

GPN-CUP 2023

«Золото — основа всего. Оно способно свергать королей и возводить замки. Именно золото — та причина, которая привела вас сюда...»

Управляющий «Главным Пристанищем Надежды» (бывший)

Поздравляем! Теперь вы — новый управляющий сетью Главное Пристанище Надежды (куда пропал предыдущий является загадкой), которая занимается продажей товаров для здоровья — эстусом (зелье восстановления здоровья), эльфийской пылью и целебными травами. У вас имеется 5 филиалов в различных городах:

- 1) Фалькония
- 2) Анор Лондо
- 3) Врата Балдура
- 4) Нокрон
- 5) Кеджистан.

И именно вы решаете какую цену необходимо установить на каждый из ваших товаров в каждом из городов каждый день!



Но все не так просто: Фракция Астрального Святилища предписывает всем продавцам целебных снадобий согласовывать цену на 90 дней вперед. Кроме того, их правила гласят:

- одна цена на товар должна держаться больше, чем 3 дня (то есть нельзя установить цену на один день, а на следующий день ее сменить).
- запрещается изменение цены больше, чем на 1 золотой за раз (то есть нельзя резко изменить цену с 3 золотых на 4,50 золотых, но можно с 3 золотых на 4 золотых, а затем через 3 дня поднять до 4,5 золотых)
- не искушайтесь поставить слишком высокую цену — все откажутся у вас покупать, А если Древние Боги заметят, что ваша цена на 20% выше, чем у конкурентов, они могут наказать вас за вашу алчность (большим штрафом)

Таков суровый закон Древних Богов и за его нарушение вас ждет суровое наказание (и большой штраф). Ваша задача — составить такое расписание цен, чтобы получить максимальную прибыль за следующие 90 дней (а не то, возможно, вы узнаете, куда пропал предыдущий управляющий)



Исходные данные:

1. Данные о транзакциях.

От предыдущего управляющего вам достались тщательные записи о продажах. Но не такие тщательные, как вам бы того хотелось. В них могут встречаться различные аномалии и ошибки.

	product	price	amount	place	datetime
0	Целебные травы	3.90	1.242125	Анор Лондо	2216-01-02 16:28:00
1	Целебные травы	3.90	-0.079689	Анор Лондо	2216-01-02 16:21:00
2	Целебные травы	3.90	0.882450	Анор Лондо	2216-01-02 03:03:00
3	Целебные травы	3.90	0.621377	Анор Лондо	2216-01-02 03:12:00
4	Целебные травы	3.90	1.367161	Анор Лондо	2216-01-02 05:33:00
...
875031	Эстус	9.28	1.079104	Фалькония	2218-09-27 14:50:00
875032	Эстус	9.28	1.580617	Фалькония	2218-09-27 06:16:00
875033	Эстус	9.28	1.738492	Фалькония	2218-09-27 11:14:00
875034	Эстус	9.28	0.430340	Фалькония	2218-09-27 22:24:00
875035	Эстус	9.28	0.090248	Фалькония	2218-09-27 17:50:00

product — название продукта (эстус, эльфийская пыльца, целебные травы);

price — цена одной унции продукта в золотых. Одна золотая делится на 100 серебряных, поэтому цена указывается до сотых;

amount — объем проданного товара в унциях;

place — город продажи (один из пяти, описанных выше);

datetime — дата и время продажи.

2. Данные о ценах конкурентов.

Благодаря широкой сети информаторов, вам удалось собрать сведения о ценах на товары, схожие с вашими, у конкурентов. Данные это не полные, имеют пропуски и ошибки. Но в целом, по ним можно постараться восстановить картину и понять, как цены конкурентов влияют на ваши продажи.

	place	product	competitor	price	date
0	Анор Лондо	Целебные травы	Арториас&Со	3.78	2216-01-04
1	Анор Лондо	Целебные травы	Арториас&Со	3.78	2216-01-05
2	Анор Лондо	Целебные травы	Арториас&Со	3.78	2216-01-06
3	Анор Лондо	Целебные травы	Арториас&Со	3.78	2216-01-09
5	Анор Лондо	Целебные травы	Арториас&Со	3.78	2216-01-11
...
39451	Фалькония	Эстус	Светлые Души	14.99	2218-09-17
39453	Фалькония	Эстус	Светлые Души	15.36	2218-09-20
39454	Фалькония	Эстус	Светлые Души	15.36	2218-09-22
39455	Фалькония	Эстус	Светлые Души	15.36	2218-09-23
39456	Фалькония	Эстус	Светлые Души	15.36	2218-09-24

date — дата наблюдения за ценой конкурента;
product — название продукта;
competitor — название лавки конкурента;
place — город, в котором была замечена данная цена;
price — цена продукта у конкурента.

3. Данные о погоде

Очевидно, что на продажи влияют и внешние факторы, такие как погода, поэтому мы решили собрать информацию об этом.

	place	hot	rain	snow	date
0	Фалькония	0	0	0	2216-01-02
1	Анор Лондо	1	0	0	2216-01-02
2	Врата Балдура	0	0	1	2216-01-02
3	Нокрон	0	0	0	2216-01-02
4	Кеджистан	0	0	1	2216-01-02

date — дата наблюдения;
place — место наблюдения;
rain — был ли дождь в это время в этом месте;
hot — была ли непереносимая жара в это время в этом месте;
snow — был ли снег в это время в этом месте.

4. Данные о себестоимости.

Для того чтобы рассчитать прибыль, необходимо знать, сколько было затрачено на производство того или иного товара. Отметим, что цена на производство меняется от времени и места производства товара.

	place	product	cost	date
0	Анор Лондо	Целебные травы	2.07	2216-01-02
1	Анор Лондо	Целебные травы	3.07	2216-01-26
2	Анор Лондо	Целебные травы	3.30	2216-02-12
3	Анор Лондо	Целебные травы	4.04	2216-03-06
4	Анор Лондо	Целебные травы	3.88	2216-03-27

date — дата наблюдения;
place — место наблюдения;
product — продукт;
cost — затраты на производство одной унции продукта.

Все данные хранятся в магическом формате `.parquet`, про который можно узнать [здесь](#).

Задача

Составить расписание цен для 5 городов, 3 продуктов на 90 дней в соответствии с правилами, описанными ранее. Результат отправляется Древним Богам в формате `.parquet`. Всего должно быть 1350 строчек ($5 \times 3 \times 15$). Результат должен содержать следующие колонки:

day_number — день, для которого выставляете цену (от 1 до 90 включительно);

product — продукт (эстус, эльфийская пыльца, целебные травы);

place — город наблюдения (Фалькония, Анор Лондо, Врата Балдура, Нокрон и Кеджистан);

price — цена в золотых (округленная до сотых — серебряные)

Пример: первые 15 дней продаж в городе Фалькония для целебных трав.

	day_number	product	place	price
0	1	Целебные травы	Фалькония	10.97
1	2	Целебные травы	Фалькония	10.97
2	3	Целебные травы	Фалькония	10.97
3	4	Целебные травы	Фалькония	10.97
4	5	Целебные травы	Фалькония	10.97
5	6	Целебные травы	Фалькония	10.97
6	7	Целебные травы	Фалькония	10.97
7	8	Целебные травы	Фалькония	10.97
8	9	Целебные травы	Фалькония	10.97
9	10	Целебные травы	Фалькония	10.97
10	11	Целебные травы	Фалькония	10.97
11	12	Целебные травы	Фалькония	10.97
12	13	Целебные травы	Фалькония	10.97
13	14	Целебные травы	Фалькония	10.97
14	15	Целебные травы	Фалькония	10.97

В качестве целевого результата вы должны максимизировать прибыль, которая считается по формуле:

$$\text{price} * \text{amount} - \text{cost} * \text{amount}$$

То есть выручка минус расходы.

Сверхзадачей для данной работы является попытка показать Высшим Сущностям, насколько больше вы заработали, изменяя цену, относительно того, сколько бы вы заработали, используя бэйзлайн (постоянную цену).

Формат предоставления работы:

- **Оформленный репозиторий** с кодом проекта (в виде .zip файла или ссылки на github/аналог), удовлетворяющий следующим критериям:
 - воспроизводимый
 - содержащий Jupyter Notebook (или несколько) с визуализациями и комментариями к проделанным шагам. Язык решения – python
- **Расписание цен.**
- **Оформленный отчет.** В любом формате (docx, pdf, markdown) и любом стиле изложения. Необходимо описать построенные модели и проверенные гипотезы. Также приветствуются комментарии по имеющимся данным и о том, что бы вы сделали по-другому, имея больше времени, ресурсов, маны, кофе и тд.

Критерии оценки

При оценке работ будут учитываться:

- внимательность к данным
- методичность и последовательность выводов
- полнота и корректность решения
- аккуратность и читаемость кода
- информативность и понятность отчета

Значение прибыли на тестовой части при проверке будет учитываться в последнюю очередь, но будет выполнять функции ориентира. Результаты будут опубликованы после сбора всех решений.

Немного о магии

Для чтения .parquet файлов вам понадобится pyarrow

```
conda install -c conda-forge pyarrow
```

затем

```
pandas.read_parquet('data.parquet')
```


Полезные советы (от предыдущего управляющего):

- Не забывайте про инфляцию
- Нужно искать золотую середину: слишком дорого - никто не будет покупать ваш товар, а если слишком дешево - вы рискуете потерять прибыль!
- Порои магия сложнейших моделей машинного обучения проигрывает более простым моделям
- После построения любой модели остановитесь и попробуйте интерпретировать ее, подумайте, как будете защищать ее перед Высшими Силами