

**«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»  
ИМЕНИ М.В. ЛОМОНОСОВА**

Факультет Вычислительной Математики и Кибернетики

Кафедра исследования операций

Отчет по заданию №3

Студенты:	Липатова Александра Смирнов Михаил Разумова Вера
-----------	--

Преподаватель:	Гусева Юлия
----------------	-------------

Группа:	312
---------	-----

# Постановка задачи.

**Applepen** – большая торговая сеть, которая занимается продажей всего двух продуктов: яблок и карандашей. Ее магазины расположены в различных уголках Соединенных Штатов и более 10 лет обслуживают покупателей. Недавно топ-менеджмент компании решил более активно использовать имеющиеся у них данные в принятии решений. Каждый магазин собирает информацию о:

1. Закупках (поставки яблок и карандашей два раза в месяц)
2. Продажах (лог транзакций, по записи на каждую проданную позицию)
3. Инвентарь (месячные данные общего количества яблок и карандашей на складе).

Данные доступны в формате CSV. Внутри файла данные отсортированы по дате.

К сожалению, данные никогда не консолидировались и не проверялись. Нам необходимо получить следующие данные в CSV- файлах:

1. Состояние склада на каждый день
2. Данные о состоянии склада в конце каждого дня после того как все поставки и продажи были совершены.

Подобная информация будет очень ценна менеджерам магазинов. Состояние склада должно строится на основе месячных данных об инвентаре. Известно, что люди воруют из магазинов, но сейчас нет никакой возможности узнать объем сворованного товара.

## План работы.

**1.** 1-го и 15-го числа каждого месяца происходит поставка товаров в магазин. Чтобы рассчитать актуальное количество товара суммируем все произведенные поставки и вычтем общее количество проданных товаров на текущую дату.

**2.** Каждый месяц проходит инвентаризация товара. Посчитать количество ворованного товара можно сравнив количество товара, которое должно быть на складе с реальным количеством товара. Сложим количество товара после прошлой инвентаризации с поставками за текущий месяц и вычтем реальное количество товара на складе. Результатом будет количество ворованного товара.

**3.** Агрегируем данные по году и штату. Для этого выполним пункт (2), но из расчета на весь год. Посчитаем сумму украденного и поставок для каждого штата и внесем данные в таблицу.

# Описание программы.

1. Необходимо скачать файлы с таблицами. Это можно сделать с помощью команды:

```
pip install gsutil
```

```
gsutil cp -r gs://artem-pyanykh-cmc-prac-task3-seed17 .
```

2. Разворачиваем таблицу sell в широкую таблицу, что бы появились колонки для яблок и ручек, также убираем из индекса дату в отдельный столбец

3. Объединяем таблицы закупок и продаж по дате используем метод left join. Так как у нас закупки два раза в месяц то заполняем все пустые значения предыдущим "forward fill".

4. Объединяем таблицу яблок на складе и таблицу с разницей в закупке - продаже, смотрим сколько в первый день было на складе, считаем разницу следующего с предыдущим, чтобы посмотреть сколько появлялось на складе за месяц, заполняем начальное значение тем, что бы заполнили раньше и считаем разницу

5. Все то же самое делаем для ручек.

6. Получаем таблицу с результатами по каждому магазину.

-----  
Shop 1  
-----

Задание 3

	apple_sold	pen_sold	apple_stolen	pen_stolen
year				
2006	681341	52555	78.0	91.0
2007	681281	51939	68.0	76.0
2008	685110	52032	78.0	73.0
2009	682296	52411	81.0	96.0
2010	680411	52600	68.0	96.0
2011	683083	52103	88.0	100.0
2012	685007	52578	82.0	91.0
2013	683401	52133	74.0	90.0
2014	682371	52135	90.0	87.0
2015	681763	51804	77.0	77.0

## Необходимые компоненты.

- Библиотеки

- **pandas** - библиотека предназначенная для хранения таблиц. Так же содержит огромное количество универсальных функций для их комфортной обработки.

- **numpy**

- **os.path**

- **from datetime import datetime, date** - date принимает три аргумента: год, месяц и день. **datetime.datetime** принимает несколько дополнительных аргументов: год, месяц, день, час, минута и секунда

- datetime.timedelta** отображает длительность времени.

- **from datetime import timedelta**

- Программы

- Jupyter Notebook

- LaTeX

## Вклад участников.

- Липатова Александра - Задание 1, Задание 3
- Разумова Вера - Задание 2, ReadMe
- Смирнов Михаил - Задание 2, Задание 3