

## Задание 1

У нас есть сервис для покупки дешевых авиабилетов. Основная воронка:

- Юзер зашел на сайт
- Нажал "искать билеты"
- Открыл конкретный билет
- Выбрал предложение конкретного партнера и перешел к нему на сайт
- Купил билет на сайте. Партнер платит нам комиссию за покупку билета.

Мы хотим запустить новую функциональность на экране с выдачей авиабилетов: виджет динамики цен, который показывает пользователю, насколько дешево или дорого стоят билеты по выбранному направлению в данный момент, по сравнению с обычной ценой. Это поможет пользователю понять, выгодно ли покупать билет сейчас или стоит подождать.

- Какие данные и метрики важно исследовать при предварительном анализе фичи? Какие факторы необходимо учесть?
- Какой будет дизайн теста данной фичи?
- На какие показатели ориентироваться при принятии решения об эффективности работы фичи?

## Задание 2

У нас есть датасет с данными по пользовательским поискам авиабилетов и открытиям экрана билета из выдачи. В рамках одного поиска пользователь может как открыть экран билета несколько раз, так и не открыть его ни разу. Будем считать, что все поиски и открытия билетов были сделаны 20 августа 2024 года.

Датасет лежит в файле *task\_2\_data.csv*

У каждой записи есть:

- `user_id` – ID пользователя
- `search_id` – ID поиска
- `origin` – город вылета
- `destination` – город прибытия
- `depart_date` – дата вылета
- `round_trip` – тип поиска: "В одну сторону" или "Туда-обратно"
- `ticket_signature` – ID билета
- `baggage` – конфигурация багажа в билете

2.1 Посчитайте/визуализируйте с помощью Python:

- Долю пользователей, которые открывали билет от общего количества пользователей, в зависимости от типа поиска;
- Конверсию из поиска в открытие трёх или более уникальных билетов в разбивке по городам вылета;
- Среднее количество уникальных открытий билетов в зависимости от конфигурации багажа.

2.2 Представим, что мы залили наш датасет в базу данных в таблицу **serp\_events**. Напишите SQL-запросы, с помощью которых можно:

- Посчитать какой процент пользователей от тех, кто сделал два или более поиска, открывает больше трёх уникальных билетов;
- Определить какому из городов прибытия соответствует наибольшая средняя глубина поиска (разница между датой поиска и датой вылета);
- Определить, является ли количество открытых пользователем уникальных билетов в рамках каждого города отправлением максимальным среди всех пользователей, выполнявших поиск в этом же городе.

### **Задание 3**

Мы провели A/B-тест.

Результаты лежат в файле *task\_3\_data.csv*

Группа a - дефолтный вариант. Группа b - с добавлением новой фичи на экран билета.  
Ключевая метрика - profit.

Вопросы:

- Проанализируйте и интерпретируйте результаты.
- Какие рекомендации по проведению эксперимента могли бы дать?