Место: магазин спорт.товаров

Задача: исследование эффективности проведенных маркетинговых кампаний и выявление факторов, влияющих на рост продаж

Данные: покупки клиентов, их социально-демографические признаки

* **personal\_data** — ID клиентов, их пол, возраст, образование, страна и город проживания;
* **personal\_data\_coeffs** — данные с персональными коэффициентами клиентов, которые рассчитываются по некоторой закрытой схеме (вам потребуется коэффициент personal\_coef);
* **purchases** — данные о покупках: ID покупателя, название товара, цвет, стоимость, гендерная принадлежность потенциальных покупателей товара, наличие скидки (поле base\_sale. Значение 1 соответствует наличию скидки на момент покупки) и дата покупки.

При передаче данных выяснилось, **что часть информации** о клиентах из таблицы **personal\_data была утеряна**

Маркетинговые кампании:

1. с 5 по 16 число; ID пользователей – в ids\_first\_company\_positive.txt; цель кампании – предоставление скидки 5000 клиенты через рассылку
2. с 15 по 45 число; банерная реклама; для клиентов, проживающих в городе 1134 – скидка всем каждое 15-е число месяца

В АВ-тесте, кроме людей с ids\_first\_company\_positive и клиентов из города 1134, также участвовали люди, которым скидку не предлагали (ids\_first\_company\_negative.txt), но они все были схожи по характеристикам покупателей

Задания:

1. Построить модель классификации для восстановления утерянных данных из табл. **personal\_data**
2. Проанализировать, насколько 1-я кампания была эффективнее
3. Провести кластеризацию аудитории и предложить методы работы с каждым кластером
4. Выяснить, какие товары предпочитают клиенты из кластеров и как на продажи повлияло наличие скидки
5. Провести анализ потенциальных клиентов из страны 32 города 1188: предпочтения в товарах

Комиссия: менеджеры-заказчики без технических навыков

Предоставить: отчет и презентацию, не более 10 стр. (Times New Roman-12)