

# Project description

## Context

We are an analyst at a big online store. Together with the marketing department, we've compiled a list of hypotheses that may help boost revenue.

We need to prioritize these hypotheses, launch an A/B test, and analyze the results.

## Description of the data

### Data used in the first part of the project

`hypotheses_us.csv`

- `Hypotheses` — brief descriptions of the hypotheses
- `Reach` — user reach, on a scale of one to ten
- `Impact` — impact on users, on a scale of one to ten
- `Confidence` — confidence in the hypothesis, on a scale of one to ten
- `Effort` — the resources required to test a hypothesis, on a scale of one to ten. The higher the `Effort` value, the more resource-intensive the test.

### Data used in the second part of the project

`orders_us.csv`

- `transactionId` — order identifier
- `visitorId` — identifier of the user who placed the order
- `date` — of the order
- `revenue` — from the order
- `group` — the A/B test group that the user belongs to

`visits_us.csv`

- `date` — date
- `group` — A/B test group
- `visits` — the number of visits on the date specified in the A/B test group specified

Make sure to preprocess the data. There might be mistakes in the original datasets; for example, some of the visitors might have gotten into both group A and group B.

## Part 1. Prioritizing Hypotheses

The file `hypotheses_us.csv` contains nine hypotheses on boosting an online store's revenue with `Reach`, `Impact`, `Confidence`, and `Effort` specified for each. The task is to:

- Apply the `ICE` framework to prioritize hypotheses. Sort them in descending order of priority.
- Apply the `RICE` framework to prioritize hypotheses. Sort them in descending order of priority.
- Show how the prioritization of hypotheses changes when you use `RICE` instead of `ICE`. Provide an explanation for the changes.

## **Part 2. A/B Test Analysis**

We carried out an A/B test and got the results described in the files `orders_us.csv` and `visits_us.csv`.