Description:

The goal of the project is to create a Flink application which will read from Kafka clicks and displays, detect some suspicious/fraudulent activities, and output the suspicious events into a file.

Please download and run this docker-compose file : https://github.com/Sabmit/paris-dauphine/blob/master/docker/kafka-zk/docker-compose.yml

To run it, simply execute `docker-compose rm -f; docker-compose up` in the same directory as the docker-compose.yml file.

By running this docker-compose, it will:

- Create a Kafka cluster with two topics: "clicks" and "displays"
- Launch a python script which will send events to those two topics and display them in the terminal

This generator simulates few suspicious/fraudulent patterns that you should detect using Flink. There are 3 distinct patterns we want you to find.

Requirements:

- Project should be made by group of 2 people, we let you make the groups. Please send us an email with the team members within a week (each group send 1 mail)
- Deadline for the project: End of may
- You can push the project into a Git repository (preferable) or zip the whole project and send it by email.
- You can choose any language/platform (Java/Scala/Python - MacOS/Linux/Windows)
- You can output to a file (easier), or anything else (Kafka is a must)
- You can ask any question on slack, we will try to be as reactive as possible
- Hint: A normal Click Through Rate is around 10%
- You may need to do offline analysis to find outliers, it would ease your work.