

# Business Analytics with **pandas** & **matplotlib**: an introduction to Project 2

# Project objectives

---

## 1. Clean the data

It seems like the data extraction process has corrupted some numbers... There's no time to ask for another extraction — solve it!

## 2. Assess data quality

Information across different tables should match... what will you do if it doesn't?

## 3. Answer business questions

This is what management asked, but you're free to come up with your own questions.

## 4. Analyze discounts

Provide clarity and context in a company debate: are discounts beneficial?

# Project overview

---

## Week 1

- Exploring & cleaning the data
- Learning pandas
  - Pandas basics
  - Grouping & summarizing data
  - Advanced pandas
- Learning matplotlib & seaborn
- Crafting lots of plots!
- Extracting the first insights

## Week 2

- Learning string operations & regex (creating categories)
- Creating a business presentation supported by a narrative and visual objects.
- Delivering a convincing presentation

# We will work in groups. Your group will belong to one of the following teams:

## Group 1

Sabrina  
Ilda  
Hana

## Group 2

Sandrine  
Christiaan  
Anand

## Group 3

Rias  
Kristina  
Mahsa

## Group 4

Sevgi  
Andreas  
Akbar

## Group 5

Dónal  
Konstantinos  
Ranjit

# Group work guidelines

## **Everyone must accomplish the main tasks of the project individually**

The objective of having a group is to have colleagues with whom to compare results, discuss common problems, ask and receive some help and build a common presentation. But each and every student has to follow the platform and do the main tasks.

## **Periodically, the group will meet the instructor**

Before the meeting, the group should put all the individual work in common and share the main findings / problems with the instructor.

### **DAY STRUCTURE**

**CHALLENGE REVIEW  
+ NEW DAILY  
CHALLENGE**

### **LUNCH**

**GROUP MEETING  
WITH INSTRUCTOR**

**ONE-TO-ONE'S**