Участник: Куляскин М.О.

Направление: DS

В данном кейсе передо мной была поставлена задача: из имеющихся данных посчитать, как каждая акция, во время ее проведения, повлияла на объем приобретенных напитков. В решении я использовал библиотеку для анализа данных pandas.

1. Исследовательский анализ. Перед тем, как начинать решать задачу, я загрузил все таблицы в Jupyter-Notebook и просмотрел описание их полей в ТЗ. Каждая акция из таблицы promo проводилась в определенный промежуток времени. В таблице effect есть данные о дополнительных литрах (то есть о том, что нам нужно в конечном счете найти) на каждый день и напиток в соответствующем городе. В таблице drinks есть информация о принадлежности каждого продукта к алкогольным или безалкогольным. Это поможет нам при сопоставлении акций и эффекта от них. Остальные таблицы я не использовал, т.к. они не предоставляют информации об акциях или их связи с эффектом. При анализе используемых таблиц я исправил некоторые недоработки в данных: проверил типы колонок и переопределил нужные, заменил или удалил пропущенные значения, добавил необходимые дополнительные строки.
2. Решение задачи. Все мое решение заключается в том, чтобы сопоставить каждой акции суммарный эффект от нее. Для этого я написал несколько функций, считывающих данные о наличии акции в том или ином городе, типе продукта (алкогольный или безалкогольный) на который действует предложение и перечень всех дат проведения акции. Далее для каждой акции я нахожу дополнительный эффект и добавляю его в таблицу promo.
3. Итоги. Мы получили таблицу promo с дополнительным полем volume\_litr, которое показывает, сколько дополнительных литров принесла каждая акция в период ее проведения. По полученным данным мы могли бы оценить зависимость затрат на акцию и объема бонусных продаж, которые она дала.