

Project description

You've done beautifully in the Practicum course, and you've been offered an internship in the analytical department at Yandex.Afisha. Your first task is to help optimize marketing expenses.

You 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

You are going to study:

- How people use the product

- When they start to buy

- How much money each customer brings

- When they pay off

Instructions for completing the project

Step 1. Download the data and prepare it for analysis

Store the data on visits, orders, and expenses in variables. Optimize the data for analysis. Make sure each column contains the correct data type.

File paths:

- `/datasets/visits_log_us.csv`. [Download dataset](#)

- `/datasets/orders_log_us.csv`. [Download dataset](#)

- `/datasets/costs_us.csv`. [Download dataset](#)

Step 2. Make reports and calculate metrics:

Product

- How many people use it every day, week, and month?

- How many sessions are there per day? (One user might have more than one session.)

- What is the length of each session?

- What's the user retention rate?

Sales

When do people start buying? (In KPI analysis, we're usually interested in knowing the time that elapses between registration and conversion — when the user becomes a customer. For example, if registration and the first purchase occur on the same day, the user might fall into category Conversion 0d. If the first purchase happens the next day, it will be Conversion 1d. You can use any approach that lets you compare the conversions of different cohorts, so that you can determine which cohort, or marketing channel, is most effective.)

How many orders do they make during a given period of time?

What is the average purchase size?

How much money do they bring? (LTV)

Marketing

How much money was spent? Overall, per source and over time.

How much did customer acquisition from each of the sources cost?

How worthwhile were the investments? (ROI)

Plot graphs to display how these metrics differ for various devices and ad sources and how they change in time.

Step 3. Write a conclusion: advise marketing experts how much money to invest and where.

What sources/platforms would you recommend? Back up your choice: what metrics did you focus on? Why? What conclusions did you draw after finding the metric values?

Format: Complete the task in Jupyter Notebook. Enter the code in *code* cells and text explanations in *markdown* cells. Apply formatting and headings.

Description of the data

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

How will my project be assessed?

Your project will be assessed on these criteria. Read over them carefully before you start the project.

Here's what project reviewers look for when assessing your project:

How you prepare the data for analysis

What graphs you plot for metrics

How you interpret the resulting graphs

How you calculate and interpret each parameter

How you back up your recommendations for marketing experts and what metrics you use

Whether you follow the project structure and keep the code tidy

The conclusions you come to

Whether you leave comments at each step