



DATA ANALYTICS

Unit 1: Introduction

Mamatha.H.R

Department of Computer Science and
Engineering

DATA ANALYTICS

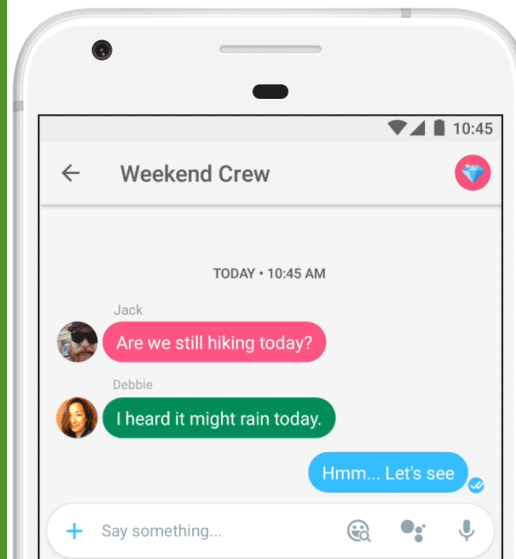
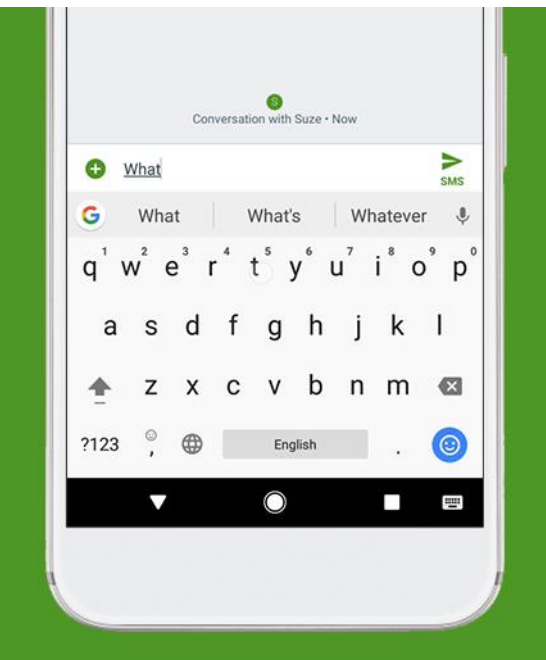
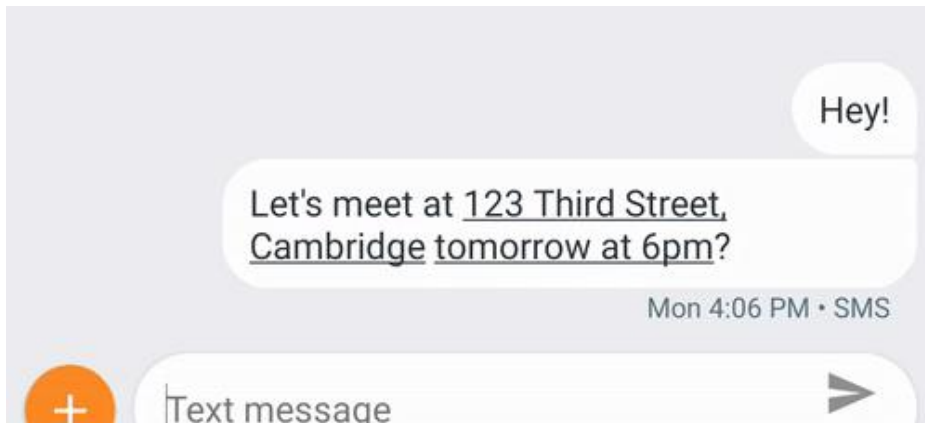
Unit 1:Introduction

Mamatha H R

Department of Computer Science and Engineering

DATA ANALYTICS

What can be done by Data Analytics?



DATA ANALYTICS

What can be done by Data Analytics?



Data can do a lot: sports



DATA ANALYTICS

History



DATA ANALYTICS

What is Data Analytics?

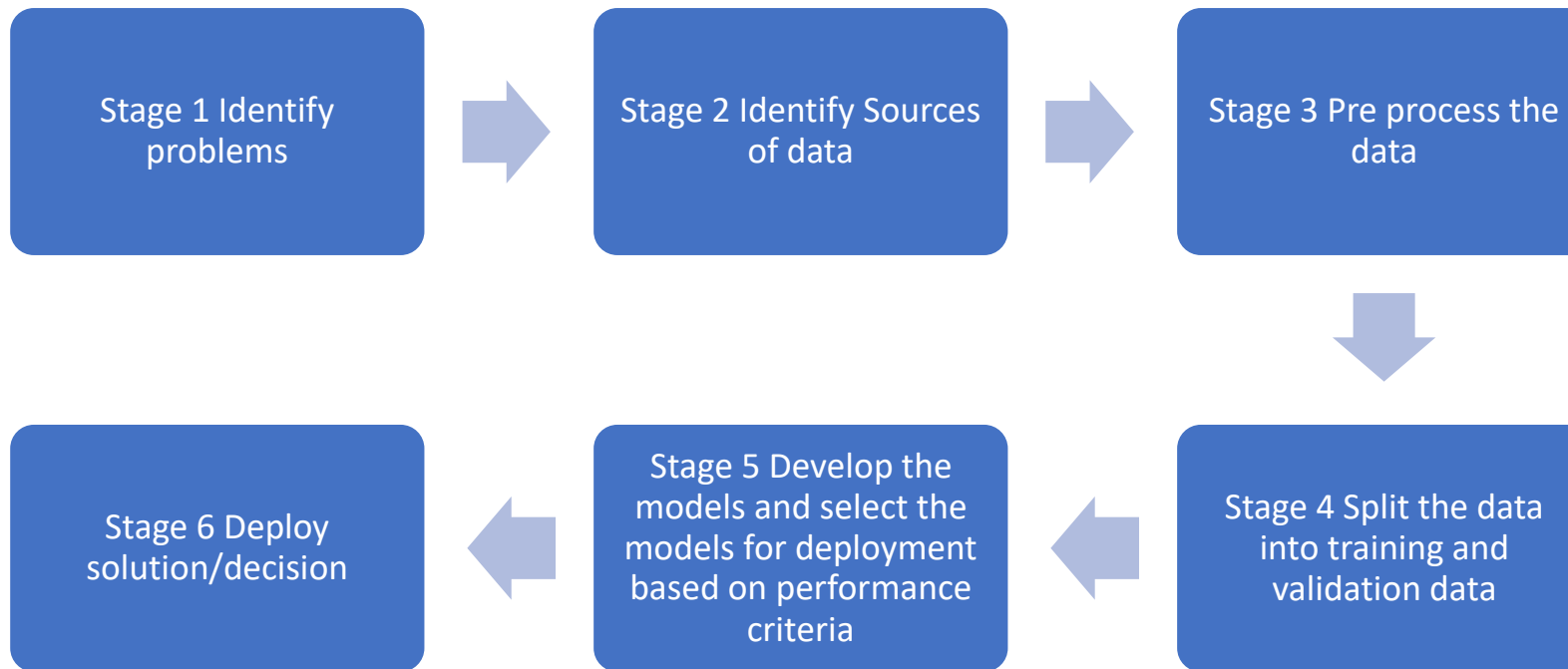
Data analytics (DA) is the process of examining data in order to draw conclusions about the information they contain, increasingly with the aid of specialized systems and software.

<https://searchdatamanagement.techtarget.com/definition/data-analytics>



What is Data Analytics?

Data analytics (DA) is a typical data-driven decision making process

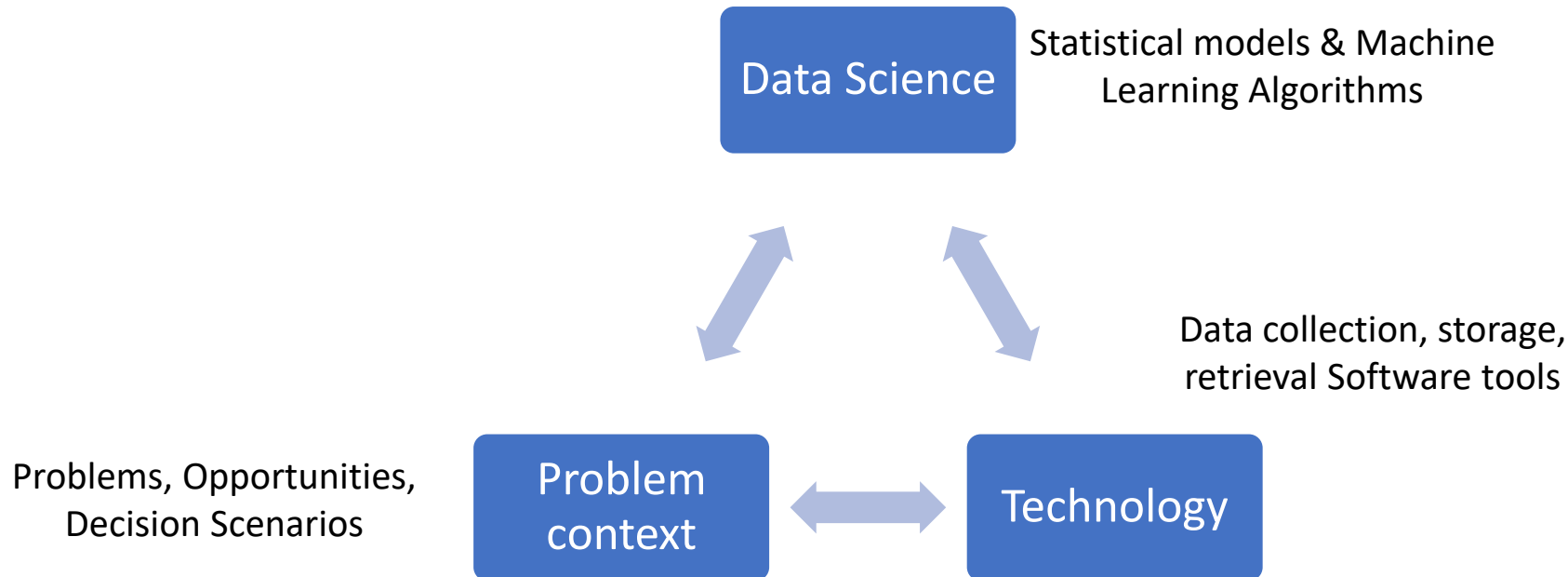


DATA ANALYTICS

What is Data Analytics?

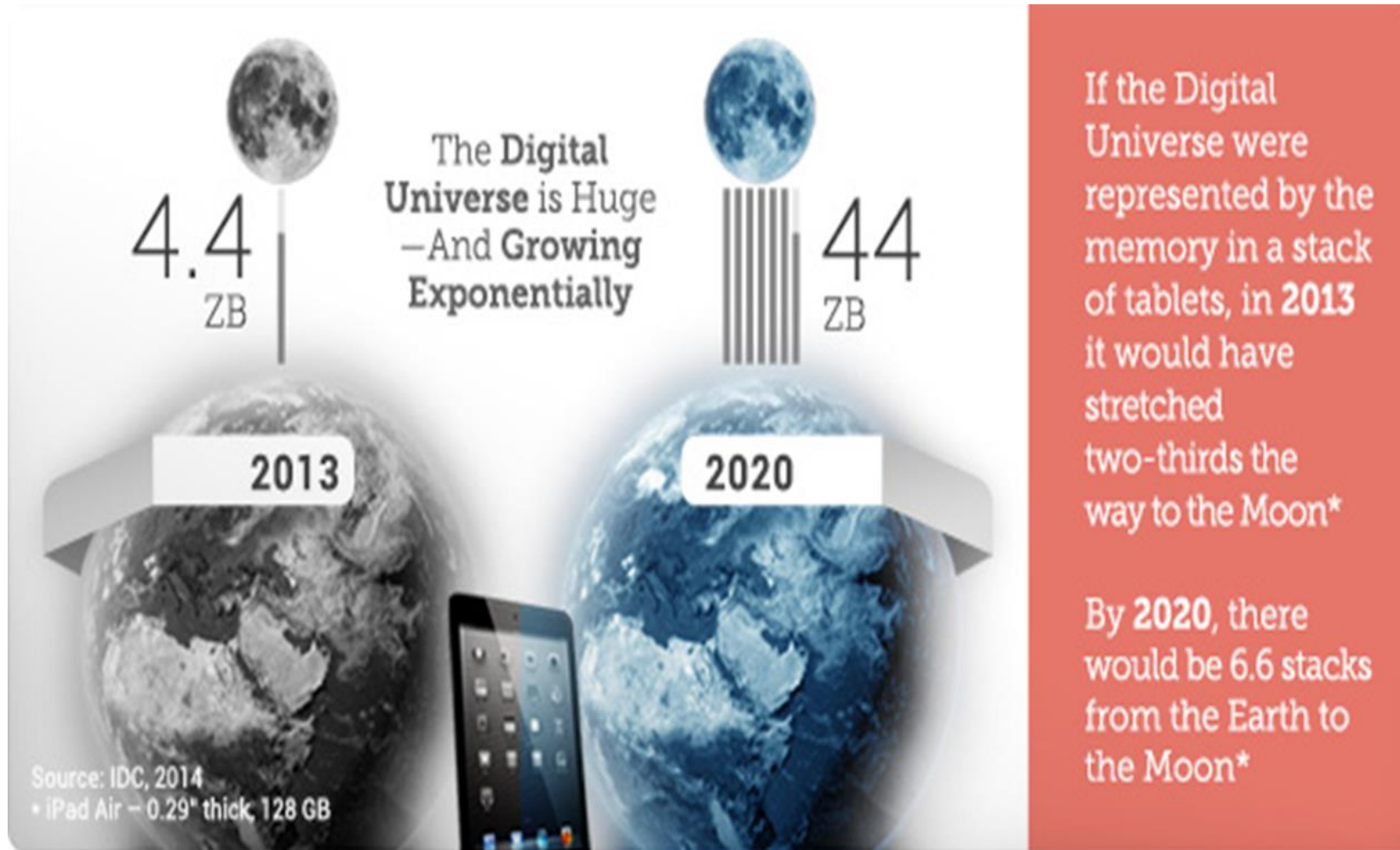
Data analytics is a set of statistical and operation research techniques, artificial intelligence, information technology and management strategies used for framing the problem, collecting data, and analyzing the data to create value to organizations.

It can be broken into three components



DATA ANALYTICS

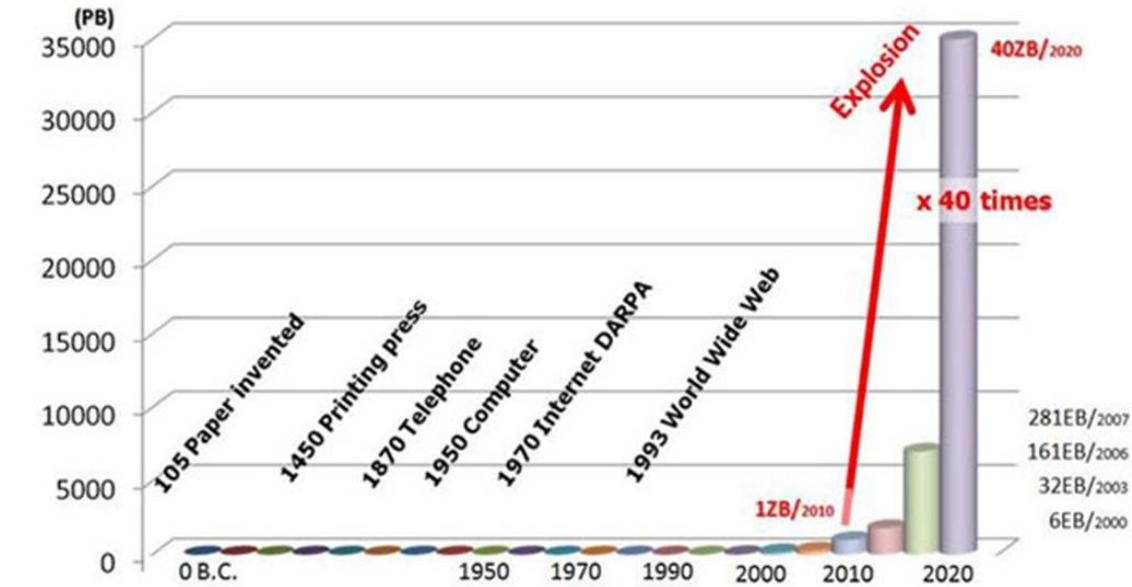
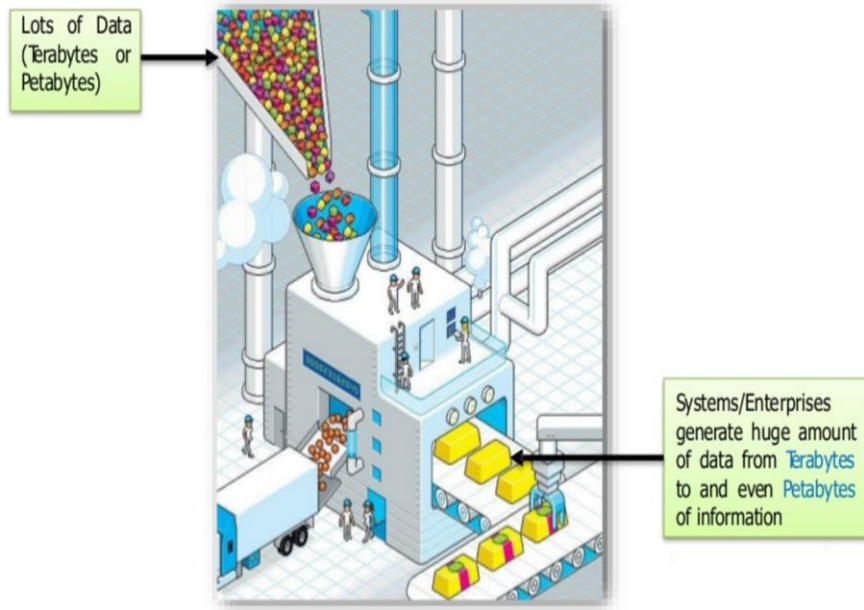
Why Data Analytics?



<https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm>

DATA ANALYTICS

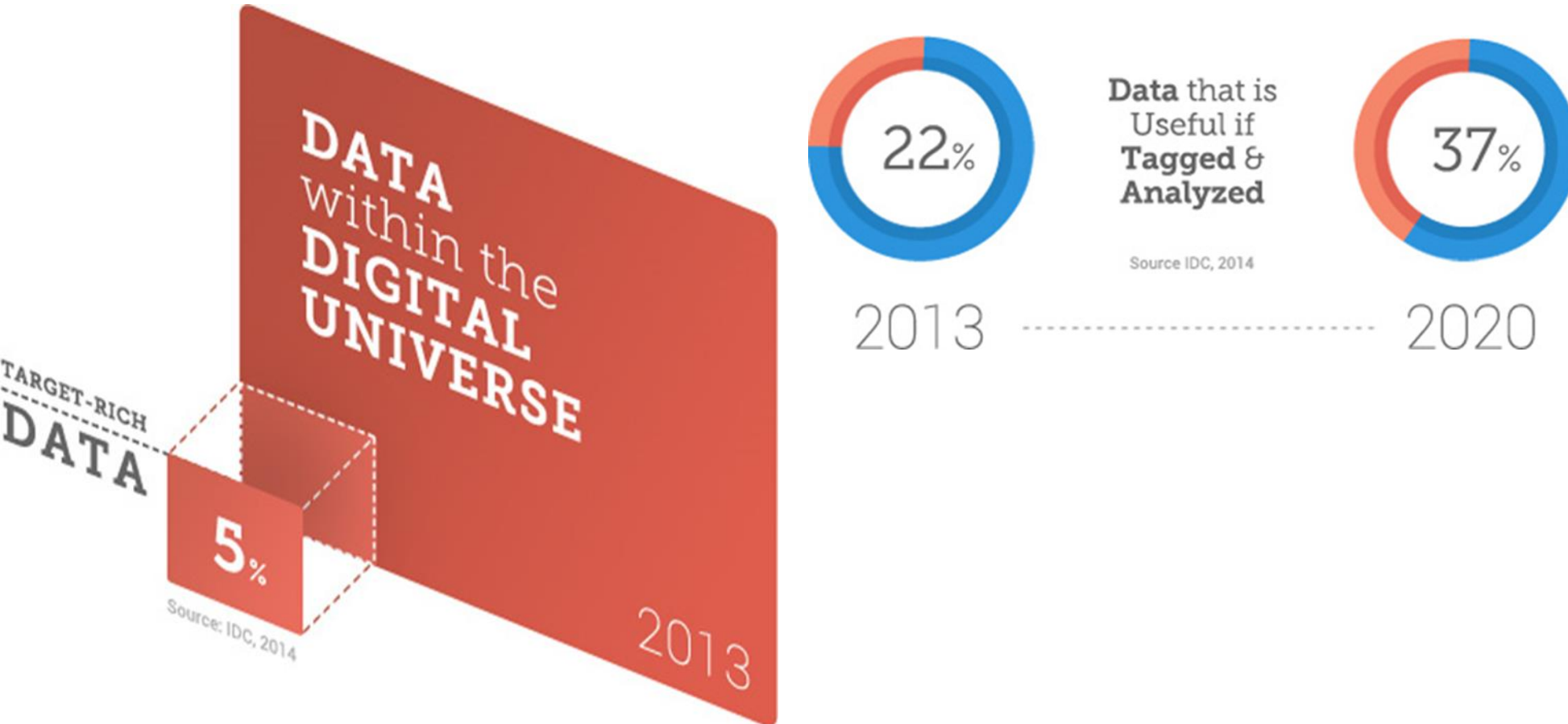
Why Data Analytics?



Source – IDC

DATA ANALYTICS

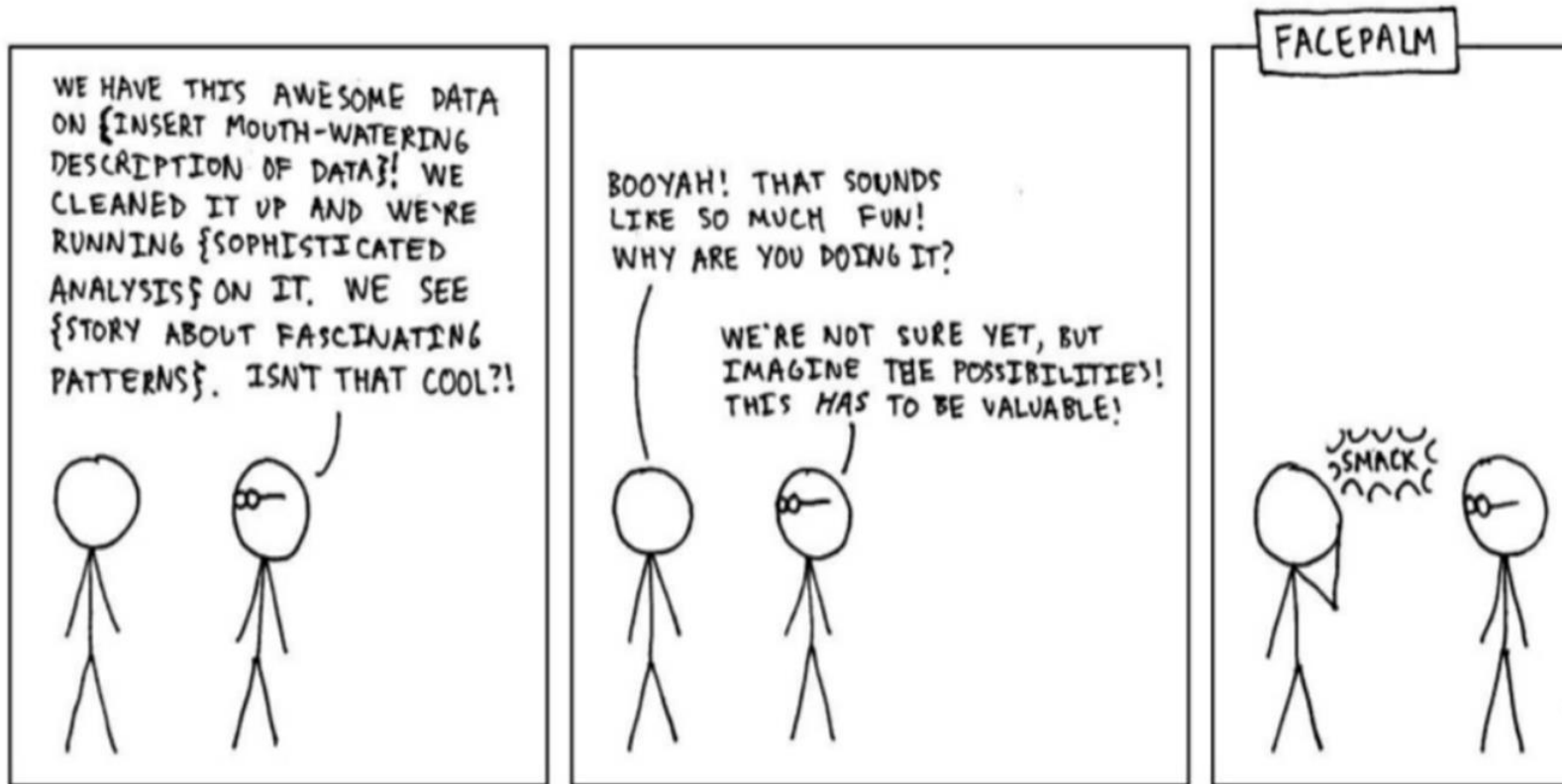
Why Data Analytics?



DATA ANALYTICS

Why Data Analytics?

No one knows how to use it



Why Data Analytics?

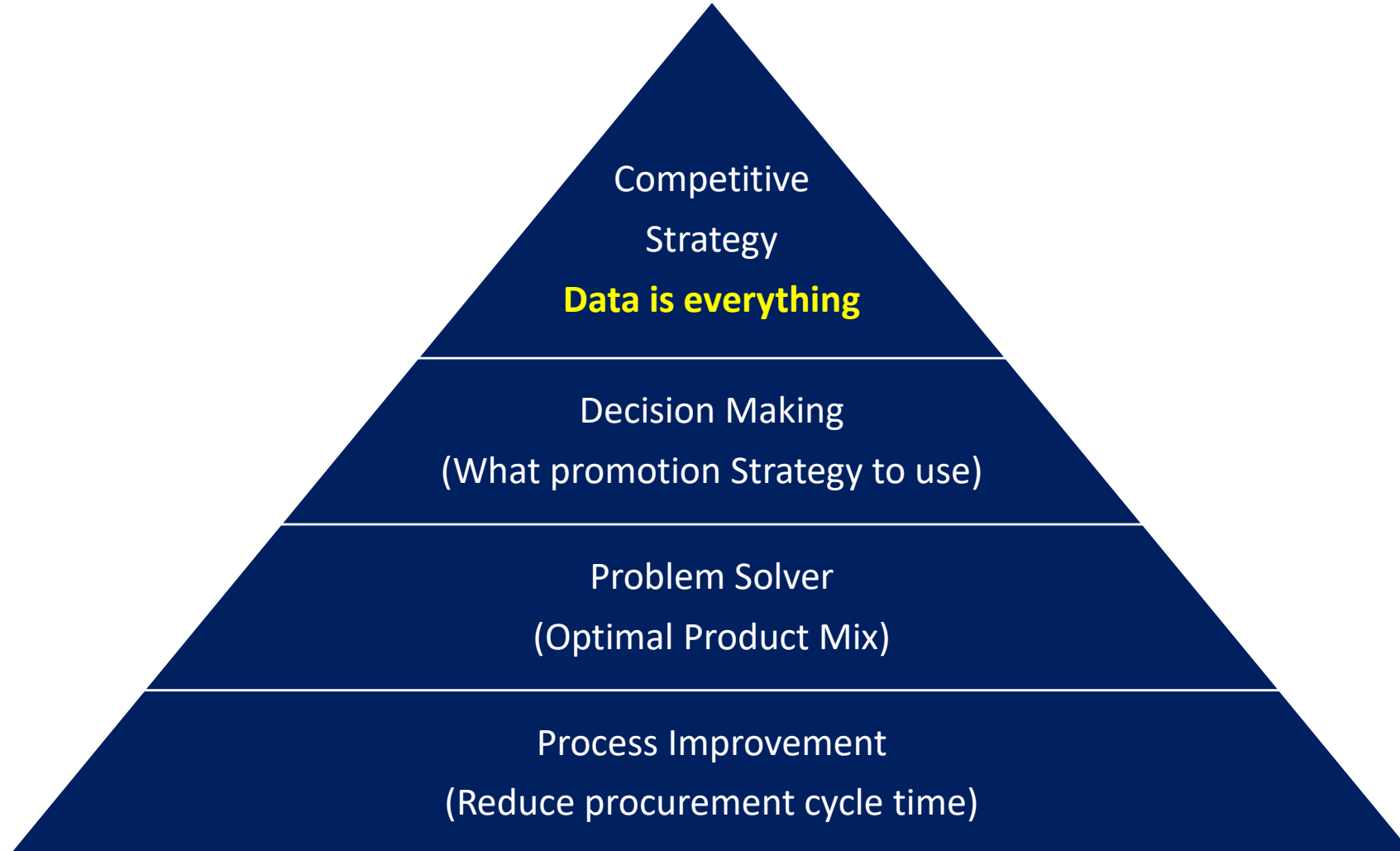
To reduce the cost of decision making

Cost of decision making fall into three categories

Cost of producing a decision: Cost of reaching a decision with the help of a decision maker or procedure

Implementation cost: Cost of actions based on decisions produced

Failure costs: Costs that account for failure of an organization's efforts on production and implementation.



DATA ANALYTICS

Analytics for Process Improvement



- Banking – Cheque clearance time
- Healthcare – Patient discharge time
- Manufacturing – Waste minimization
- Retail – Waiting time at check out counters
- E-commerce – Time to deliver the customer order

- Banking – Reduce non-performing assets, Predict Fraud
- Healthcare – Improve net promoter's score (NPS)
- Manufacturing – Reduce inventory management cost
- Retail – Assortment planning and shelf space allocation
- E-commerce – Predict customer cancellations and Fraud

DATA ANALYTICS

Analytics for Decision Making

- Banking – Loan approval and the interest rate
- Healthcare – Introducing new specialties
- Manufacturing – Whether to introduce a new product
- Retail – Markdown Pricing
- E-commerce – Promotions



Why Data Analytics?



Example: Akshaya Patra Foundation(TAPF)

Mid-day meal programme in South Bangalore

84000 school children

650 schools

35 vehicles

Problem: Vehicle routing –minimize the cost of distribution

1 vehicle-20 schools

Solution space will have $20!(2.4329 \times 10^{18})$

If a computer can evaluate one million routes per second, it would take more than 77146 years to evaluate all possible routes.

For Akshaya Patra, every rupee saved would enable them to add more children to their mid-day meal programme.

DATA ANALYTICS

Analytics is necessary for survival



Problems faced by E-commerce companies such as Amazon and Flipkart

- Forecast demand for each Stock keeping unit(SKU).
- Predict customer cancellations and returns.
- Predict customer contacts at the customer service.
- Predict what a customer is likely to purchase in the future?
- How to optimize the delivery system?

DATA ANALYTICS

The Game Changers...



- **Google**
 - Used Markov chains to rank pages
- **Proctor and Gamble**
 - Analytics as competitive strategy.
- **Target**
 - Predicts customer pregnancy.
- **Capital One**
 - Identifies the most profitable customer.
- **Hewlett Packard**
 - Developed “flight risk score” for 3,30,000 employees.
- **Obama’s 2012 presidential campaign.**
 - Persuasion Modelling.

DATA ANALYTICS

The Innovators...



- **OKCupid**: Predicts which online dating messages is most likely to get a response!
- **Polyphonic HMI**: Uses “hit song science” to predict commercial success of a song.
- **Netflix**: Predicts movie ratings by customers (RMSE is 1%).
- **Amazon.com**: 35% of sales come from product recommendations.
- **Divorce360.com**: Predicting success of [a](#) marriage!

Indian online grocery store bigbasket.com

Problem context driving analytics : "did you forget feature"

The ability to predict the items that a customer may have forgotten to order can have a significant impact on the profits of online grocers such as bigbasket.com

The ability to ask right questions is an important success criteria for analytics projects.

Indian online grocery store bigbasket.com

Technology:

To find out whether a customer has forgotten to place an order for an item

Information technology is used for data capture, data storage ,data preparation, data analysis , data share and to deploy solution.

An important output of analytics is automation of actionable items derived from analytical models which is usually achieved using IT.

Indian online grocery store bigbasket.com

Data Science:

Data science is the most important component of analytics, it consists of statistical and operations research techniques, machine learning and deep learning algorithms.

The objective of the data science component of analytics is to identify the most appropriate statistical model/machine learning algorithm that is best based on a measure of accuracy.

Example: did you forget prediction is a classification problem in which customers are classified into

1. Forget
2. Not forget

DATA ANALYTICS

Case Study

Consider the Target's Pregnancy Prediction feature .
Identify the three components of data analytics - problem context, technology and data science - explain.



Text Book:

“Business Analytics, The Science of Data-Driven Decision Making”, U.
Dinesh Kumar, Wiley 2017

DATA ANALYTICS

Image Courtesy

<https://gfycat.com/smugscratchyiaerismetalmark-acciona-data>

<https://later.com/instagram-marketing/>

<https://www.omnisci.com/technical-glossary/geospatial-analytics>

<http://siteanalystiot.com/>

<https://giphy.com/search/>

<https://www.epam.com/insights/blogs/mapping-the-analytics-continuum-in-life-sciences>





THANK YOU

Dr.Mamatha H R

Professor, Department of Computer Science

mamathahr@pes.edu

+91 80 2672 1983 Extn 834