



POLITECNICO  
MILANO 1863

# Marketing analytics

## Course introduction

AY 2024/2025 – Lucio Lamberti & Gloria Peggiani

# Rules of the game

01

# Who we are



## **Lucio Lamberti, Ph.D.**

- Full Professor @Politecnico di Milano
- Visiting professor @Shanghai Tongji University, Solvay University, Hong Kong PolyU
- Scientific Director AXD, Metaverse Marketing Lab
- Coordinator – B.R.I.E.L.
- Director of International Master in Media and Communication Management – Polimi GSoM
- Research interests: CRM, customer centricity, marketing accountability
- TT projects: development of customer centricity in organizations, renewal of customer satisfaction assessment processes in utilities, internationalization of made in Italy, analysis of the diffusion of AI/marketing automation among SMEs
- Main collaborations: Google, Meta, RAI, Huawei, Nestlé, Fastweb, Lamborghini, Novartis Group, SAP, Edison, Prada, Hyundai, Amazon, Coop, CRAI

## **Gloria Peggiani**

- PhD Candidate @Politecnico di Milano
- Main research interests: Influencer marketing and creator economy
- Other research interests: event marketing, product attachment, social media marketing
  - Collaborations with such companies as Maserati, Fastweb, Huawei, Uniting Group

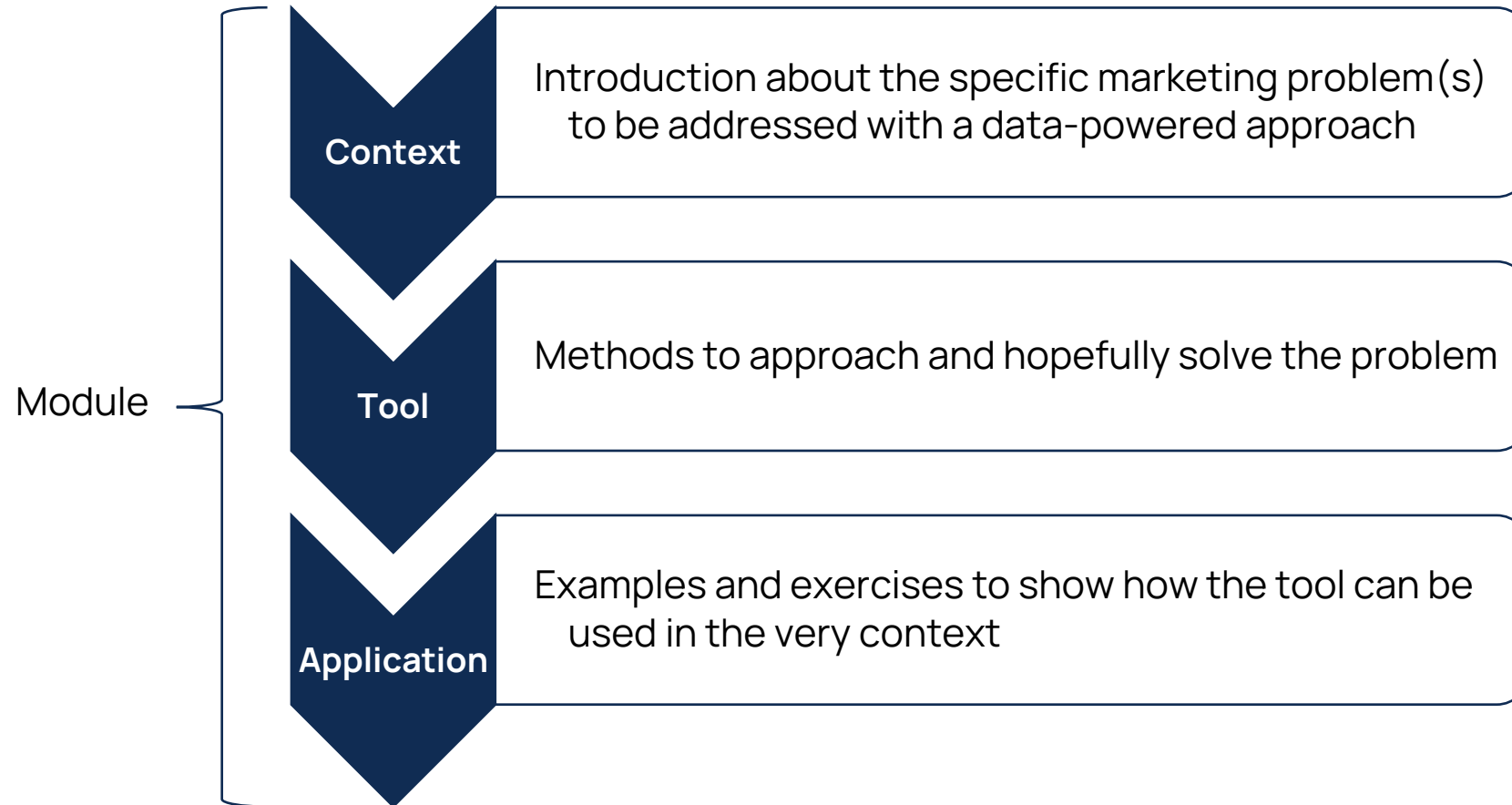


# Course objectives

To introduce tools and methods to process marketing data

To support a data-powered approach in marketing decisions

# Course approach



# Course modules

## Customer (e)valuation

- How to estimate the value of the relationship with a customer
- Methods: CLV, stochastic modelling

## Multivariate Statistics for Social Research

- How to design a market research (e.g., survey) and to process and read the outcomes
- Methods: survey design and multivariate statistics for marketing

## Data monetization from CRM

- How to extract value from CRM data, past purchases and other interactions occurred with customers
- Methods: RFM Strategies, market basket analysis (MBA)

## Churn Prevention

- How to identify which customers are in risk of churning and how to retain the customers in risk of churning?
- Methods: classification models (Regression, Survival Analysis, Classification Trees & Random Forest)

## Marketing Spending Optimization

- How to pursue an optimal allocation of marketing resources to pursue specific business objectives
- Methods: marketing mix modelling, attribution modelling for communication

## Preliminary competences suited to the course

- **Strategy** (strategic planning, segmentation)
- **Marketing** (marketing process, marketing strategy and marketing decisions, branding basics)
- **Basic accounting** (investment analysis, costing)
- **Solid statistical background**

The course has been designed within a stream together with the Applied Statistics course. Attending both courses may significantly help better fill the statistical gaps the student may perceive at the beginning of the term

## Course languages and tools

- **SPSS** (cluster analysis, linear models, regression, hypothesis testing) – you will be provided with a student license
- **R** (survival analysis and random forest)
- **SMART PLS 4** (PLS-SEM)
- **Basic SQL** (useful to fully get the approach to MBA – online free courses are available, for instance:
  - [Datacamp.com](https://www.datacamp.com/courses/introduction-to-sql) introduction to SQL
  - [Codecademy.com](https://www.codecademy.com/learn/learn-sql) introduction to SQL



# Course scheduling

- **Thursday:** 2.30 p.m. – 6.00 p.m. **room B6.3.2**
- All the classes shall be available in streaming and recorded
- Complete scheduling: on WeBeep

# Didactic approach

- **Frontal lectures** (as few as possible)
- **Interactive classes** (case studies and exercises), aimed at turning theory into action
- **Assignment review sessions**

# Textbooks

Main books:

- Stephan Sorger, **Marketing Analytics: Strategic Models and Metrics**, Admirall Press
- Blanchard, Debasish, Pranshu, **Data Science for Marketing Analytics: Achieve your marketing goals with the data analytics power of Python**, Packt Publishing
- 
- Hermann, Burbary, **Digital Marketing Analytics: Making Sense of Consumer Data in a Digital World**, Que Publishing

Specific further references shall be provided for each class

# Contacts and support

We receive by appointment. To fix an appointment:

- [lucio.lamberti@polimi.it](mailto:lucio.lamberti@polimi.it)
- [gloria.peggiani@polimi.it](mailto:gloria.peggiani@polimi.it)

# Evaluation

**The mark of the assignments will be considered for all the calls of the year** (i.e. 2 calls in Jan/Feb 2025 + calls in July 2025 and in Sept 2025)

Written exam **(30%)**

Exercises and/or theoretical questions

Marketing analytics assignment **(70%)**  
in groups of 4 people

Analysis of a dataset related to a marketing issue to provide recommendations



**Kick-off:** late October  
**Delivery:** 19th Jan., 2025

**Two typologies of project work:**

Maximum Grade	Estimated effort (time and workload)
25	● ● ● ○ ○
30	● ● ● ● ●

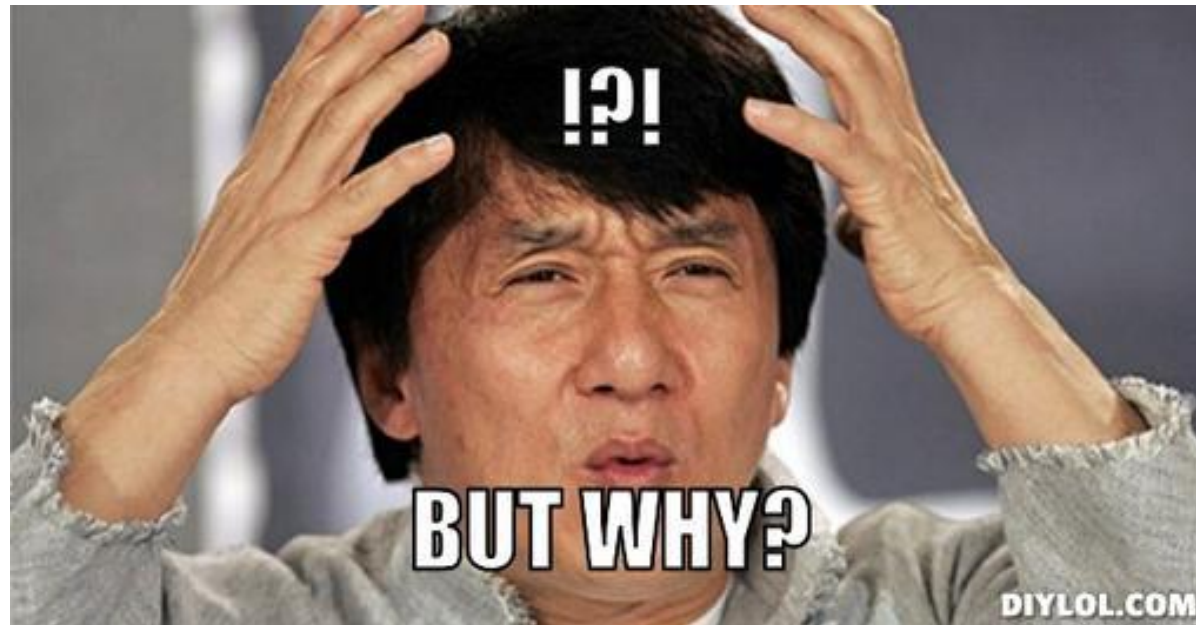


# Q&A

# Data-powered marketing

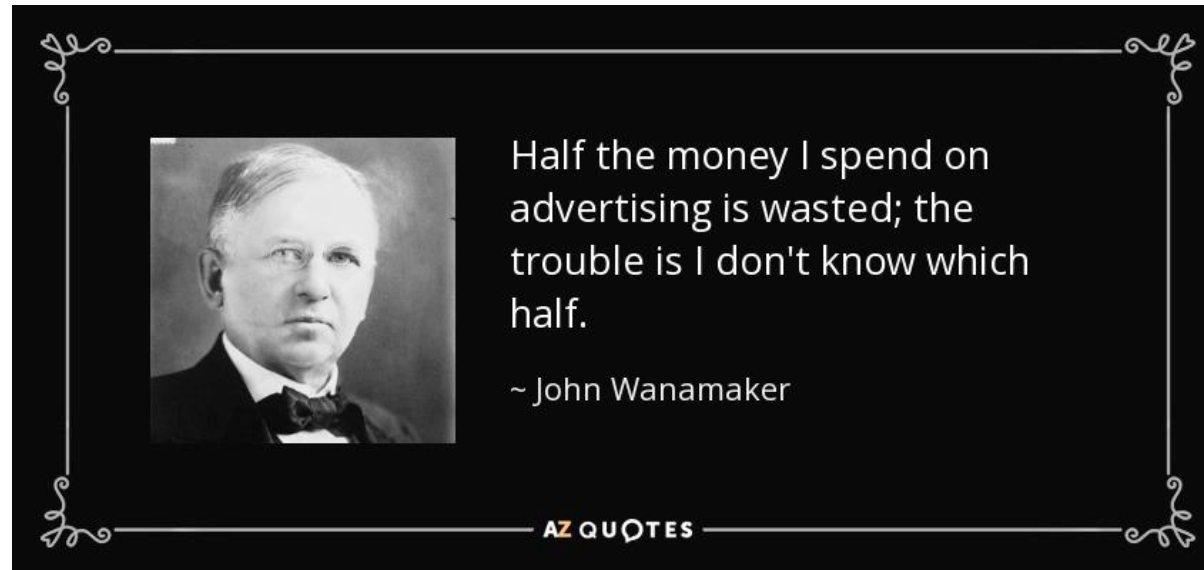
02

# Analytics for Marketing





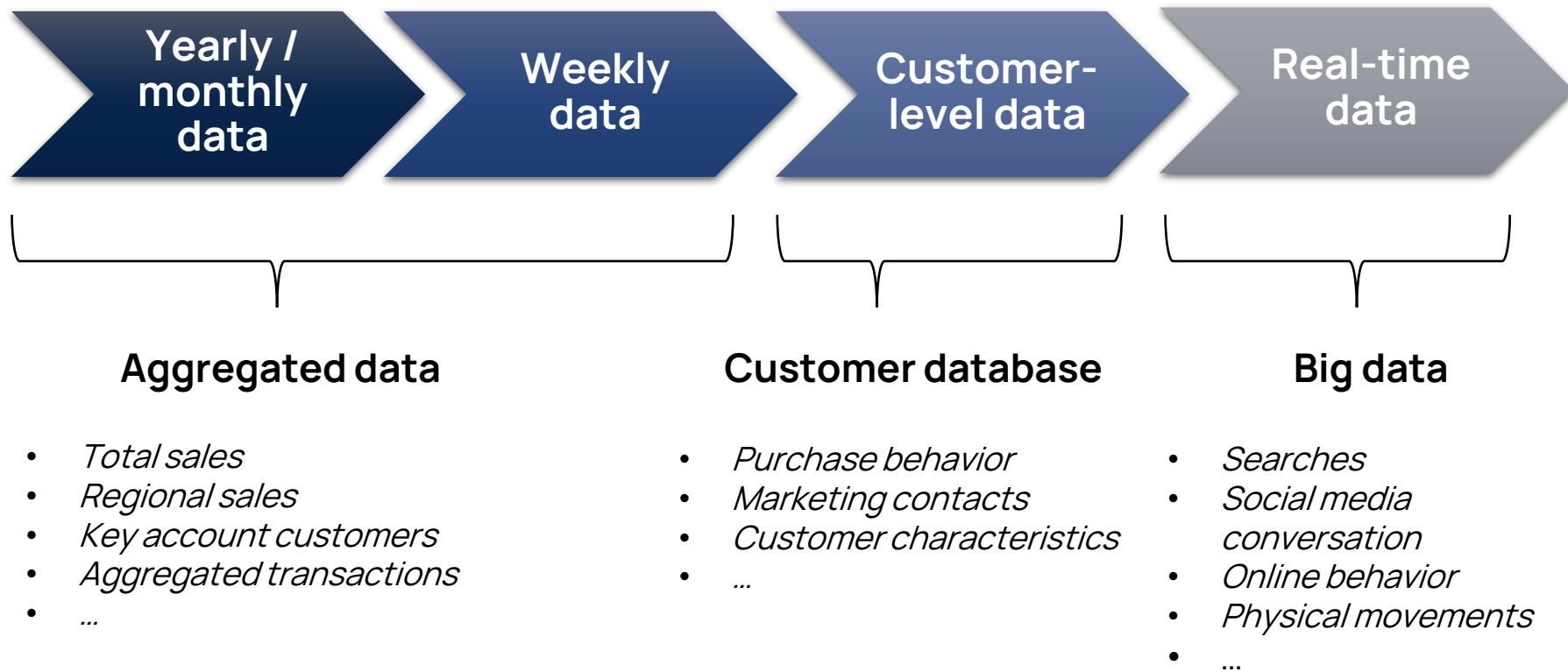
## Once upon a time in marketing...



# Data Analytics for Marketing

It is the process of **collection, management, and analysis** (descriptive, diagnostic, predictive, and prescriptive) **of data to obtain insights** into marketing performance, **maximize the effectiveness** of instruments of marketing control, and **optimize** firms' return on investment.

# Data is nothing new...



## ...but becomes a core strategic element

Data is assuming an **increasingly central role** in organizations, as marketers aim to harness data to:

- Build and maintain **customer relationships**;
- Personalize products, services, and the **marketing mix**;
- Automate **marketing processes** in real time;

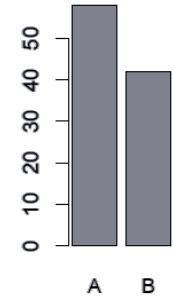
Entirely new forms of marketing have emerged, including recommendations, geo-fencing, search marketing, and retargeting.

# Possible approaches to marketing analytics

## Descriptive Analysis

Building and synthetizing data  
(indices, plots, feature selection...)

Tizio votes for A  
Caio votes for B  
...



## Diagnostic/Inferential Analysis

Extending the information gathered from the observations to the entire phenomenon  
(estimation, null hypothesis testing, ...)

Tizio votes for A  
Caio votes for B  
...



A wins the elections

## Predictive Analysis

Predicting new observations of the phenomenon  
(regression, supervised classification, ...)

Tizio votes for A  
Caio votes for B  
...



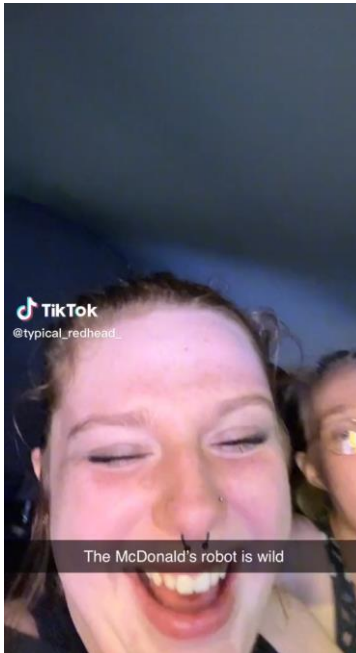
Sempronio votes for A

# The risks of being too data-driven

## Chatbots May 'Hallucinate' More Often Than Many Realize

When summarizing facts, ChatGPT technology makes things up about 3 percent of the time, according to research from a new start-up. A Google system's rate was 27 percent.

Source: [Nytimes](#)



ARTIFICIAL INTELLIGENCE / TECH / GOOGLE

## Google promised a better search experience – now it's telling us to put glue on our pizza

Imagine this: you've carved out an evening to unwind and decide to make a homemade pizza. You assemble your pie, throw it in the oven, and are excited to start eating. But once you get ready to take a bite of your oily creation, you run into a problem — the cheese falls right off. Frustrated, you turn to Google for a solution.

"Add some glue," Google answers. "Mix about 1/8 cup of Elmer's glue in with the sauce. Non-toxic glue will work."

So, yeah, don't do that. As of writing this, though, that's what Google's new AI Overviews feature will tell you to do. The feature, while not triggered for every query, scans the web and drums up an AI-generated response. The answer received for the pizza glue query appears to be based on a comment from a user named "fucksmith" in a more than decade-old Reddit thread, and they're clearly joking.

Source: [TheVerge](#)



The New York Times

## 260 McNuggets? McDonald's Ends A.I. Drive-Through Tests Amid Errors

Source: [Nytimes](#)

MA

The rent on an apartment is \$700 per month. Each month, you have to pay the full \$700 in rent, even if you only live there for part of the month. If you live there for 3 months and 4 days, how much do you pay in rent?

GA

To calculate how much you pay in rent for living in the apartment for 3 months and 4 days, you can break it down as follows:

1. Rent for 3 months:  
 $3 \text{ months} * \$700/\text{month} = \$2,100$
2. Additional rent for 4 days:  
Since there are typically 30 or 31 days in a month, let's assume an average of 30.5 days in a month for simplicity.  
 $4 \text{ days} / 30.5 \text{ days (average month)} * \$700/\text{month} = \text{approximately } \$91.80$

Now, add the rent for 3 months and 4 days together:

$\$2,100 \text{ (for 3 months)} + \$91.80 \text{ (for 4 days)} = \$2,191.80$

So, you would pay a total of \$2,191.80 in rent for living in the apartment for 3 months and 4 days.

MA

That's incorrect. Remember, you have to pay the FULL \$700 RENT for each month. You still have to pay the full \$700 even if you only live there for a few days out of the month.

GA

You are absolutely right, and I apologize for the previous incorrect response. If you have to pay the full \$700 in rent for each month, regardless of how many days you live there, then the calculation is much simpler.

For 3 months and 4 days, you pay the full rent for 3 months and an additional 4 days:

$3 \text{ months} * \$700/\text{month} = \$2,100$   
 $4 \text{ days} * \$700/\text{month} = \$93.33 \text{ (approximate)}$

So, you would pay a total of  $\$2,100 + \$93.33 = \$2,193.33$  in rent for living in the apartment for 3 months and 4 days, considering the full monthly rent is due even if you live there for part of the month.

# SOFTWARE FOR MARKETING ANALYTICS

03

## Excel (or other type of spreadsheet)

Excel is a **powerful, flexible, and easy-to-use tool to manage and process data**. It is excellent to provide you overview of the data, to perform basic analyses (which have great added value nonetheless), and to create tables and graphs for reports.

Excel is not a statistical analytical software. As such, it's flexibility of managing dataset may also lead to a lack of data structure for less experienced users. It cannot perform analysis with complex algorithm. It becomes less efficient in handling huge amount of data.



## SPSS, STATA, SAS

These are the **leading software packages for statistical analysis**. They are **commonly used in social science**, and can perform various types of statistical analysis on large datasets. They often have GUI (Graphic User Interface) so that it is relatively easy to use. They are well developed and robust.

A specific business usually does not require all the types of analysis available in these packages. On the other hand, some particular analyses and statistical models useful for the businesses are not always included in the standard package. The cost of license could be high.

# R, Python

**R** and **Python** are two languages that are very popular in data science. The two have many similar function but relatively speaking, **R is more used in statistical analysis, while Python is more used in machine learning**. They are open-source software (i.e. many contributors create analytical packages to their libraries), therefore they are very powerful that almost any types of analyses can be performed. They are also free to use.

Being open-source and free, an important drawback is that they are not as user friendly as the softwares developed for applied research. They are also less robust, and lack service and assistance.

# ERP / CRM Systems

**ERP (Enterprise Resource Planning) / CRM systems** such as SAP, Oracle, Salesforce, and many smaller vendors, provide the function of **data management and certain analytical functions**. They can quickly (or in real-time) produce reporting as it is directly connected to the data source; but usually it is more about routine reporting, visualizing, and monitoring, rather than uncovering the latent patterns and relationship (statistical analysis).

An ERP / CRM system suitable for the business is important for the systematic collection and management of data. However, it could need to be integrated with other analytical softwares.

## Business intelligence software such as Tableau, Google Analytics

There are many **business intelligence software**, many also cloud-based, to help businesses to make sense of data in specific areas. Mostly, they focus on **summarizing, visualizing, and reporting data**, instead of analyzing latent patterns and relationships. They are very easy to use. Many could produce routine reports automatically and in real-time. They can be useful in monitoring the specific business areas through data. But they should also be integrated with other analytics for decision making.

# GETTING STARTED

04

## Getting started

A spanish marketplace\* has launched in 2017 an eCommerce service for grocery. The customers access the platform, select the store and see the available products in the store.

Some customers have a loyalty card that can be used also when they go shopping directly in the stores.

Two delivery methods are available:

- Home delivery
- Click and collect

Two payment methods are available:

- Cash on delivery
- Credit card/digital payment when the order is sent

You have a basic transaction report for some 6 months

*\* Facts, figures and players are fictionary, but realistic*

# Getting started

User id	Store_id	Loyalty card	Sent Date	Delivery date	Delivery type	Reservation Interval	Payment type	# items	Ticket (€)
Customer5904	0517	N	01/10/18	01/10/2018	Click and collect	16:00:00 - 18:00:00	Cash on delivery	9	9,22 €
Customer2007	0134	Y	01/10/18	01/10/2018	Click and collect	08:00:00 - 10:00:00	Cash on delivery	10	17,01 €

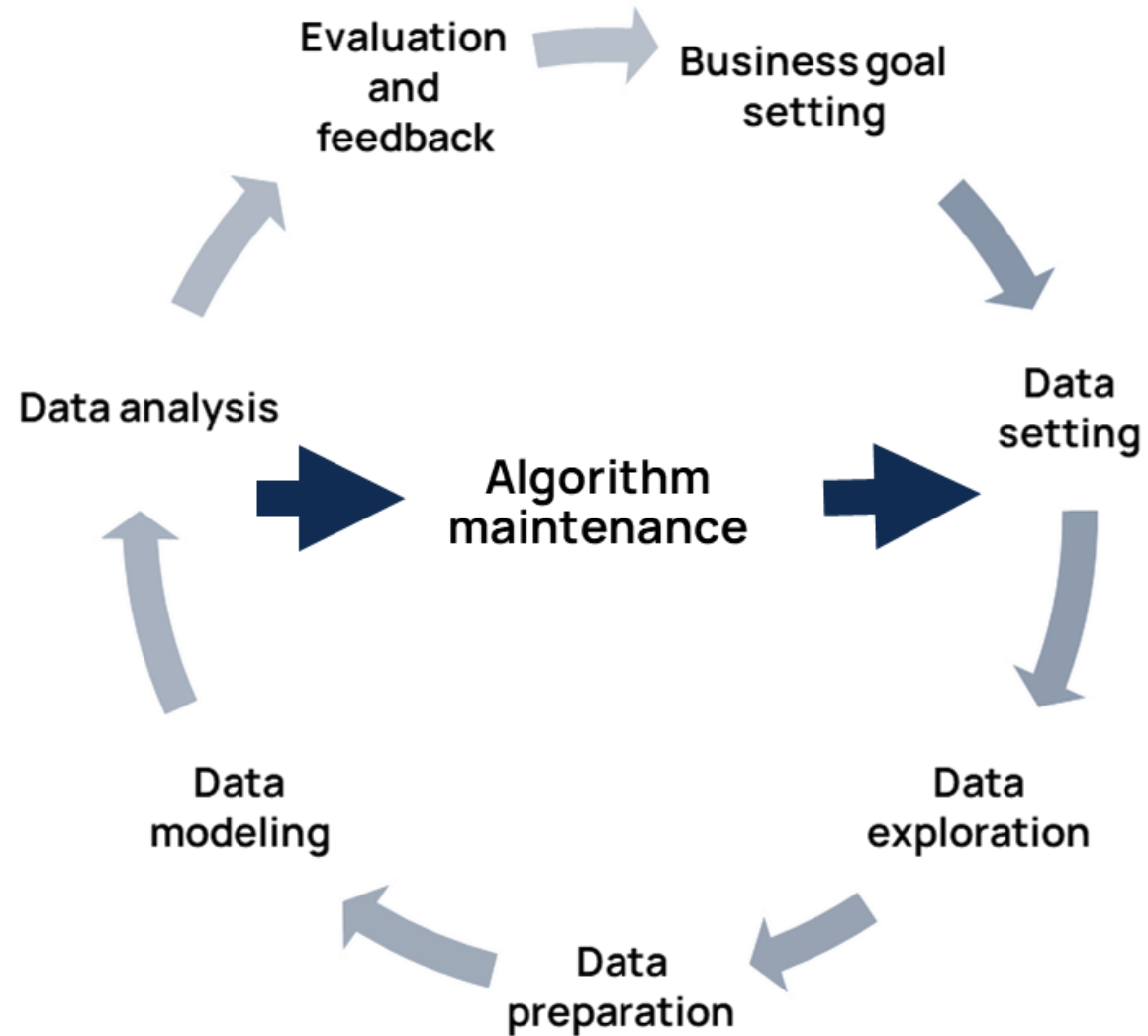
Field	Description
User id	eCommerce user identifier
Store_id	Store identifier. The first two figures represent the region (1 = Madrid area, 2 = Catalunya, 3 = Pays Basque, 4 = Valencia area, 5 = Southern area, 6 = Galizia and western area, 7 = central Spain) , the other 2 figures represent the id of the store within the region
Loyalty card	eCommerce users may or may not have a loyalty card. The loyalty card can bu used both online and offline
Sent Date	Date of the order
Delivery date	Date of the delivery
Delivery type	Home delivery vs click and collect
Reservation Interval	Time of delivery
Payment type	Cash on delivery vs credit card
# items	Number of items in the order
Ticket (€)	Total amount for the order

# Assignment

- What kind of **sense** would you like to make out of the data?
- What kind of **analyses** can be run starting from these data?
- Prepare a **roadmap** for data analysis (objectives & possible analyses)



# A virtuous cycle





**POLITECNICO**  
**MILANO 1863**

Lucio Lamberti – Gloria Peggiani

AY 2024/2025