Project description

We've been offered an internship in the analytical department at Yandex.Afisha. Our first task is to help optimize marketing expenses. We have:

- Server logs with data on Yandex.Afisha visits from June 2017 through May 2018
- Dump file with all orders for the period
- Marketing expenses statistics

We have been provided with 3 datasets(visits_log_us.csv, orders_log_us.csv,costs_us.csv)

Description of the data are as follows:

The `visits` table (server logs with data on website visits):

- *Uid* — user's unique identifier
- *Device* — user's device
- *Start Ts* — session start date and time
- *End Ts* — session end date and time
- *Source Id* — identifier of the ad source the user came from

All dates in this table are in YYYY-MM-DD format.
The `orders` table (data on orders):

- *Uid* — unique identifier of the user making an order
- *Buy Ts* — order date and time
- *Revenue* — Yandex.Afisha's revenue from the order

The `costs` table (data on marketing expenses):

- source_*id* — ad source identifier
- *dt* — date
- *costs* — expenses on this ad source on this day