Задание

*Анализ личности клиента — это подробный анализ клиентов компании. Он помогает компании лучше понять своих клиентов и облегчает изменение продукции в соответствии со специфическими потребностями, поведением различных типов клиентов.*

*Анализ личности клиента помогает компании модифицировать свой продукт в соответствии с целевыми потребителями из различных сегментов потребителей. Например, вместо того, чтобы тратить деньги на маркетинг нового продукта для каждого клиента в базе данных компании, компания может проанализировать, какой сегмент клиентов с наибольшей вероятностью купит продукт, и затем продвигать продукт только для этого сегмента.*

*Выполните задания, используя шаблон* [*тетрадки-jupyter.*](https://disk.yandex.ru/d/5JuQhFGbLBKUcw)

*Вам предоставлен* [*Dataset*](https://disk.yandex.ru/d/wKxSUIO9J3PWIg)*, проведите полный цикл анализа данных (EDA), что проходили ранее. Этапы выполнения задания (наш pipeline):*

1. Загрузка и предобработка данных. Очистить данные и удалить/заполнить пропуски.
2. Описательный анализ данных, просмотр данных и вывод статистик
3. Постройте необходимые графики для анализа
4. Проверьте гипотезы
5. Сделайте отчет-вывод по исследованию

*Что хочет узнать заказчик?*

1. Сколько в среднем клиенты тратят на каждый тип товара? На какой товар надо сделать акцент?
2. Какой средний возраст покупателя? Зависит ли возраст клиента от его покупок во всех категориях?
3. Тратят ли клиент больше денег, если у него много детей и иждивенцев?
4. Зависят ли покупки от образования? Что предпочитают люди чаще покупать в зависимости от образования?
5. Женатые люди тратят больше или меньше денег на покупку золота?
6. Люди, которые покупают товар на сайте, чаще берут что-то со скидкой?
7. Зависит ли жалоба клиента от его трат?
8. Люди, принявшие предложение о покупке лишь в последней компании, тратят больше, чем люди, принявшие сразу предложение первой компании?

*ЗАДАНИЕ ПРО (не обязательно к выполнению)*

1. Разбейте данные на несколько выборок: по образованию, по уровню заработка и по семейному статусу. Опишите все эти выборки и скажите — отличаются ли они между собой? Опишите общие закономерности для каждого типа — на что чаще тратит женатый человек и жалуется ли он? Люди с PhD покупают фрукты? И т.д.
2. Разбейте данные на две выборки — по сумме трат всех товаров: 1 — больше медианы трат, 0 — меньше. Сравните выборки между собой и сделайте вывод — на кого надо опираться при модификации товаров магазинам?
3. Проведите дополнительный анализ, который поможет дать больше информации для заказчика.

*Описание столбцов:*

| Название | Описание |
| --- | --- |
| ID | Уникальный идентификатор клиента |
| Year\_Birth | Год рождения клиента |
| Education | Уровень образования клиента |
| Marital\_Status | Семейное положение клиента |
| Income | Годовой доход семьи клиента |
| Kidhome | Количество детей в семье клиента |
| Teenhome | Количество подростков в домохозяйстве клиента |
| Dt\_Customer | Дата регистрации клиента в компании |
| Recency | Количество дней с момента последней покупки клиента |
| Complain | 1, если клиент жаловался в течение последних 2 лет, 0 — в противном случае |
| MntWines | Сумма, потраченная на вино за последние 2 года |
| MntFruits | Сумма, потраченная на фрукты за последние 2 года |
| MntMeatProducts | Сумма, потраченная на мясо за последние 2 года |
| MntFishProducts | Сумма, потраченная на рыбу за последние 2 года |
| MntSweetProducts | Сумма, потраченная на сладости за последние 2 года |
| MntGoldProds | Сумма, потраченная на золото за последние 2 года |
| NumDealsPurchases | Количество покупок, совершенных со скидкой |
| AcceptedCmp1 | 1, если клиент принял предложение в 1-й кампании, 0 — иначе |
| AcceptedCmp2 | 1, если клиент принял предложение во 2-й компании, 0 — иначе |
| AcceptedCmp3 | 1, если клиент принял предложение в 3-й кампании, 0 — иначе |
| AcceptedCmp4 | 1, если клиент принял предложение в 4-й кампании, 0 — иначе |
| AcceptedCmp5 | 1, если клиент принял предложение в 5-й кампании, 0 — иначе |
| Response | 1, если клиент принял предложение в последней кампании, 0 — иначе |
| NumWebPurchases | Количество покупок, совершенных через веб-сайт компании |
| NumCatalogPurchases | Количество покупок, сделанных с помощью каталога |
| NumStorePurchases | Количество покупок, сделанных непосредственно в магазинах |
| NumWebVisitsMonth | Количество посещений веб-сайта компании за последний месяц |