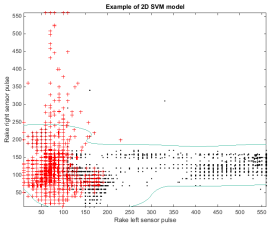
POLITECNICO
MILANO 1863

Quality Data Analysis
aa. 2022-2023

Bianca Maria Colosimo – biancamaria.colosimo@polimi.it

My cv

QUALITY DATA ANALYSIS



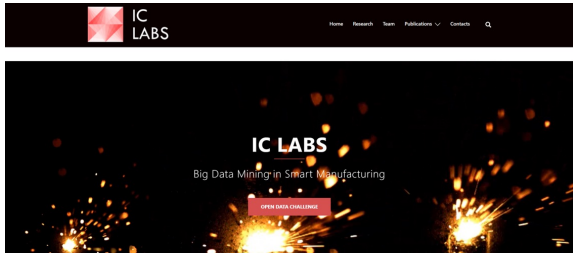
Georgia Tech

PennState

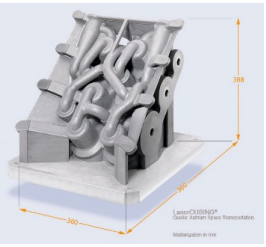
ASQ
The Global Voice of Quality

informs

enbis European Network for Business and Industrial Statistics



MANUFACTURING



SIEMENS

LEONARDO

INGERSOLL Machine Tools

TRUMPF

MANUFUTURE-EU

esa

agenzia spaziale italiana

MIT

ETH Zürich

TUM

Bianca M Colosimo - Polimi

Statistical methods and models for complex data
800 years of research to understand a complex world

1222-2022 800 ANNI

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE STATISTICHE

Bianca M Colosimo

Full Professor

MSc and PhD Politecnico di Milano
Post Phd – Penn State (USA)
Activities & Collaboration – MIT, GA Tech, Chalmers, TU Munich

Lab: Co-founder of the **AddMe Lab**, **3D cell Lab**, **IC Labs**

Senior Editor- Department Editor:
Progress in Additive Manufacturing- Additive Manufacturing Letters
Informs Journal of Data Science – IISE Transactions
Past Editor-in Chief Journal of Quality Technology

Strategy and Roadmapping:
Council members of **Enbis**, **ASQ**, **Informs QSR** and
Member of the European Commission's platform **Manufuture**
SC of the **Vanguard Initiative on 3D Printing** - Board - **EIT Manufacturing**,

Included among the top 100 Italian woman scientists in **STEM** (https://100esperte.it/)

Useful Information

Useful information

1. Timetable:

Data	Dove	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00
Lunedì	BL.27.06			QUALITY DATA ANALYSIS lezione (dal *20/02/2023* al *29/05/2023*)									
Martedì													
Mercoledì	L.13				QUALITY DATA ANALYSIS lezione (dal *22/02/2023* al *31/05/2023*)								
Giovedì	LM.3					QUALITY DATA ANALYSIS lezione (dal *23/02/2023* al *01/06/2023*)							
Venerdì													
Sabato													

2. Slides and data set: https://webeep.polimi.it/

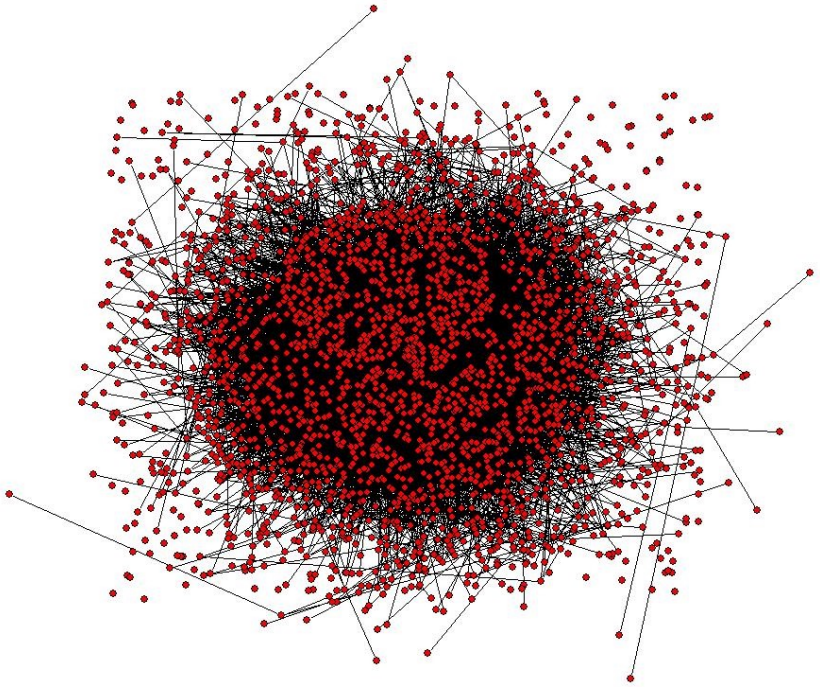
Quality Data Analysis – THE TEAM

Lectures

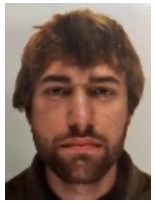
Prof. Bianca Maria Colosimo
(biancamaria.colosimo@polimi.it)
Full Professor, Deputy Head of Dept
Editor of Journal of Quality Technology

Recitations
(esercitazioni)

Matteo Bugatti
(matteo.bugatti@polimi.it)
Assistant Professor



Teaching assistants (lab projects)



Filippo Bracco, PhD Candidate
Patrizia Gironi, PhD Candidate
Giovanni Zanderigo, PhD Candidate

Industrial and scientific connections

bio pros

Our partners

Fraunhofer IPA

Fraunhofer IGB

POLITECNICO MILANO 1863

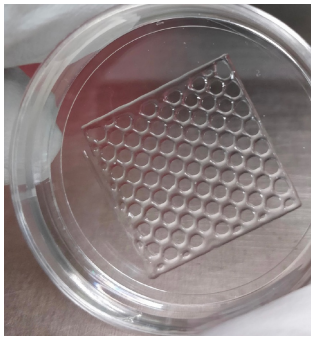
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

NecstGen

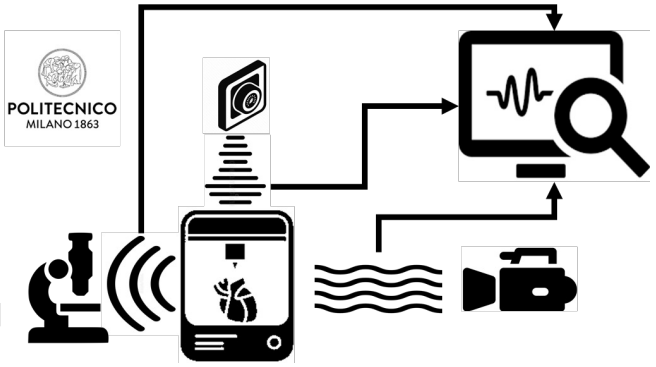
ELVESYS MICROFLUIDICS INNOVATION CENTER

EurA

CELLINK A BICO COMPANY



CARIPLO – POLIMI & HUMANITAS
SMART BIOPRINTING



POLITECNICO MILANO 1863

AVIO

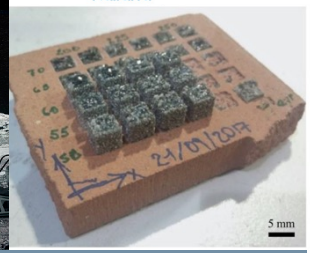
PRIMA INDUSTRIE

POLITECNICO DI TORINO

FONDAZIONE E. AMALDI

POLITECNICO MILANO 1863

asi agenzia spaziale italiana



3D PRINTING OF LUNAR REGOLITH

SCIENTIFIC COOPERATIONS

Georgia Tech

MIT

TUM

Industrial and scientific connections

MICS

Made in Italy Circolare e Sostenibile

PUBLIC RESEARCH ENTITIES

- POLITECNICO MILANO 1863
- Consiglio Nazionale delle Ricerche
- Politecnico di Bari
- Politecnico di Torino
- SAPIENZA UNIVERSITÀ DI ROMA
- UNIVERSITÀ DEGLI STUDI DI BERGAMO
- UNIVERSITÀ DEGLI STUDI DI BRESCIA
- UNIVERSITÀ DEGLI STUDI FIRENZE
- FEDERICO II
- UNIVERSITÀ DEGLI STUDI DI PALERMO
- UNIVERSITÀ DEGLI STUDI DI PADOVA

PRIVATE RESEARCH COMPANIES

- AEFFE
- CAMOZZI GROUP
- CAVANNA
- ITALTEL
- itema
- LEONARDO
- NATUZZI
- Prima Additive
- SACMI
- scm
- ITALIAN LEATHER RESEARCH INSTITUTE
- ThalesAlenia Space

PROJECT KEY FACTS

- 3A-ITALY works on 48% of Value Added of the Italian GPD
- The 12 public entities represent 77% of Italian research capacity on PE11 disciplines
- More than 80 Key Exploitable Results
- Private co-funding: 52.3 M€
- Budget for SMEs and Startups/Spinoffs: 26.5 M€
- Budget for Open Calls: 50.3 M€
- Budget for Equipment, Materials and Licenses: 15.7 M€
- New RTDa positions: 135
- New PhD and research fellows positions: 228
- Endorsement of more than 100 external stakeholders
- Endorsement of more than 20 Industrial Associations
- 50 dissemination and technology transfer roadshow events + Made in Italy Forum

Some of our current research partners (industrial research contracts):

TRUMPF

eMOTION of MOTION.

Avio Aero
A GE Aviation Business

Baker Hughes

INGERSOLL
Machine Tools

Biancamaria.colosimo@polimi.it

Awards

Almost every year – one or more of our students are awarded with premio UCIMU (best thesis award)

Syllabus of the course

Objectives

Nowadays, an impressive amount of data can be collected in real industrial scenarios (Industry 4.0).

The course presents a set of **quantitative tools and methods** for **managing, modeling, monitoring data** in industrial and business scenarios.

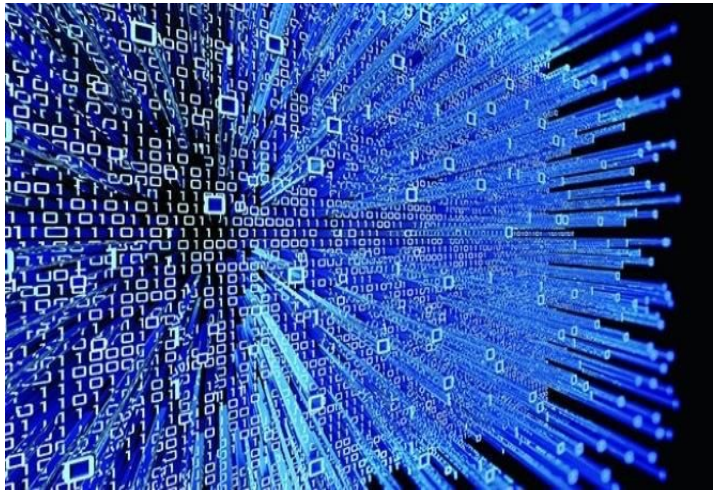
Specific attention is given to **quality** data, i.e., all the key indicators of products and processes which play a relevant role in creating added value for the company.

Target

After successfully completing the course, students should be able to do the following:

1. Understand the philosophy and basic concepts of **quality data modelling**, monitoring and improvement.
2. **Extract relevant information from complex, high-dimensional data**
3. Identify **models to predict the expected pattern** of quality and performance indicators
4. Design and use **appropriate tools** to design and manage data in industrial and service scenarios.

Expected Results



Lectures will provide the basic tools to understand functions and processes, using appropriate quantitative tools for data analysis.

Recitations in computer labs will show how the learned tools can be effectively used to design new solutions using a scientific approach to face the problems at hand (applying knowledge and understanding).

The **lab project** will foster an additional insight to **develop** new ideas and solutions in business and industrial scenarios (making judgements and learning skills).

Requirements

A five-credit course in Statistics is strongly suggested/required.

Students should mainly know:
- basics of statistical distributions (normal, poisson, binomial, Chi-square, F)
- Confidence intervals and Hypothesis tests

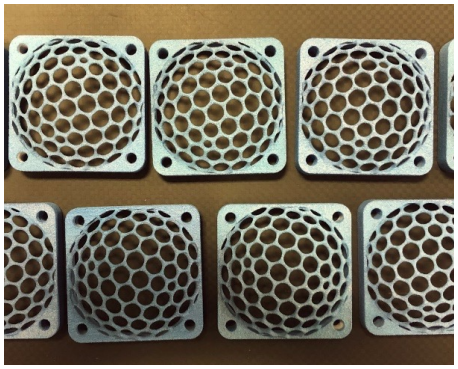
Project activity

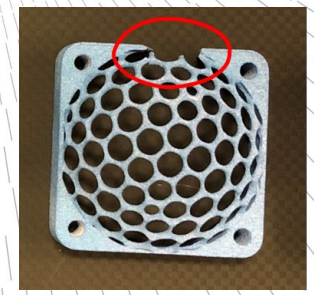
MADE - Competence Center Industria 4.0 | Milano, Italia
https://www.made-cc.eu

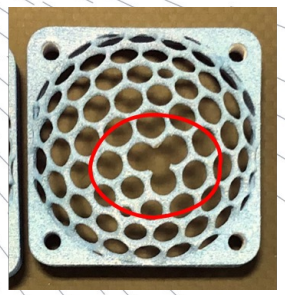
In-line vision-based quality analysis for automated defect detection

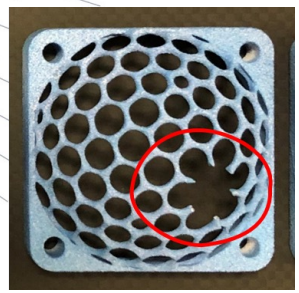
(in collaboration with Rockwell Automation and Siemens)

Test case: 3D printed fan covers (with HP Multijet-Fusion)









Problem: design and test your own statistical method to detect quality errors and anomalies affecting the 3D printed parts based on in-line gathered images

Topics

1.Quality engineering & Industry 4.0: data analysis as a basic tool for modeling, monitoring, control and improve.
2.Quality data modeling:
- Standard assumptions and related tests;
- Modeling patterns via linear models;
- Modeling autocorrelated data via time series analysis;
- Modeling survey data: Categorical and ordinal data

3.Quality monitoring of continuous variables
- Traditional statistical process control (SPC) for the mean and the variance
- SPC for autocorrelated data: Problems of traditional control charts for autocorrelated data;
- Model based and model-free approaches for quality control of autocorrelated data.

4.Quality modelling and monitoring for "big" data streams: multivariate data
- Modeling multivariate data
- Dimensional reduction via Principal Component Analysis
- Control chart for multivariate data - controlling the mean and the covariance

Mech. engineers (8CFU)
Final week around May 15- 20

5. Toward zero-defect manufacturing: process quality and product specifications. Capability analysis. Univariate and multivariate control charts for small shifts (EWMA, CUSUM).
6. Modelling and monitoring attribute and qualitative data: control chart for defective rates and survey data
7. Quality Improvement – The role of improvement in the six-sigma roadmap. Quality improvement via empirical model building (for management engineering only- hints).

All the others:
Management engineering & food engineering (10 CFU)

Exam

Textbooks:

D.C. Montgomery: Introduction to Statistical Quality Control – Wiley

Additional textbooks:

- L. C. Alwan "Statistical Process Analysis" – Irwin Mc Graw Hill
- E. del Castillo: "Statistical Process Adjustment for Quality Control" – Wiley

Additional material will be provided

Written exam using appropriate SW (minitab, excel, …)

Project work – optional but highly suggested

– The project score sums up to the exam score only if written exam >=18)
– maximum increase of the final grade 3/30
– groups made of min 1, max 4 students
– The project scores are valid only for the sessions of this academic year

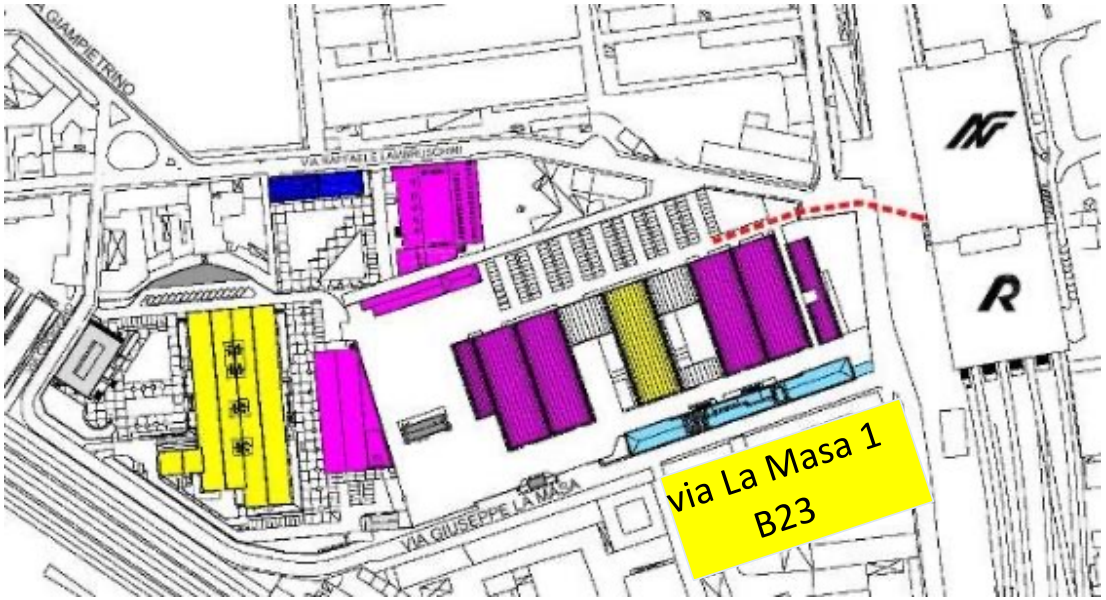
Oral exam – optional (only if written exam >=18)

– maximum increase of the final grade 3/30

Contacts

Bianca Maria Colosimo
Dipartimento di Meccanica
(Sesini Building, B23, via La Masa 1)
Tel.: 02-2399.8522 E-mail:
biancamaria.colosimo@polimi.it
Monday 14:00-15:30
(please send an email to book an appointment)

Matteo Bugatti
Dipartimento di Meccanica
(Sesini Building, B23, via La Masa 1)
E-mail: matteo.bugatti@polimi.it



Recitations on Monday

We kindly ask you to:

- Bring your laptop computer (with electrical socket if required)
- We will guide you through the process to download and install Python

Instructions will also be uploaded to the Beep website (pay attention that a specific procedure is needed to install the sw on Mac - information are available on beep)

Introduction to the course

More practically, what is this course about?

This course deals with **Quality Data Analysis**
with special attention to
Quality data modeling monitoring and control

It trains engineers to handle and design ***quantitative approach*** for
industrial data "science"

Quality Data Analysis

Quality

capability to satisfy (expressed and unexpressed) customer needs - Deming

All the information affecting the customer satisfaction

Data analysis – modeling and monitoring

Industry 4.0 Digital Transformation

All the approaches that we will study can be applicable to other data types, too

Twin transition: green and digital

Green

- Zero-defect first-time right versus zero waste:
  - reduced time to market,
  - Material savings
  - Resource savings (energy)
  - Reduced scrapes and rework

Digital

- Massive data
- Multi-stream
- high velocity
- Big data

- Fast reaction
- Decision making

- (Edge computing
- Distributed data)

I4.0 in EU

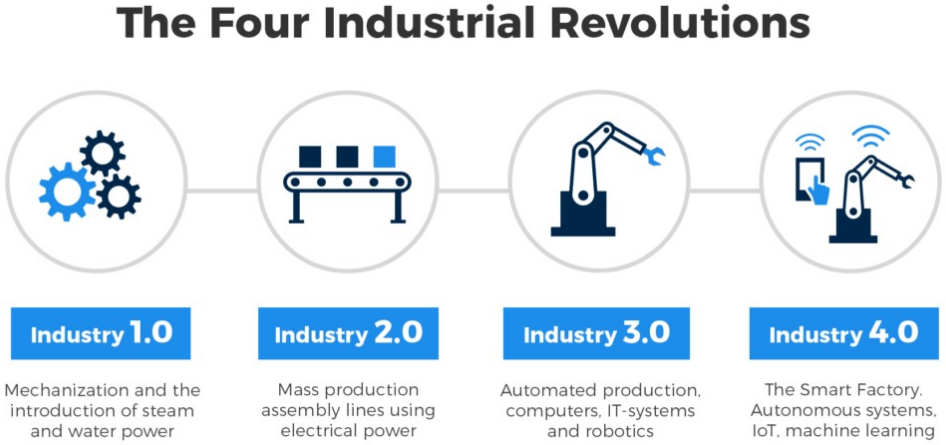
Ursula von der Leyen, President of the European Commission, said: “Europe's industry is the motor of growth and prosperity in Europe. And it is at its best when it draws on what makes it strong: its people and their ideas, talents, diversity and entrepreneurial spirit.

This is more important than ever as Europe embarks on its ambitious **green and digital transitions** in a more unsettled and unpredictable world. Europe's industry has everything it takes to lead the way and we will do everything we can to support it.”

Thierry Breton, Commissioner for Internal Market “Europe has the strongest industry in the world. Our companies - big and small - provide us with jobs, prosperity and strategic autonomy. Managing the **green and digital transitions** and avoiding external dependencies in a new geopolitical context requires radical change - and it needs to start now.”

Industry 4.0

Industry 4.0 Is Enabling A New Era Of Manufacturing Intelligence And Analytics



- 72% of manufacturing enterprises predict their use of data analytics will substantially improve customer relationships and customer intelligence along the product life cycle.
- 86% of manufacturers surveyed expect to secure simultaneous gains from both lower costs and added revenue in the next five years.
- 35% of companies adopting Industry 4.0 expect revenue gains over 20% over the next five years.
- Japan and Germany are the furthest along in digitizing internal operations and partnering across their value chains.
- Data analytics and digital trust are the foundations of Industry 4.0.

Forbes August 7th, 2017

Big data - Volume

From Harvard Business Review, 2012:

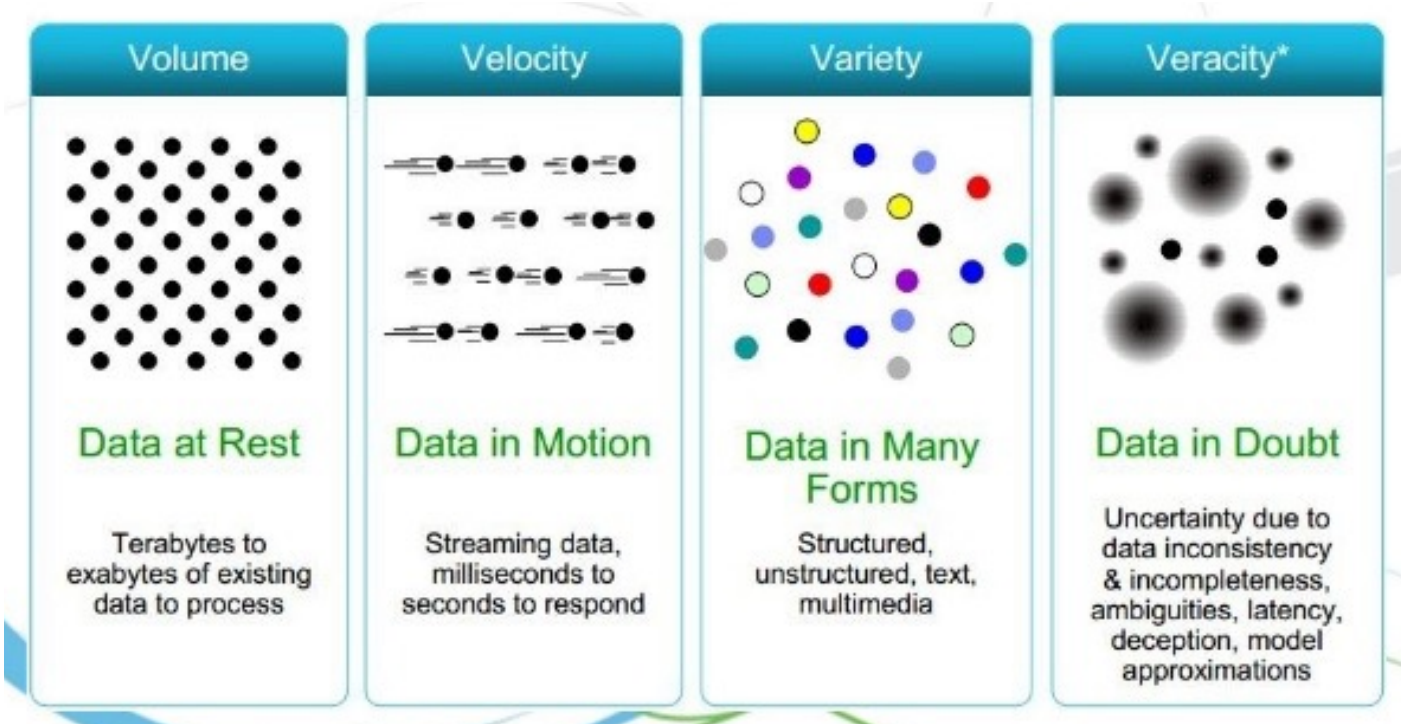
The shortage of data scientists is becoming a serious constraint in some sectors.

DATA

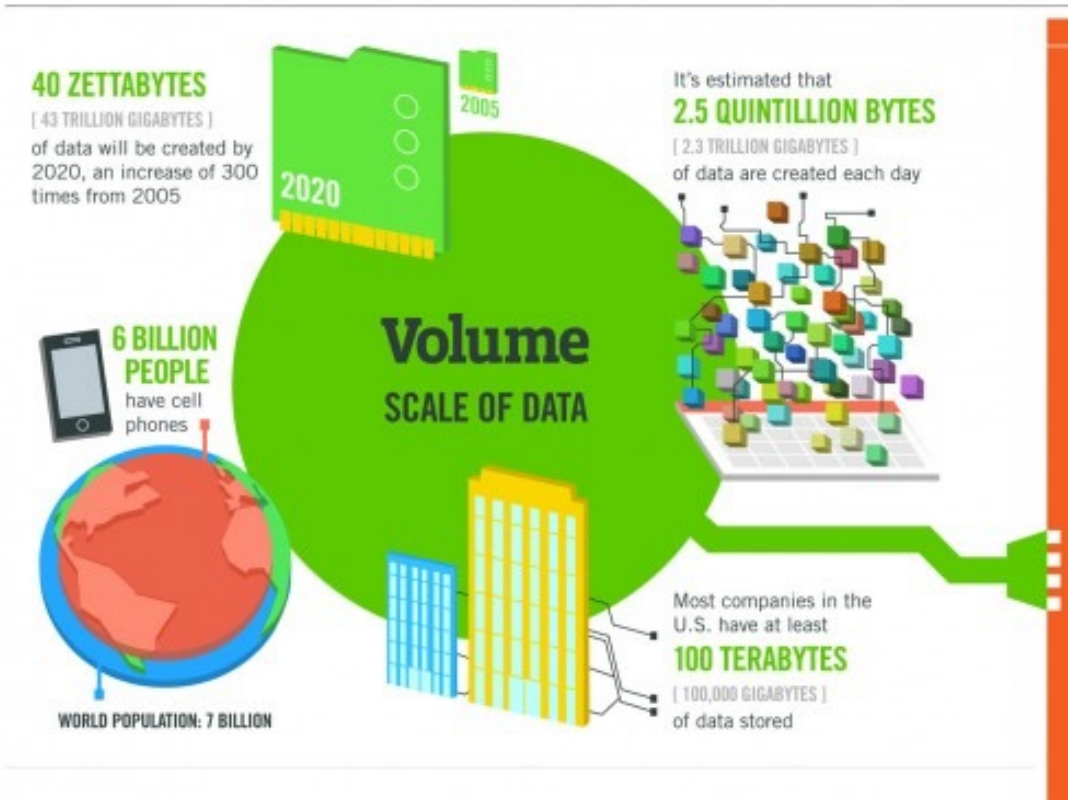
Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE



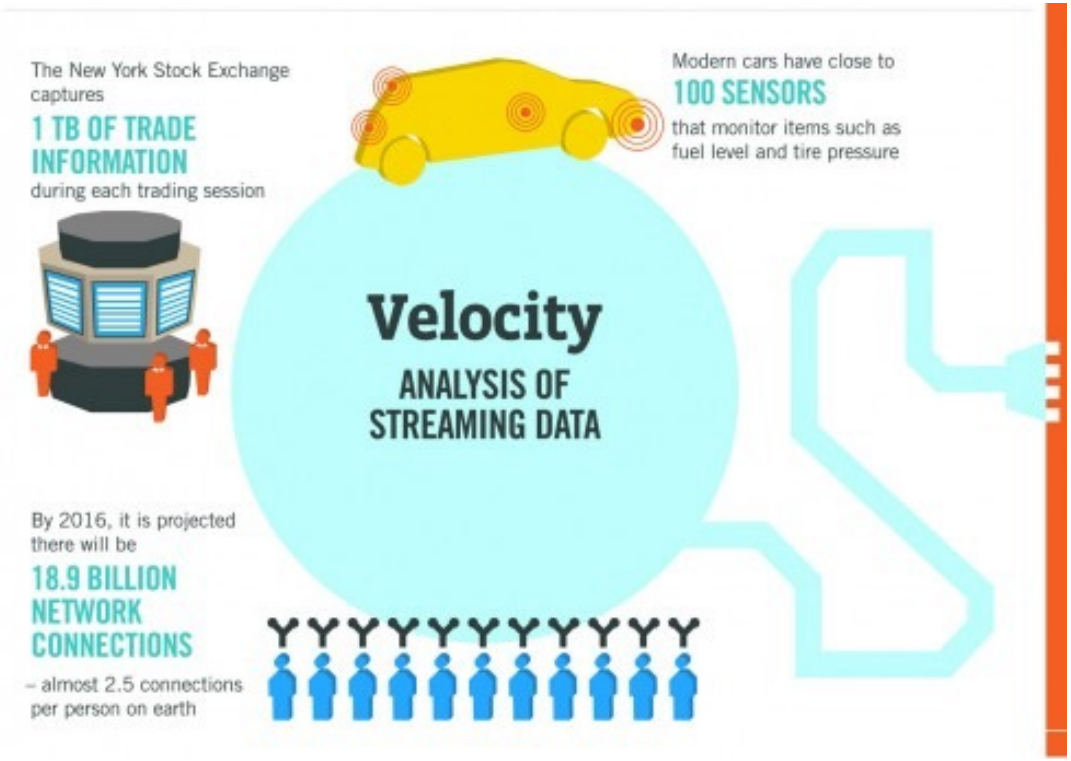
Data in the modern age



Volume refers to the **vast amounts of data** generated every second. We are not talking Terabytes but Zettabytes or Brontobytes. If we take all the data generated in the world between the beginning of time and 2008, the same amount of data will soon be generated every minute! This increasingly makes data sets too large to store and analyse using traditional database technology. With big data technology we can now store and use these data sets with the help of distributed systems, where parts of the data is stored in different locations and brought together by software.

Big data - https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know/

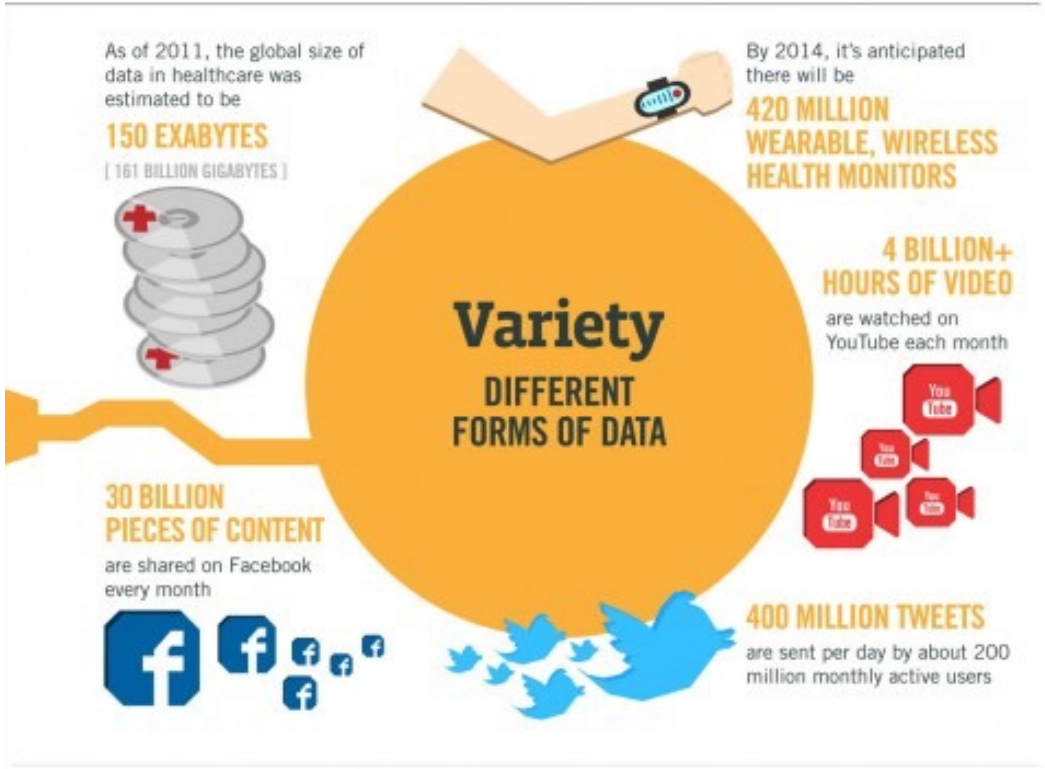
Data in the modern age



Velocity refers to the **speed** at which new data is generated and the speed at which data moves around. Just think of social media messages going viral in seconds, the speed at which credit card transactions are checked for fraudulent activities, or the milliseconds it takes trading systems to analyse social media networks to pick up signals that trigger decisions to buy or sell shares. Big data technology allows us now to analyse the data while it is being generated, without ever putting it into databases.

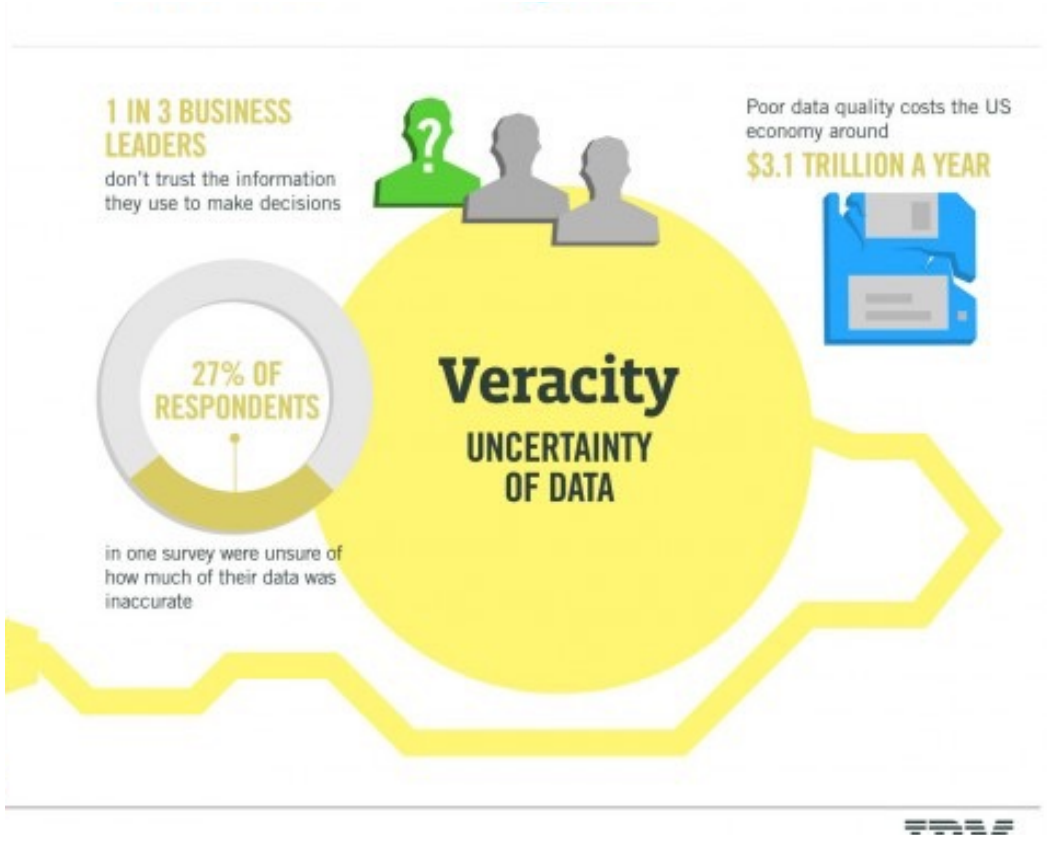
Big data - https://www.linkedin.com/pulse/20140306073407-64875646-big-data-the-5-vs-everyone-must-know/

Data in the modern age



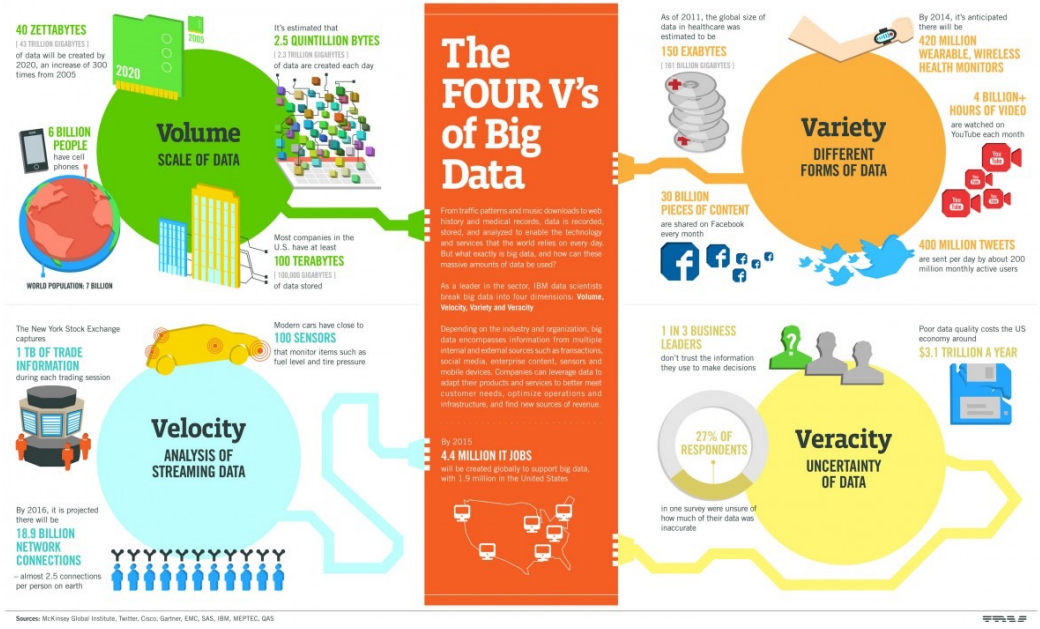
Variety refers to the **different types of data** we can now use. In the past we focused on structured data that neatly fits into tables or relational databases, such as financial data (e.g. sales by product or region). In fact, 80% of the world's data is now unstructured, and therefore can't easily be put into tables (think of photos, video sequences or social media updates). With big data technology we can now harness differed types of data (**structured and unstructured**) including messages, social media conversations, photos, sensor data, video or voice recordings and bring them together with more traditional, structured data.

Data in the modern age



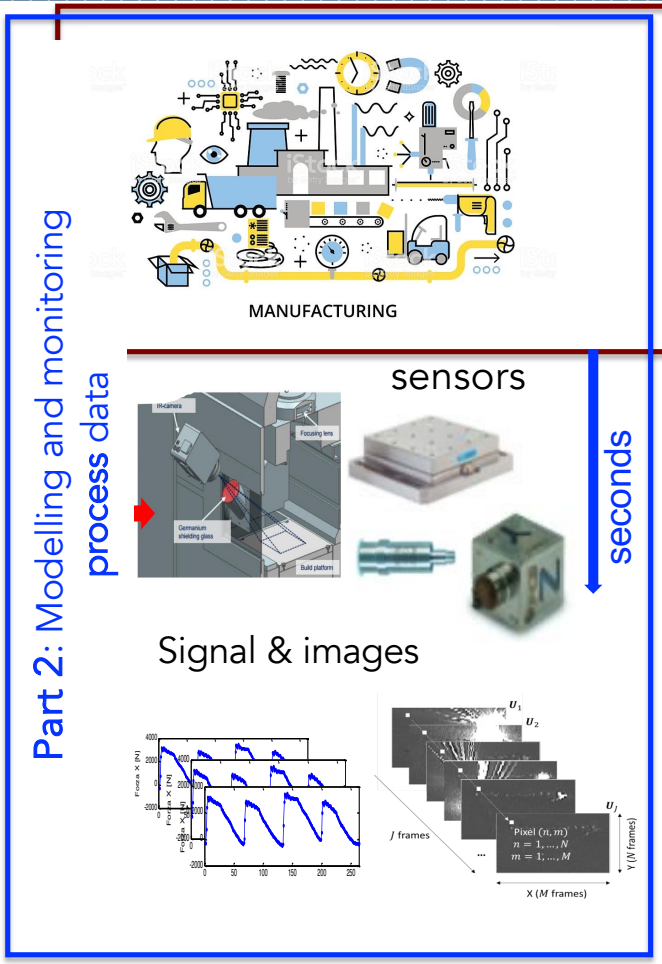
Veracity refers to the messiness or trustworthiness of the data. With many forms of big data, quality and accuracy are less controllable (just think of Twitter posts with hash tags, abbreviations, typos and colloquial speech as well as the reliability and accuracy of content) but big data and analytics technology now allows us to work with these type of data. The volumes often make up for the lack of quality or accuracy.

Data in the modern age



Value: Then there is another V to take into account when looking at Big Data: Value! It is all well and good having access to big data but unless we can turn it into value it is useless. So you can safely argue that 'value' is the most important V of Big Data. It is important that businesses make a business case for any attempt to collect and leverage big data. It is so easy to fall into the buzz trap and embark on big data initiatives without a clear understanding of costs and benefits.

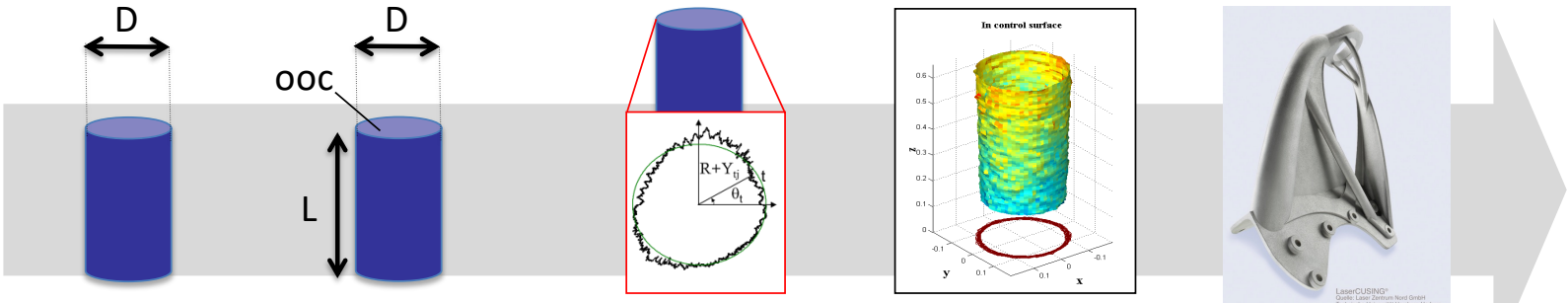
Quality Data Analysis



We are assisting to a paradigm shift in product and process data

- from simple to complex data (image, point clouds, voxels, network data)
- from normal size to massive data
- from slow to fast data
- from single to multiple streams

Product data



$X = D \in \mathbb{R}$

One quality characteristic

X=[D,L,OOC,…]'

$\mathbf{X} \in \mathbb{R}^p$

p quality characteristics

$X = f(\theta) + \varepsilon$

Quality profiles

$X = f(\theta, z) + \varepsilon$

Quality Surfaces

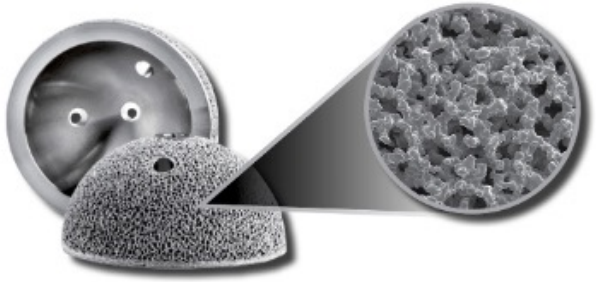
?

Complex shapes

Lightweight structures

Complex shapes (topological optimization)

Internal channels and features

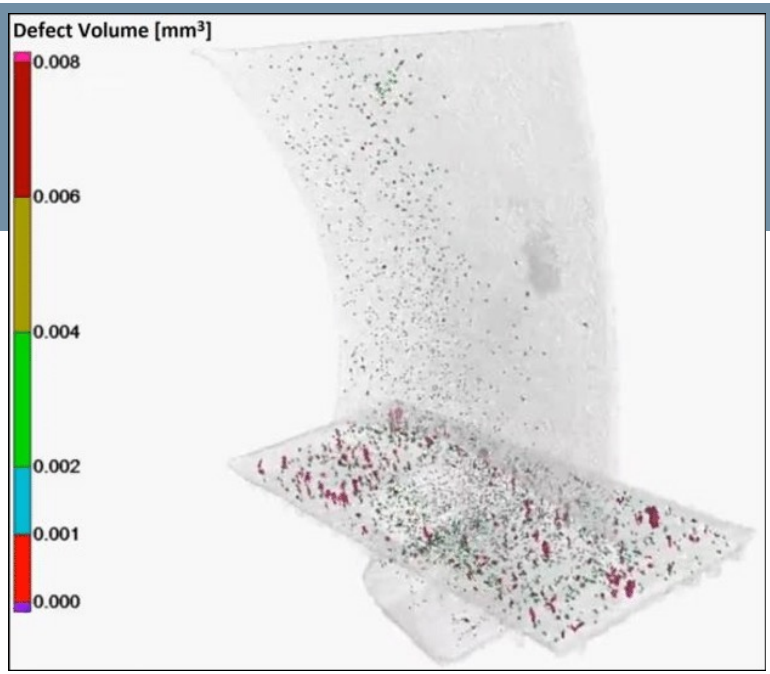


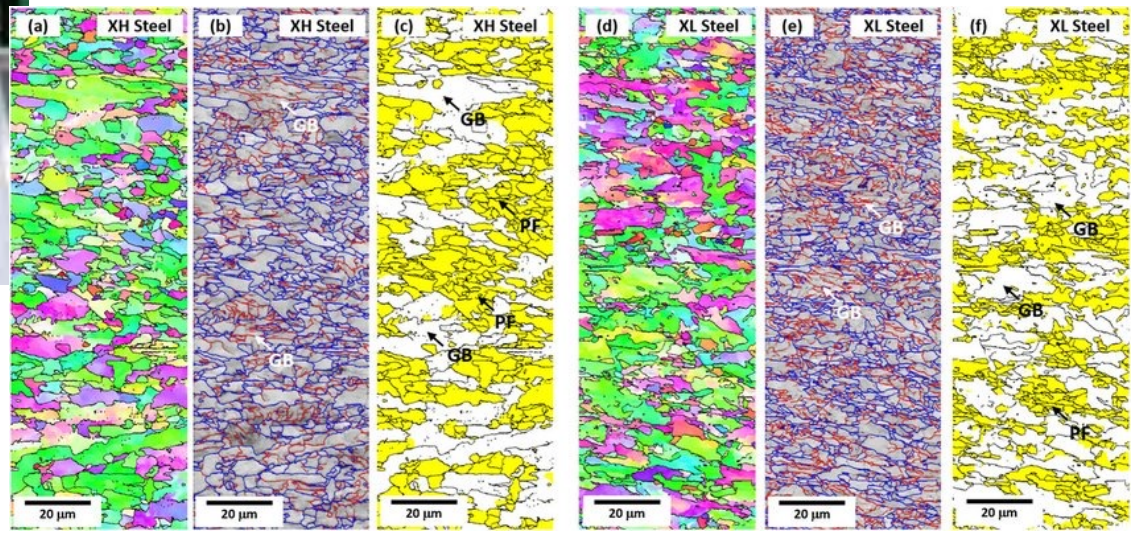
Surface and functional patterns

Airbus A380 bracket

Source: Concept

Product data

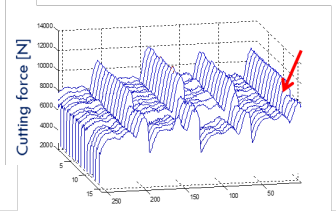




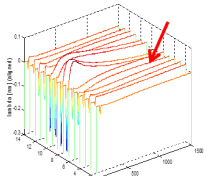
Data 4.0

signals

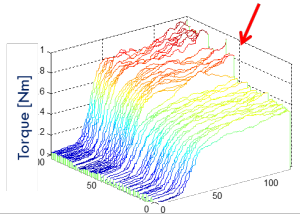
Milling hard-to-cut materials



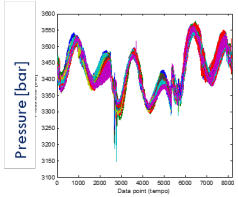
Drilling of multi-layer materials (aerospace)



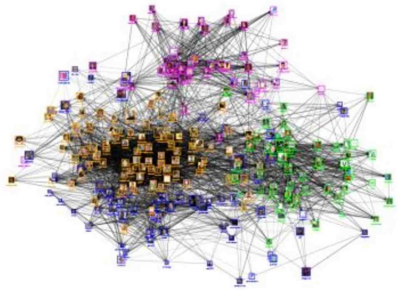
Tapping monitoring



Waterjet and Abrasive Waterjet cutting



Network data

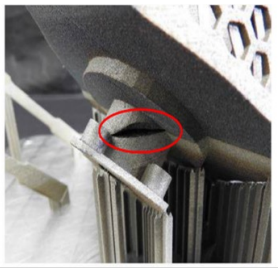


Social network data

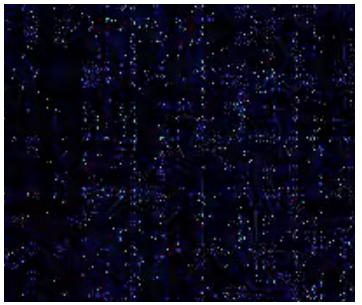
Images /video

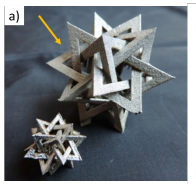
POST-SCAN

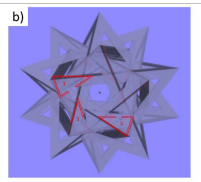
DEFECT

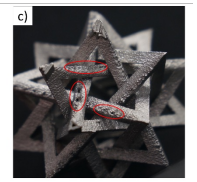


Red line: nominal contour
Blue area: detected departure between in-situ reconstruction and nominal contour









Why data analysis is so important?

✓Having data is not enough if we don't know how to let them speak.

✓"Data analysis" does not mean "Complex analysis"

✓Very often data are not appropriately analyzed because appropriate tools and methods are not known.

"One important idea is that science is a means whereby learning is achieved, not by mere theoretical speculation on the one hand, nor by the undirected accumulation of practical facts on the other, but rather by a motivated iteration between theory and practice."

George E. P. Box (1976)

✓"*Essentially, all models are wrong, but some are useful.*" Box, G. E. P., and Draper, N. R., (1987)

Quality Data Analysis

Quality

capability to satisfy (expressed and unexpressed) customer needs - Deming

All the information affecting the customer satisfaction

Data analysis – modeling and monitoring

Industry 4.0 Digital Transformation

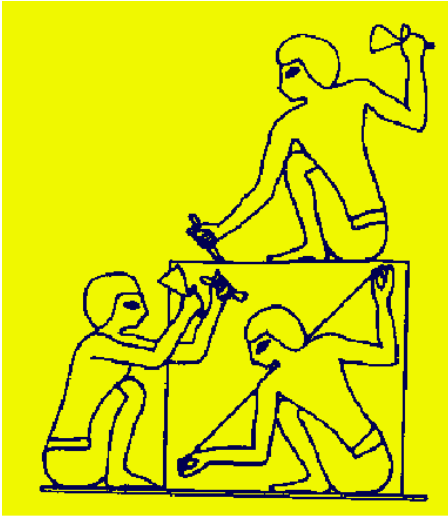
All the approaches that we will study can be applicable to other data types, too

Quality & Customer satisfaction

> Quality—you know what it is, yet you don't know what it is. But that's self-contradictory. But some things are better than others, that is, they have more quality. But when you try to say what the quality is, apart from the things that have it, it all goes poof! There's nothing to talk about. But if you can't say what Quality is, how do you know what it is, or how do you know that it even exists? If no one knows what it is, then for all practical purposes it doesn't exist at all. But for all practical purposes it really does exist.

Quality: a never ending problem old story

XX b.C.
"When standards were repeatedly violated, *Phoenicians inspectors* cut the hands of people who did the work."



1450 b.C. Egyptian tomb: with the help of a chord, the inspector is checking whether the stone faces are perpendicular or not

50 a.C. roman water main (aqueduct): built to support the wind at a speed of 215 Km/h

(max wind speed measured at that time 100 Km/h)

standard were the same for all the regions in the empire

Quality and its relevance

Quality issues lead NextMove to bail on its massive Tesla Model 3 order

The German specialist EV rental company experienced a number of issues with the first 15 Model 3s of a 100-car order and canceled the remainder.

Kyle Hyatt August 16, 2019 3:45 PM PDT

ES

Quality control solutions: Configuring technologies for EV motor production

By Zeiss | 04 November 2021

The Volkswagen components plant in Salzgitter manufactures key components for the new APP 310 electric drive destined for the ID.3 model – including an innovative hairpin stator. Zeiss has worked with Volkswagen to develop a comprehensive measuring solution to ensure the quality of this sophisticated engine design

https://www.businessinsider.com/tesla-fixing-quality-issues-nearly-5000-model-3-owners-say-2019-10?IR=T

https://www.cnet.com/roadshow/news/tesla-model-3-nextmove-german-cancels-order/

Quality and its relevance: Food industry

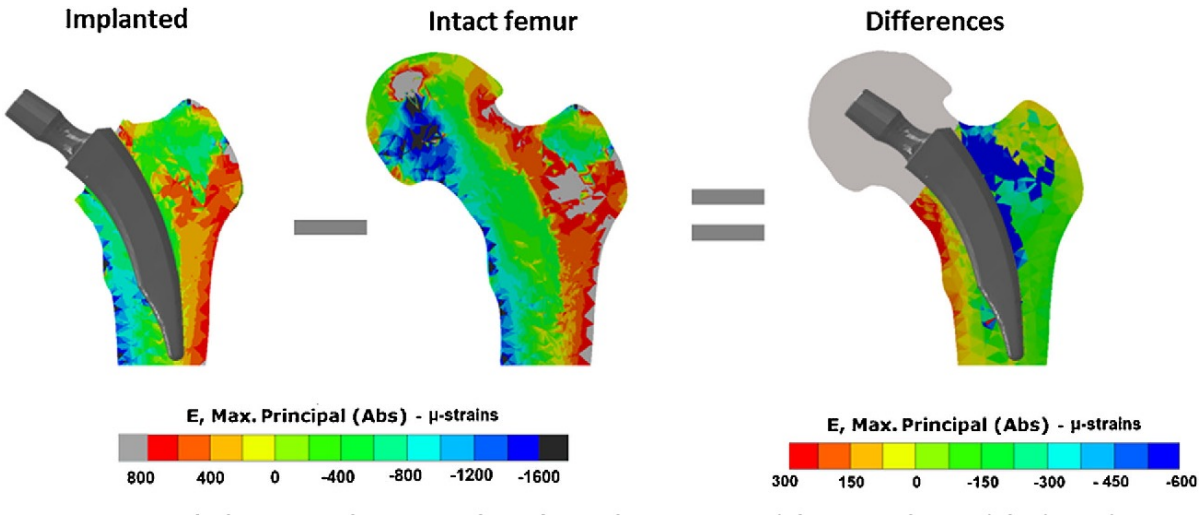
The secret of Coca Cola – is Statistical Process Control!

Importance of Quality Control in Food Industry

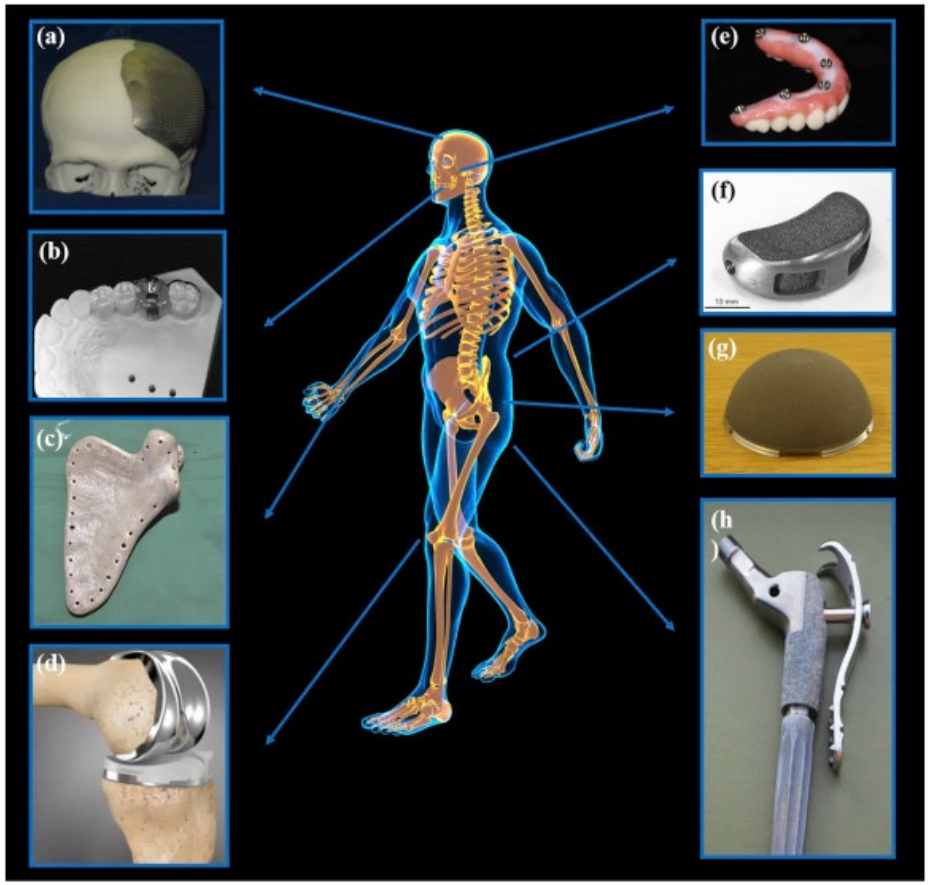
- **Reduced production cost**
- **Better goodwill (reputation)**
- **Increase sales (social media!)**
- **Facilitates Pricing**
- **Improved techniques of production**
- **Higher employee morale**

Quality and its relevance: medical implants

Quality is a key element in biomedical applications

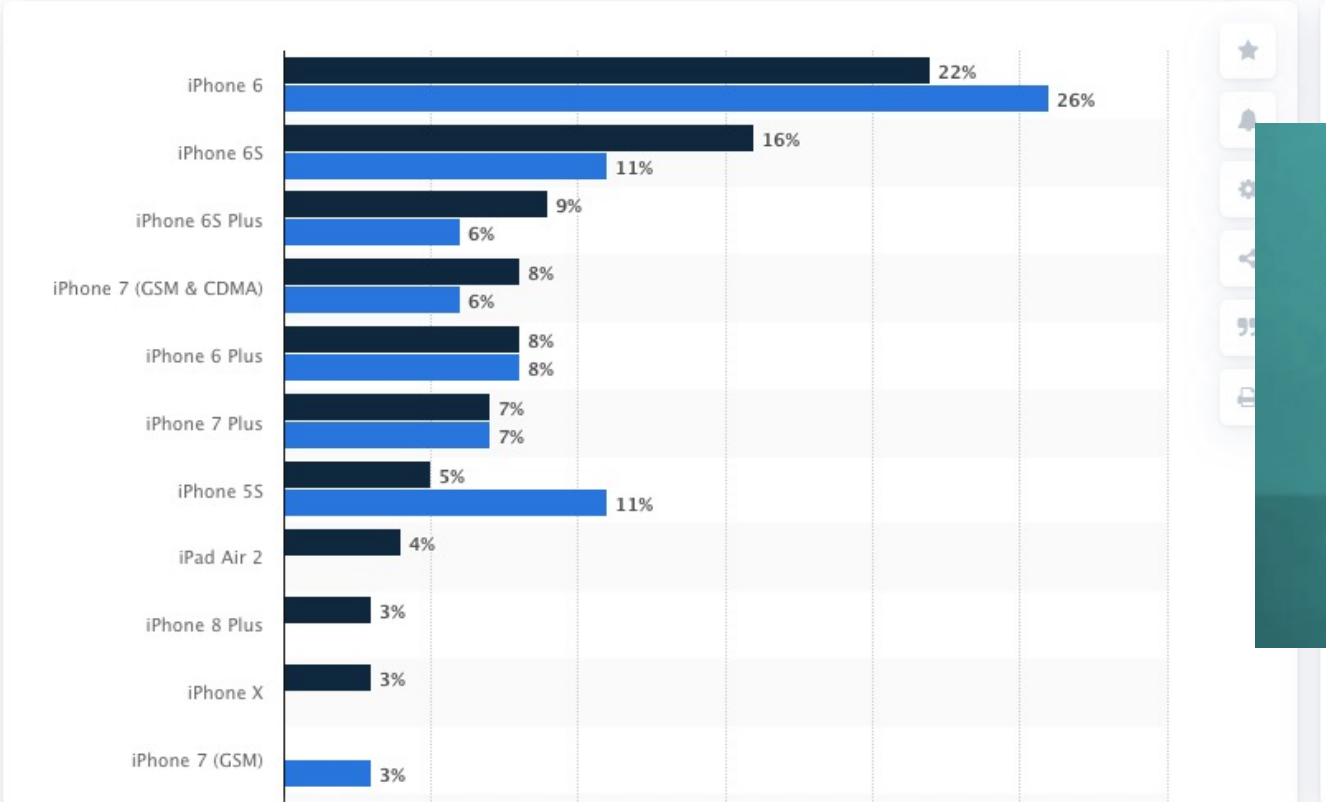


Product quality inspection via X-ray CT



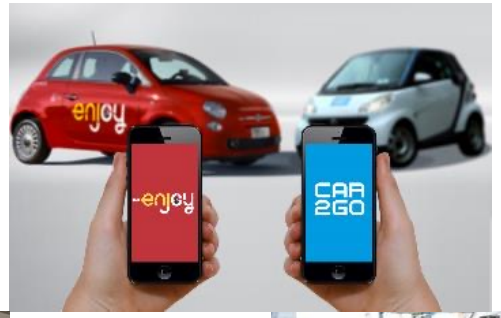
Qualty and its relevance

iphone- failure rate by model worldwide



https://www.statista.com/statistics/804359/iphone-failure-rate-by-model-worldwide/

Quality in services: Quality and customer satisfaction



Examples
Customer care
Teaching evaluation

Examples
Car sharing
Railway/airplane transportation

Salute

HOME ALIMENTAZIONE E FITNESS MEDICINA E RICERCA SALUTE SENO ONCOLINE

Le classifiche degli ospedali che funzionano meglio: il nord primeggia, il sud insegue

La cardiochirurgia è una specialità che si misura abbastanza bene, visto che gli interventi che vengono realizzati in questi reparti super specialistici sono soprattutto due, il bypass e la sostituzione delle valvole. Anche qui lavora meglio chi lavora di più e si valuta la riuscita dell'intervento andando a vedere la mortalità dopo 30 giorni dall'uscita dalla sala operatoria. Ecco quali sono le eccellenze nel nostro Paese.

Bypass aortocoronarico mortalità a 30 giorni (strutture con volumi più alti, almeno 250 casi rilevati da Agenas a biennio)

1 - San Raffaele, Milano - 0%
2 - Cliniche Gavazzeni, Bergamo - 0,37%
3 - Santa Chiara, Trento - 0,50%
4 - Hesperia Hospital, Modena - 0,55%
5 - Spedali Civili, Brescia - 0,62%
6 - Fondazione Giovanni Paolo II, Campobasso - 0,66%
7 - Ospedale di Lecco, Lecco - 0,73%
8 - Città di Lecce, Lecce - 0,74%
9 - Poliambulanza, Brescia - 0,80%
10 - Ismett, Palermo - 0,85%
Media italiana, 2,36%

Examples
Hospitals

Zero defects and zero-waste

Zero Defects

Improving Product Quality and **Eliminating** Waste

Zero defects means getting it right the first time, every time!

How much do quality failures cost your company?
Quality defects have significant costs associated with them – some of the most obvious being money, time, resources, and lost reputation. And programs to eliminate quality defects can be expensive and time consuming

https://www.mindtools.com/pages/article/newTMC_87.htm