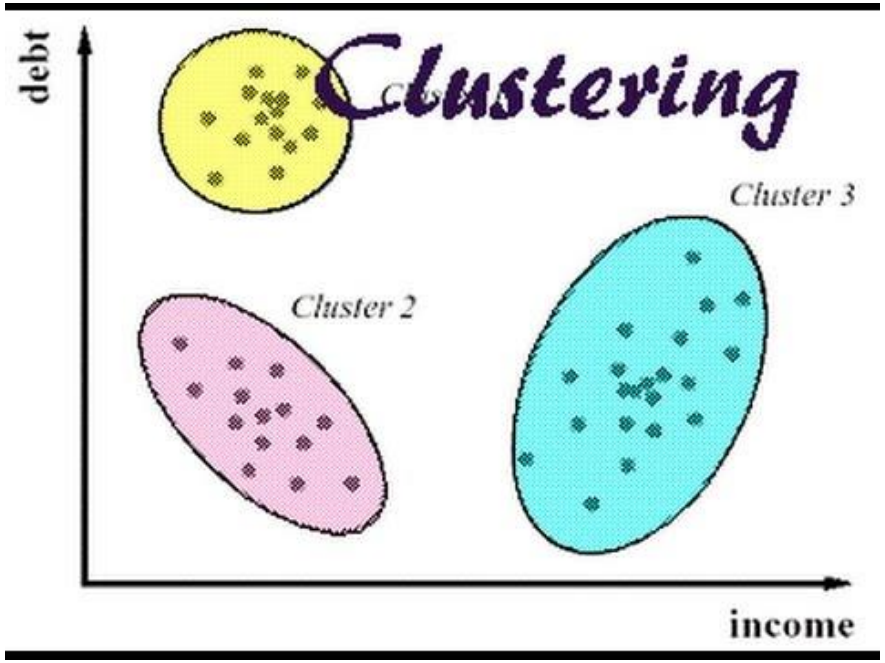
defour analytics

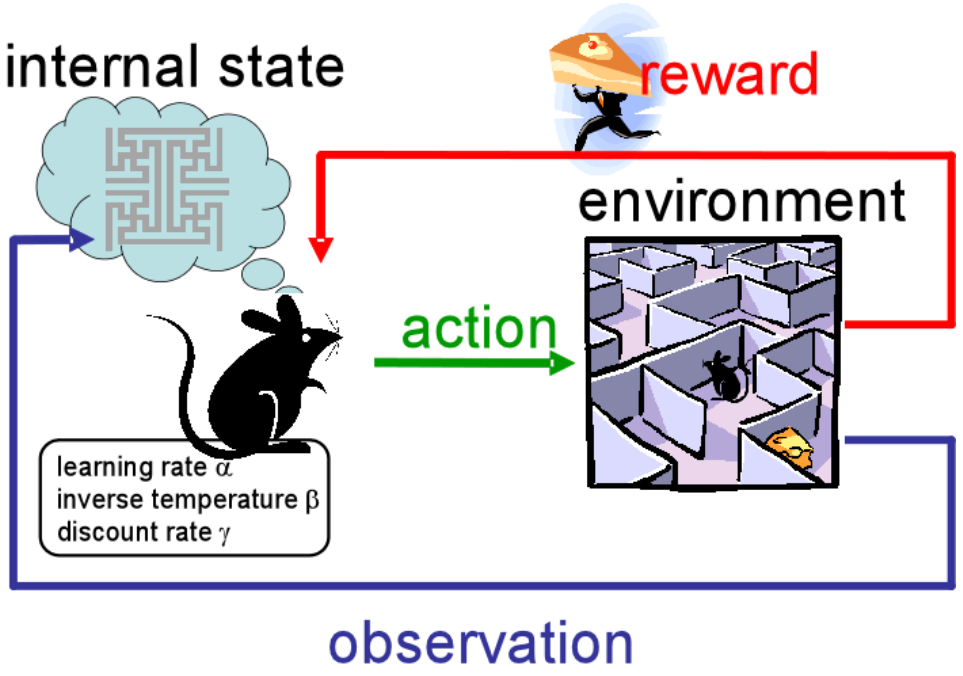
2. Anomaly Detection Algorithm



3. Regression Algorithm

5. Clustering Algorithm





6. Reinforcement learning Algorithm

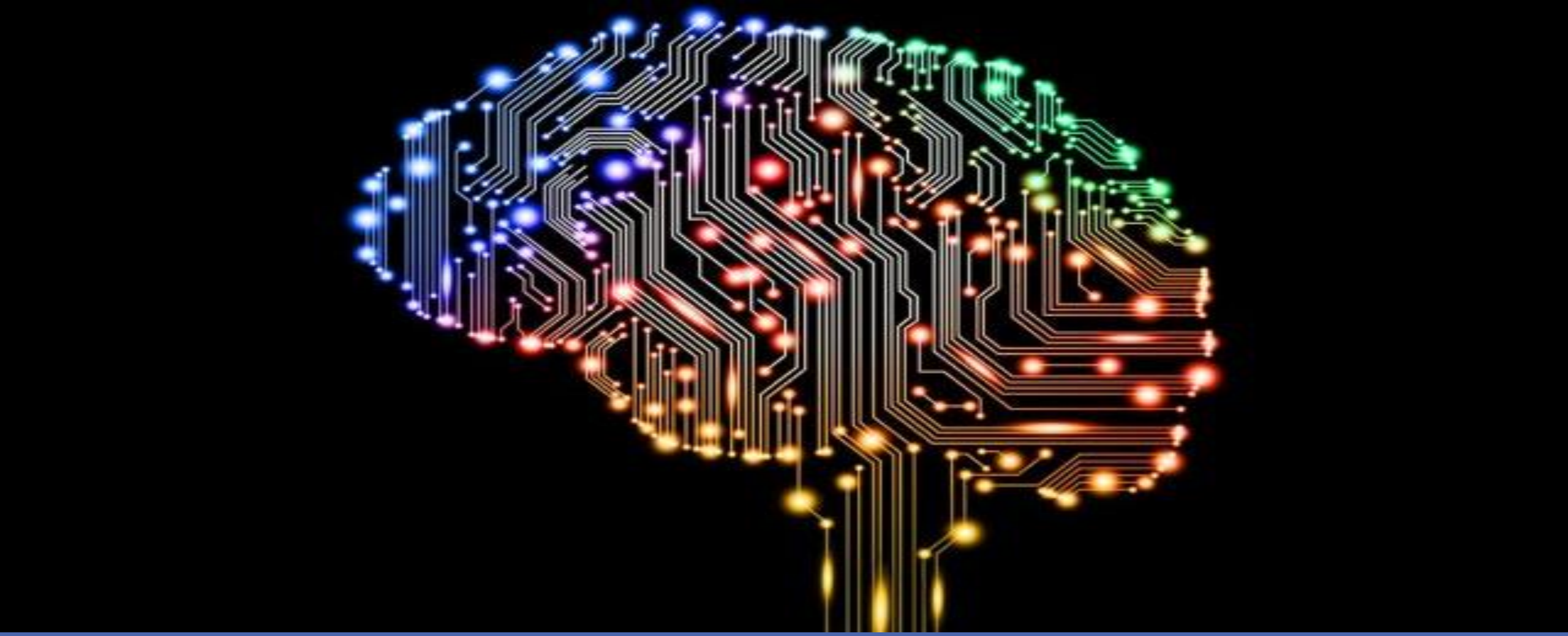
Applications of Data Science

defour analytics

Internet search: Search engines make use of data science algorithms to deliver best results for search queries in fraction of seconds.

Digital Advertisements: The entire digital marketing spectrum uses the data science algorithms - from display banners to digital billboards. This is the main reason for digital ads getting higher CTR than traditional advertisements.

Recommender systems: The recommender systems not only make it easy to find relevant products from billions of products available but also adds a lot to user experience. A lot of companies use this system to promote their products and suggestions in accordance to the user's demands and relevance of information. The recommendations are based on the user's previous search results.



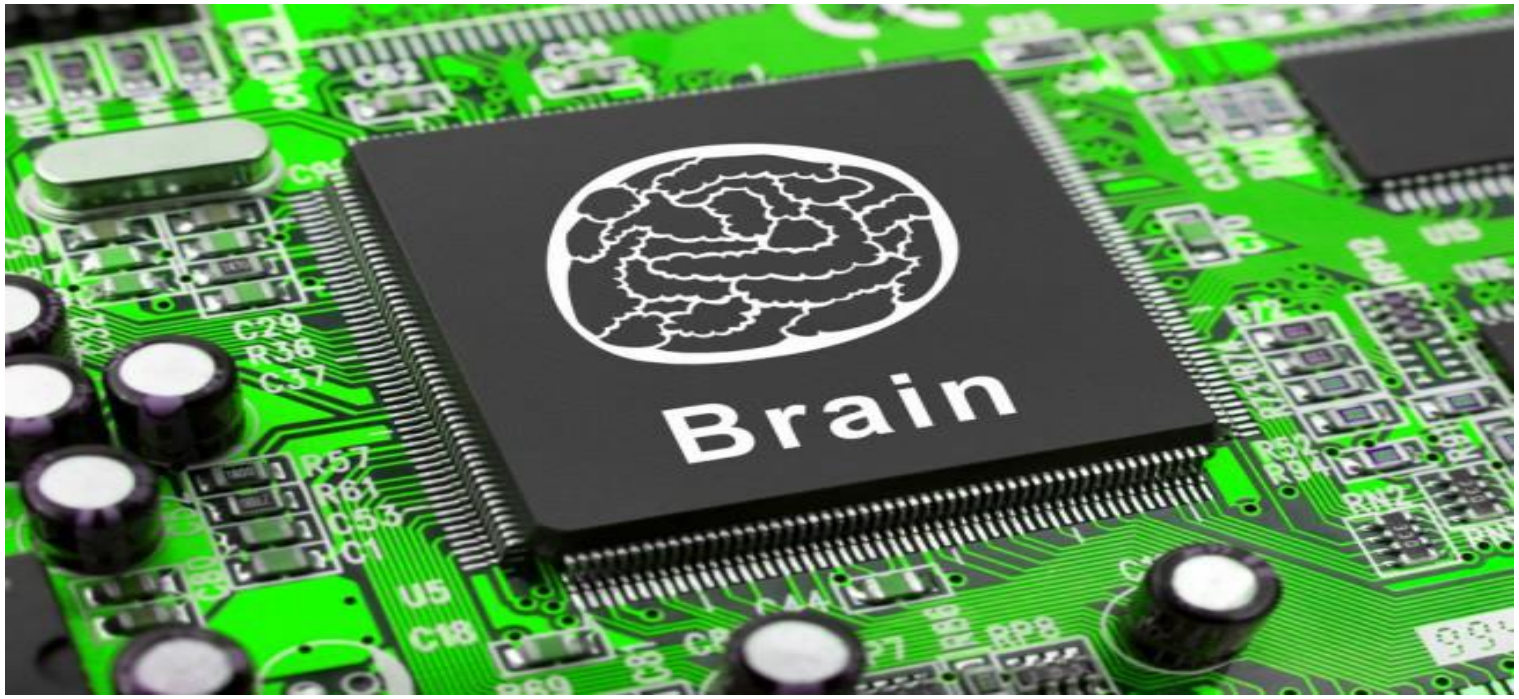
Artificial Intelligence
&
Machine learning

Artificial Intelligence

defour analytics

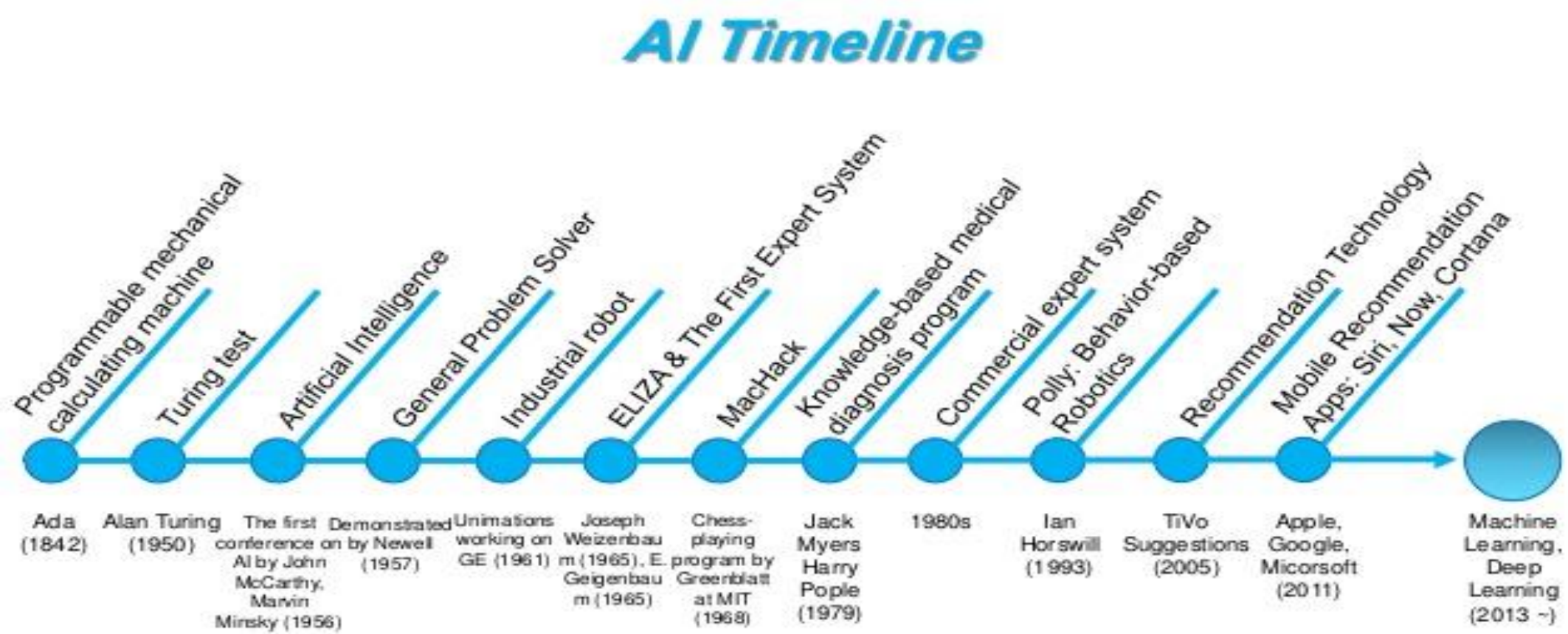
The capability of a machine to imitate intelligent human behavior

It can also be defined as a branch of computer science dealing with the simulation of intelligent behavior in computers



Evolution of AI

defour analytics



Artificial Intelligence is constantly pushing the boundaries of what machines are capable of.

But could machines ever become better than us?

Yes !!!!

Lets consider some cases …

- Search the web quicker
- Work in deadly environments
- Translate in many languages
- Deliver a correct medical diagnosis

Research Areas in AI

The **Expert Systems** are the computer applications developed to solve complex problems in a particular domain, at the level of extra-ordinary human intelligence and expertise.

The expert systems are capable of –

- Advising
- Instructing and assisting human in decision making
- Demonstrating
- Deriving a solution
- Diagnosing
- Explaining
- Interpreting input
- Predicting results
- Justifying the conclusion
- Suggesting alternative options to a problem

Fuzzy logic is useful for commercial and practical purposes.

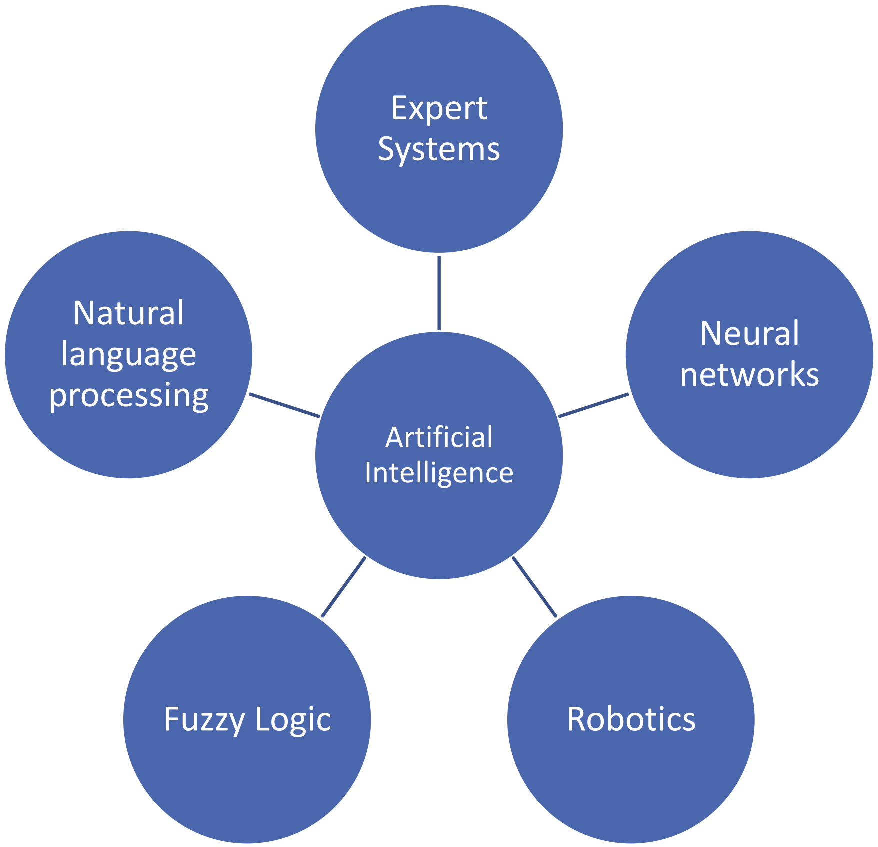
- It can control machines and consumer products.
- It may not give accurate reasoning, but acceptable reasoning.
- Fuzzy logic helps to deal with the uncertainty in engineering.

The input and output of an NLP system can be:

- Speech
- Written Text

number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs.

- Aerospace
- Automotive
- Industrial
- Medical
- Anomaly detection



Applications Of AI

defour analytics

- Gaming
- Natural Language Processing
- Expert Systems
- Vision Systems Speech Recognition
- Handwriting Recognition
- Intelligent Robots

Machine learning

defour analytics

- Machine learning is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed.
- It focuses on the development of computer programs that can change when exposed to new data.
- It uses that data to detect patterns in data and adjust program actions accordingly.

Applications of Machine Learning

defour analytics

- The heavily hyped, self-driving Google car? The essence of machine learning.
- Online recommendation offers such as those from Amazon and Netflix
- Knowing what customers are saying about you on Twitter? Machine learning combined with linguistic rule creation.
- Fraud detection! One of the most important uses in our world today.

Types of Machine Learning

defour analytics

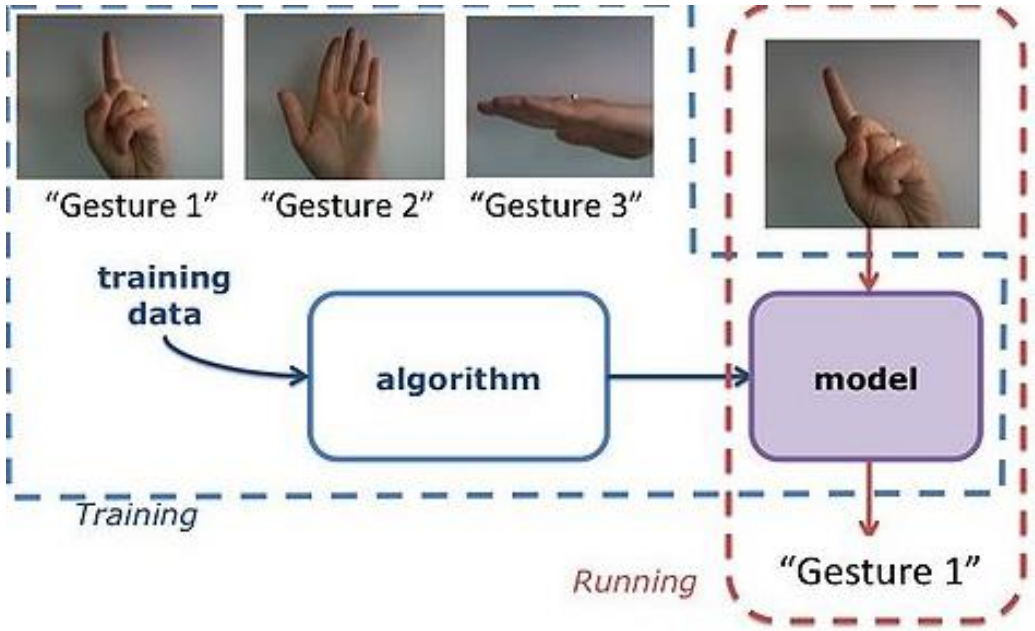
Machine learning algorithms are often categorized as:

Supervised Learning

- Makes machine learn explicity
- Data with clearly defined output is given
- Predicts outcomes/Future
- Resolves regression and classification problems

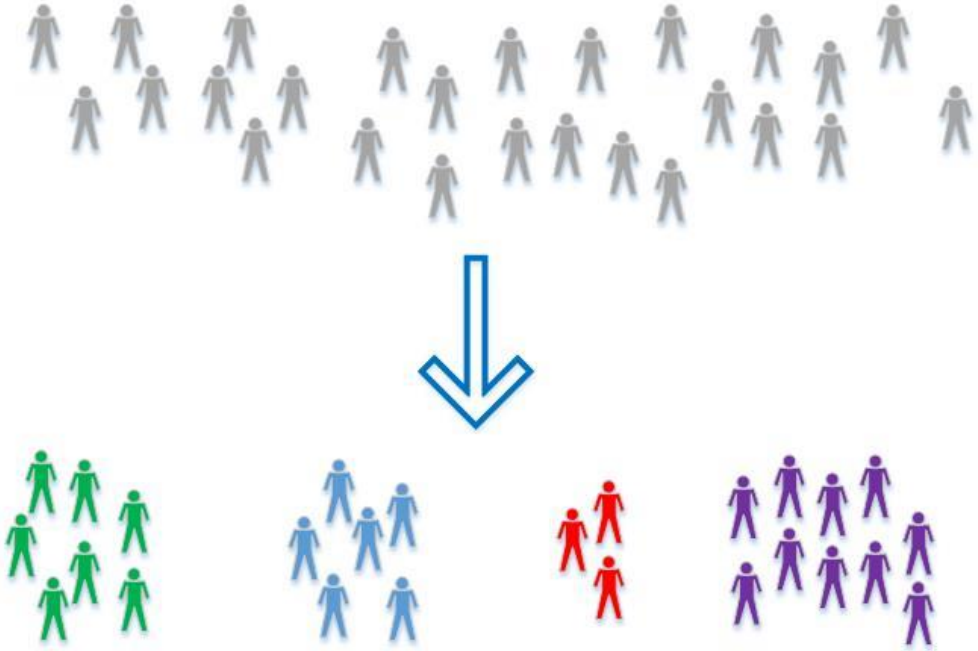
Unsupervised Learning

- Machines understands the Data (Identifies Patterns /Structure)
- Evaluation is qualitative or indirect
- Does not predict or find anything specific
- Helps in clustering



Supervised Learning

Unsupervised Learning



Difference between AI and Machine Learning

defour analytics

- AI and Machine learning are interchangeably used
- But while AI and machine learning are very much related, they are not quite the same thing.
- AI is a branch of computer science attempting to build machines capable of intelligent behavior
- Whereas ML is the science of getting computers to act without being explicitly programmed

Lets consider a basic example of Google!

AI is making major shifts in society by better medical diagnosis, self driven cars, etc..

AI is the science and machine learning is the algorithms that make the machines smarter.

defour analytics

Internet of Things (IOT)

defour analytics

The Internet of Things (IoT) is a system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction.

What exactly is a *'Thing'*, in the Internet of Things??

It can be a person with a heart monitor implant, a farm animal with a biochip transponder, an automobile that has built-in sensors to alert the driver when tire pressure is low -- or any other natural or man-made object that can be assigned an IP address and provided with the ability to transfer data over a network.

Applications of IOT

- Smart home….
- Wearables. ...
- Smart City. ...
- Industrial internet. ...
- Connected car. ...
- Connected Health (Digital health/Telehealth ...
- Smart retail.

IOT and Analytics

defour analytics

Gartner previously estimated that IoT adoption will grow to 26 billion Internet-enabled devices installed by 2020.

There will a lot of fast moving data --- BIG DATA!!!!

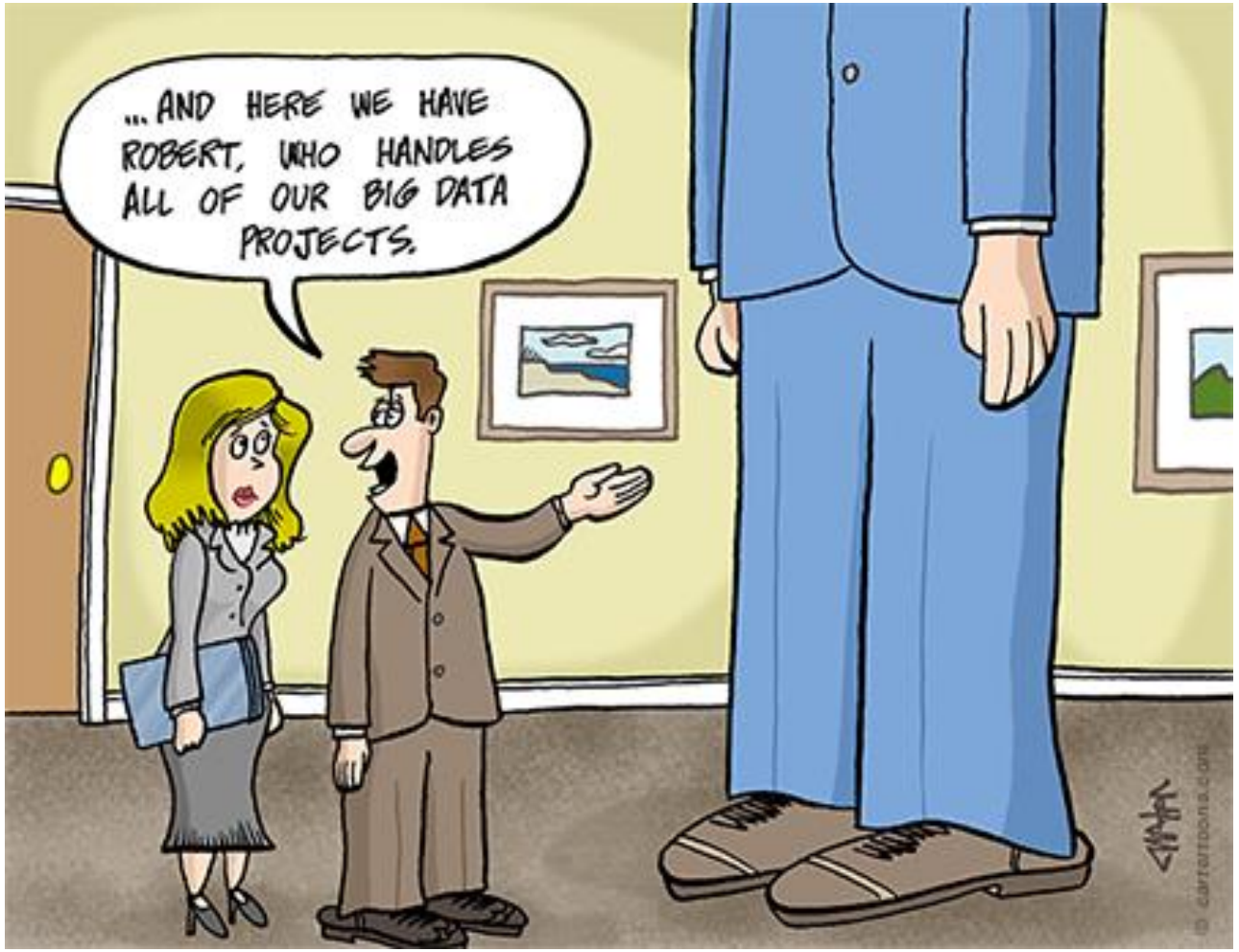
The critical challenge will using this data when it is still in motion – and extracting valuable information from it.

It will be exciting about the possibilities that IoT brings in terms of diversifying data sources and enabling analytics to be applied to new areas, such as the monitoring of sensors attached to manufacturing devices.

IoT could be a boon to analytics by serving as a major source of data.

Video On IOT

defour analytics


...AND HERE WE HAVE ROBERT, WHO HANDLES ALL OF OUR BIG DATA PROJECTS.

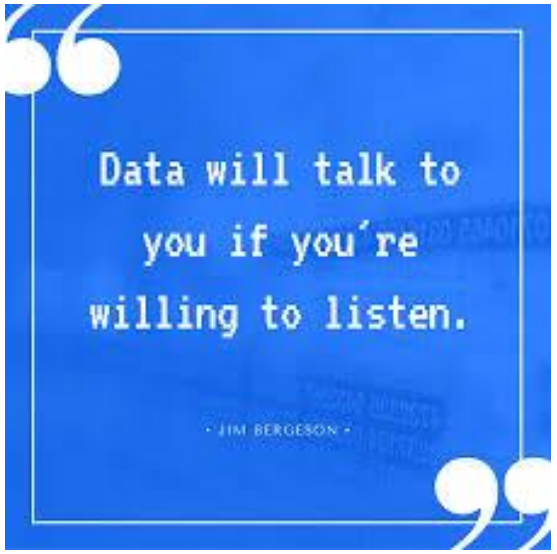
BIG DATA

What is BIG DATA??

it is a collection of large datasets that cannot be processed using traditional computing techniques. Big data is not merely a data, rather it has become a complete subject, which involves various tools, techniques and frameworks.

But it’s not the amount of data that’s important. It’s what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.

90% of the world’s data was generated in the last few years.



What comes under BIG Data

defour analytics

The BIG Data include 3 Vs

defour analytics

Volume: Organizations collect data from a variety of sources, including business transactions, social media and information from sensor or machine-to-machine data.

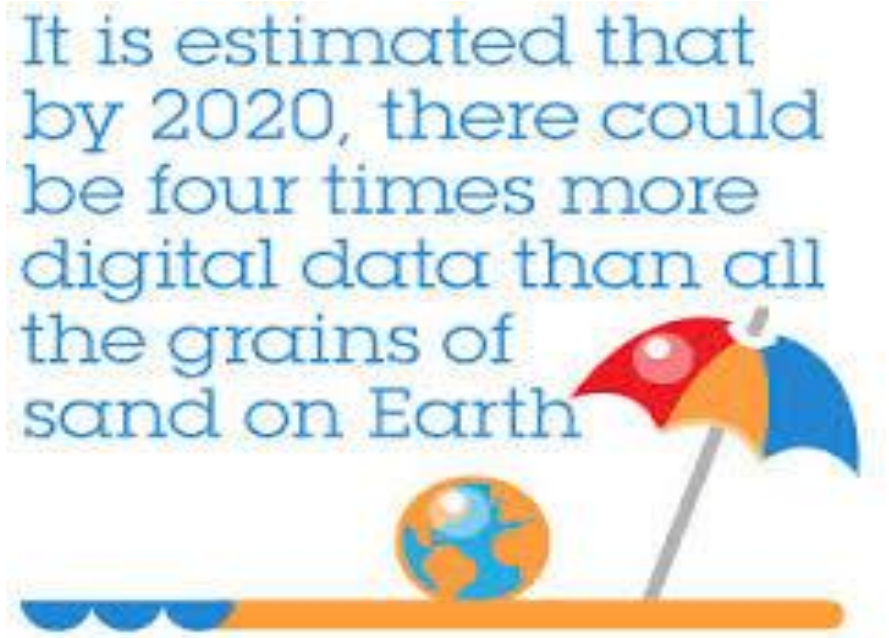
Velocity: Data streams in at an unprecedented speed and must be dealt with in a timely manner.

Variety: Data comes in all types of formats
Majorly it will be of three types.
Structured data : Relational data.
Semi Structured data : XML data.
Unstructured data : Word, PDF, Text, Media Logs, Videos, Emails, Audios, etc

Big Data Challenges

The major challenges associated with big data are as follows:

- Capturing data
- Curation
- Storage
- Searching
- Sharing
- Transfer
- Analysis
- Presentation



To meet these challenges we have solution ……

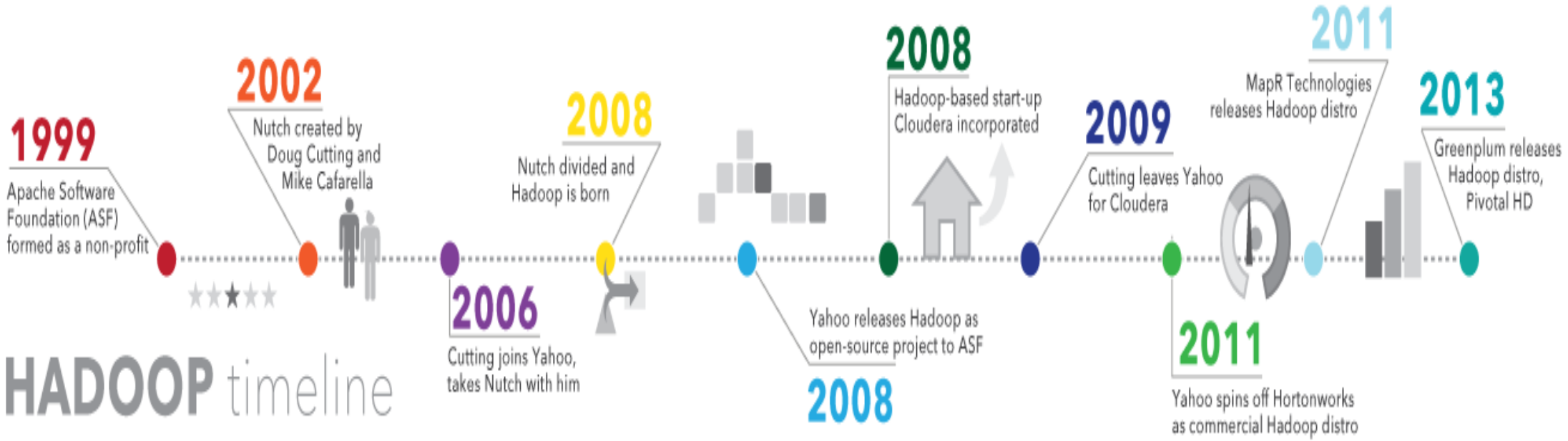

HADOOP
BIG DATA
OPEN SOURCE
BUSINESS
STORAGE
ANALYSIS
INTERNET
DATA
SECURITY
CLOUD
ANALYTICS
SYSTEMS
MANAGE
SHARE
RISK
TREND
SOCIAL
MEDIA
ENTERPRISE
PROJECTS
LEARNING
COMPUTING
VOLUME
SOLVED
SOFTWARE
TECHNOLOGIES
DATABASE
NETWORKS
COMPLEX
LARGE
SERVER
OPEN SOURCE

Hadoop!!

Hadoop is an open-source software framework for storing data and running applications on clusters of commodity hardware.

It provides massive storage for any kind of data, enormous processing power and the ability to handle virtually limitless concurrent tasks or jobs.

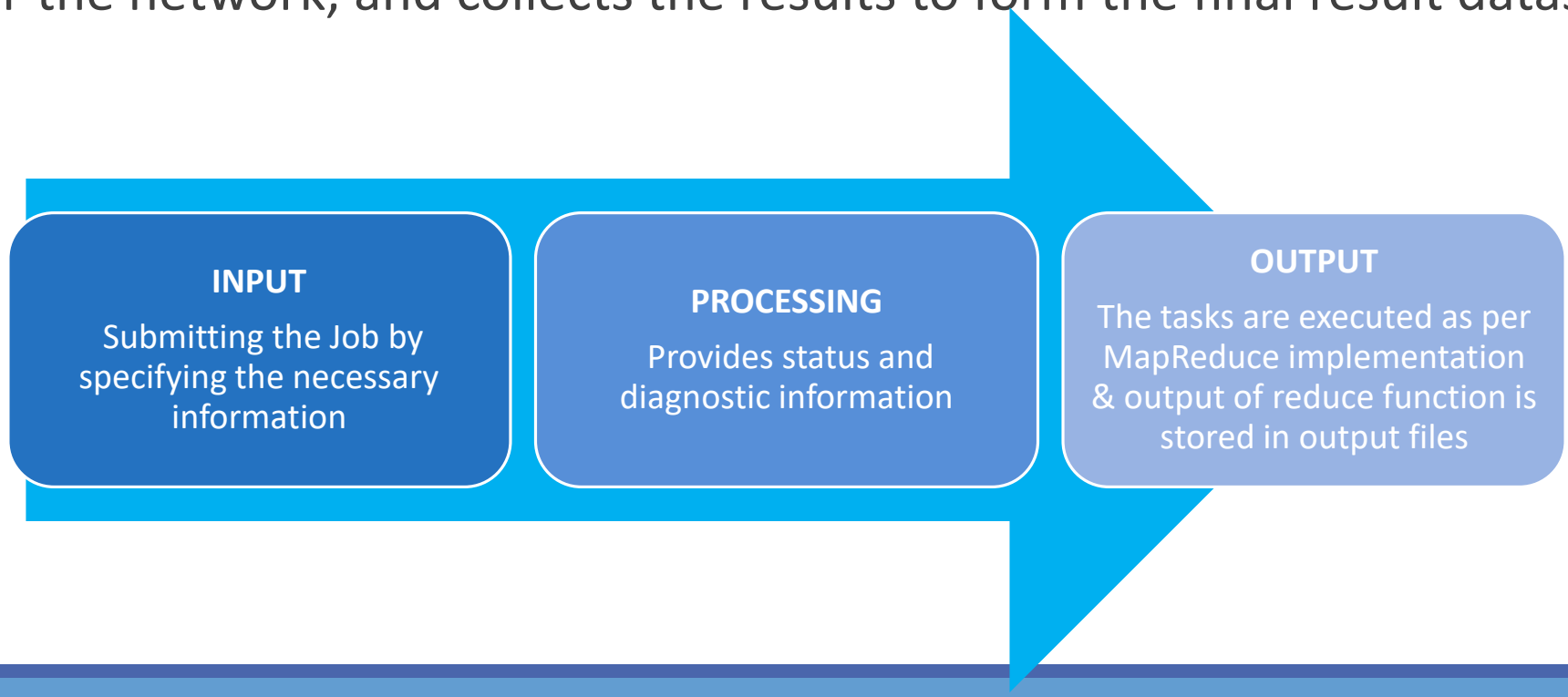
Hadoop History..



How It Works?

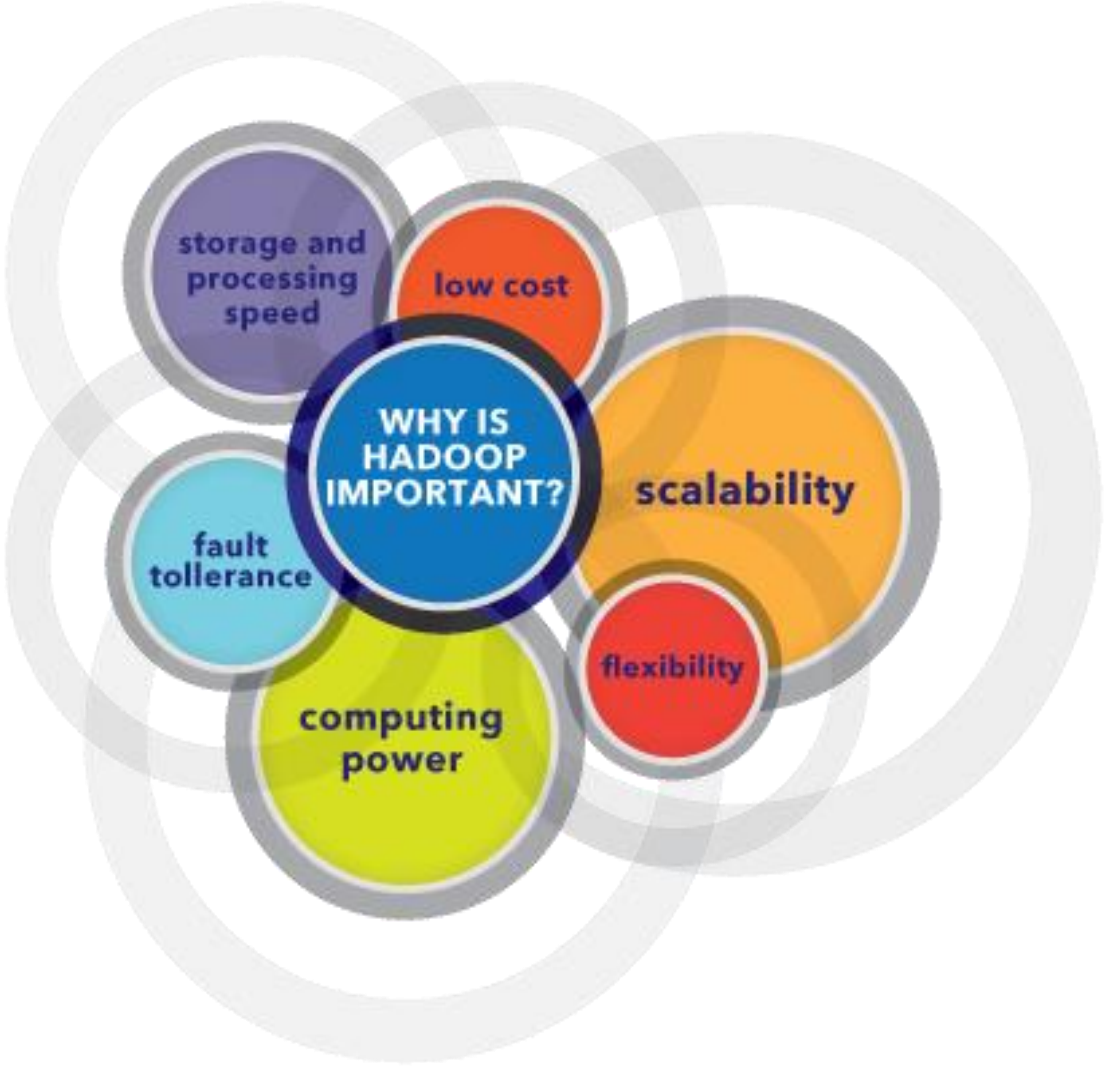
Hadoop runs applications using the MapReduce algorithm, where the data is processed in parallel on different CPU nodes.

This algorithm divides the task into small parts and assigns those parts to many computers connected over the network, and collects the results to form the final result dataset.



Why is Hadoop important?

defour analytics

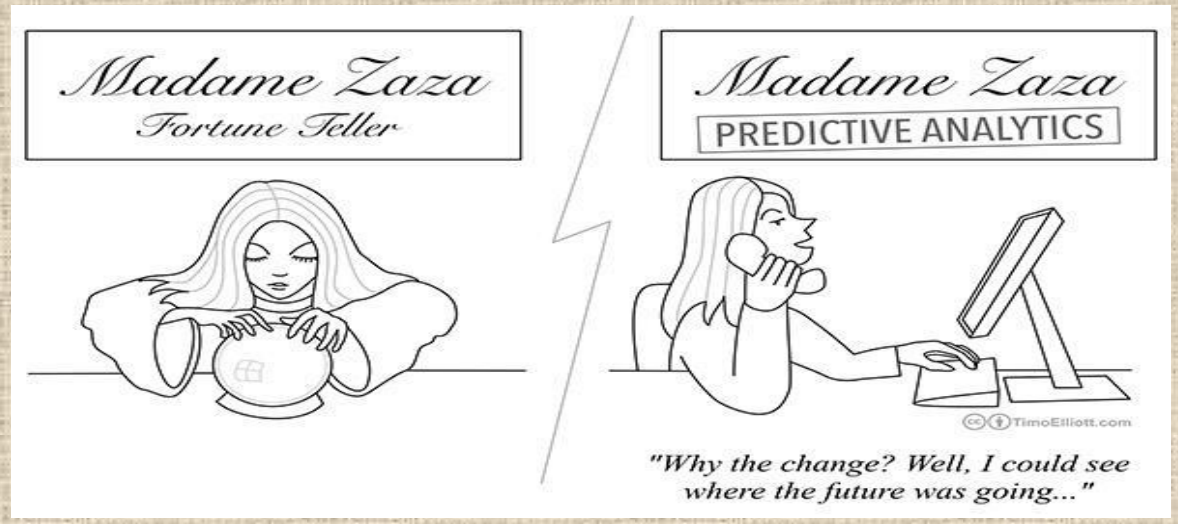


Advantages of Hadoop

defour analytics

- Hadoop framework allows the user to quickly write and test distributed systems. It is efficient, and it automatic distributes the data and work across the machines and in turn, utilizes the underlying parallelism of the CPU cores.
- Hadoop does not rely on hardware to provide fault-tolerance and high availability (FTHA), rather Hadoop library itself has been designed to detect and handle failures at the application layer.
- Servers can be added or removed from the cluster dynamically and Hadoop continues to operate without interruption.
- Another big advantage of Hadoop is that apart from being open source, it is compatible on all the platforms since it is Java based.


CONSULTANTS SAY THREE QUINTILLION BYTES OF DATA ARE CREATED EVERY DAY.
IT COMES FROM EVERYWHERE. IT KNOWS ALL.
ACCORDING TO THE BOOK OF WIKIPEDIA, ITS NAME IS "BIG DATA."
BIG DATA
DilbertCartoonist@gmail.com


Madame Zaza
Fortune Teller
Madame Zaza
PREDICTIVE ANALYTICS
TimoElliott.com
"Why the change? Well, I could see where the future was going..."


OUR BIG DATA ANALYSIS TELLS US THAT ONLY THE TOP PERFORMERS LEAVE FOR HIGHER PAY.
SINCE YOU'RE STILL HERE, IT MEANS YOUR PERFORMANCE IS AVERAGE AT BEST.
THAT'S NOT FAIR!
THAT'S WHAT ALL THE AVERAGE PEOPLE SAY.
DilbertCartoonist@gmail.com
Dilbert.com
1-27-14 ©2014 Scott Adams, Inc. /Dist. by Universal Uclick

defour analytics

Data Visualization

- "...to convey information through visual representations."
- "...produces (interactive) visual representations of abstract data to reinforce human cognition; thus enabling the viewer to gain knowledge about the internal structure of the data and causal relationships in it."
- "…involve detection, measurement, and comparison, and is enhanced via interactive techniques and providing the information from multiple views and with multiple techniques"

Why Visualization?

defour analytics

- Accelerates the identification of hidden patterns in data

"A picture is worth a thousand words"

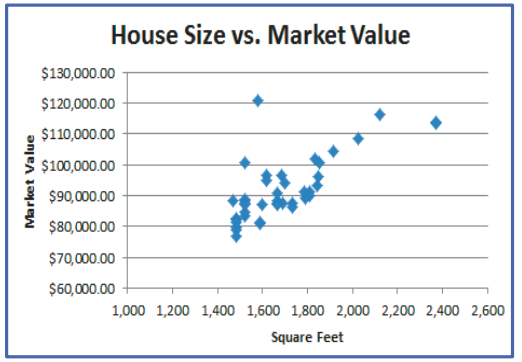
- Represents large quantities of data coherently and efficiently
- The power to visualize and graphically represent results, ideas, solutions, and problems in multiple dimensions, as well as to manipulate data and virtually collaborate with others, is the next big revolution in technology.

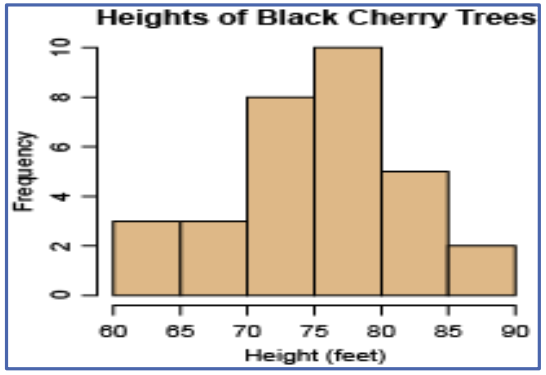
Types of graphs

defour analytics

Scatter Chart

- Suggests correlation between two variables.
- Ability to show nonlinear relationships between variables.



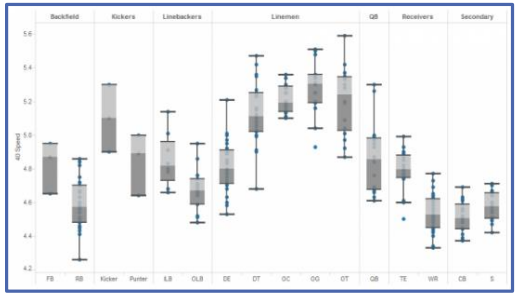


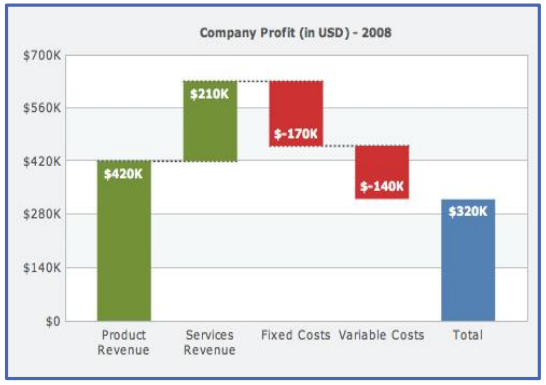
Histogram

- Shows entire distribution of one particular variable.

Box and whisker plot

- Display differences between subpopulations in your data.





Waterfall chart

- A waterfall chart can be used for analytical purposes, especially for understanding or explaining the gradual transition in the quantitative value of an entity which is subjected to increment or decrement.

defour analytics

Dashboards

What's a Dashboard?

A dashboard is a single screen that tracks a number of key metrics in real time.

That may not sound very exciting…

Lets take an example of **Sales and Marketing Dashboard :**

The place where the entire revenue creation team tracks your progress and collaborates around opportunities.

CEO Cockpit!!

Now that's exciting!!

With a well-designed dashboard, you can see:

- If you're attracting enough leads
- How well they're progressing through the sales cycle
- How well you're converting them
- How much revenue is coming your way
- Where the revenue is coming from
- How your campaigns and sales people are performing
- How happy your customers are…

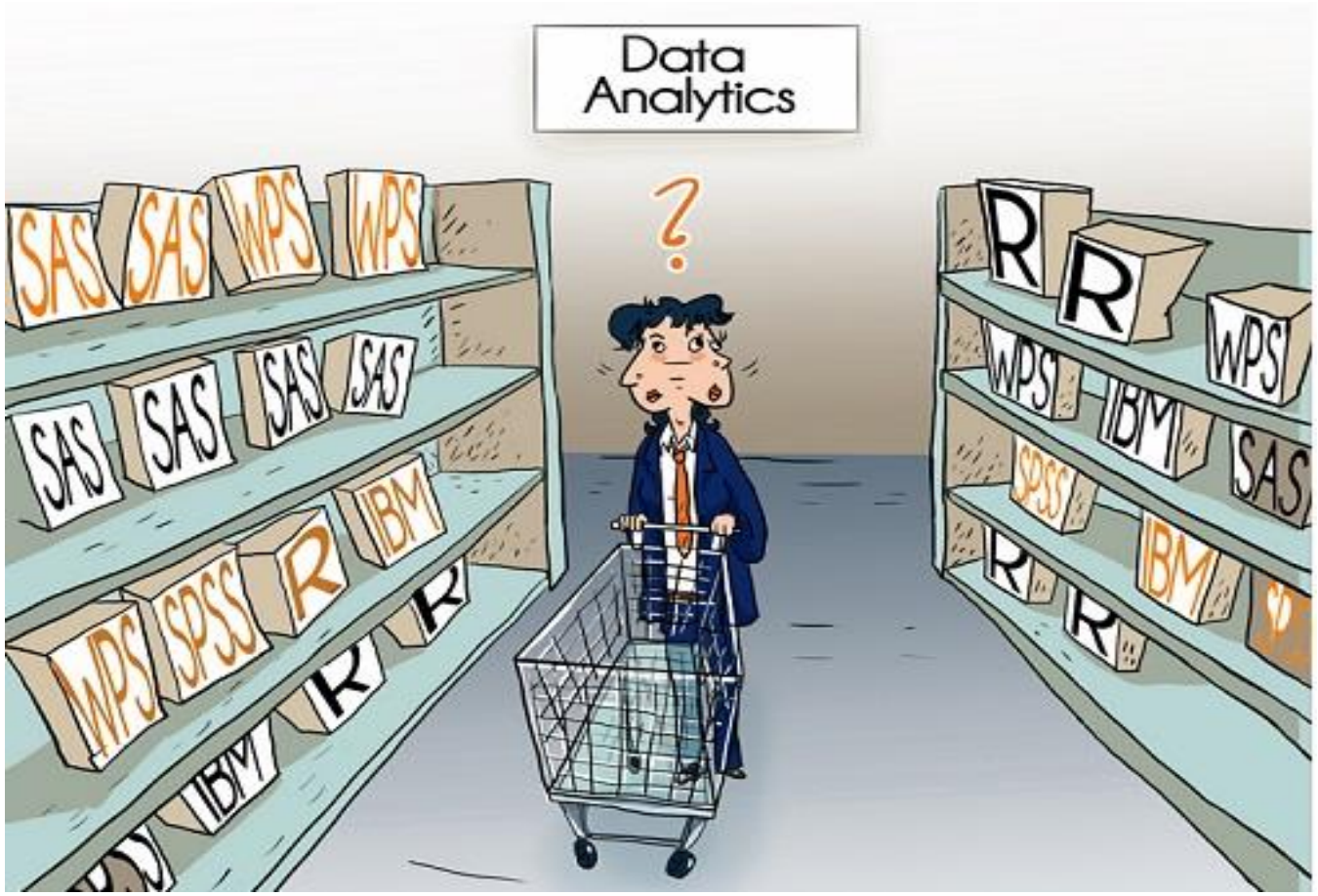
And that's just a small sample.

Why are dashboards so important?

defour analytics

- Dashboards are important because they align your entire sales and marketing organization around the metrics you most need to track.
- They’re open on screens all the time. Everyone refers to them.
- The most important meetings happen around them.
- The critical reports feed into them.
- A great dashboard helps streamline, automate and accelerate the entire sales and marketing operation.

Technology Support



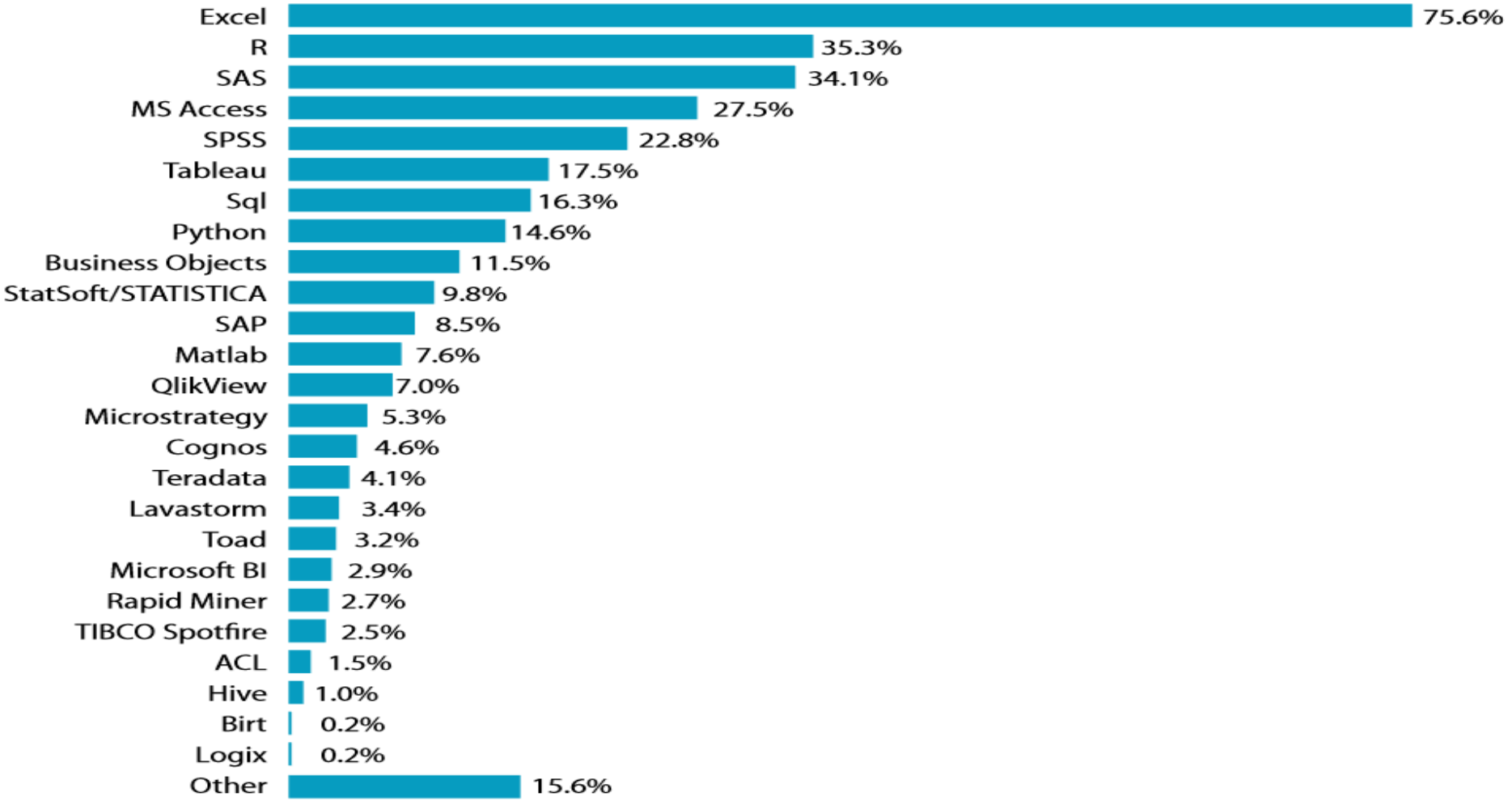
Latest Trends

defour analytics

- Business analytics is a fast growing field and there are many tools available in the market to serve the needs of organizations.
- The range of analytical software goes from relatively simple statistical tools in spreadsheets (ex-MS Excel) to statistical software packages (ex-KXEN, Statistica) to sophisticated business intelligence suites (ex-SAS, Oracle, SAP, IBM among the big players).
- Open source tools like R and Weka are also on the popularity chart.
- There are also some other popular data analytics tools for visualization today, like Tableau, Qlikview, Google maps!

defour analytics

What self-service analytic tool are you currently using?



Tools We Use

defour analytics

R is a data crunching Mammoth which can perform numerous functions and is most widely used!

Alteryx

- Repeatable workflow
- No coding
- Volume and variety of Data
- Quick outputs
- Advanced Analytics

Tableau is popular for its visualization capabilities

Lets Summarize!!

- Business Analytics and its related terms such as Data science, Big Data, Data Mining, etc. has become a powerful tool for companies.
- A Business Analyst is someone who combines analytics knowledge with strong domain expertise and are valuable to businesses because they tend to focus on the role of analytics in the context of a business problem or an opportunity.
- According to the Harvard Review, “Data Scientist is the hottest job of 21st Century”.
- Gartner, a global research firm, predicts that by 2020 nearly 4.4 million new jobs will be created globally by Business Analytics.

defour analytics