**Описание**

На этот раз данные имеют следующую структуру:

записываются для каждого пользователя, совершившего покупки, каждый день

для каждой даты есть своя папка, внутри неё – папки для каждого пользователя

внутри каждой папки есть файл data.csv, где и хранятся данные

Схематично выглядит так:

└── data

├── 2020-12-30

│ ├── FirstName\_LastName1

│ │ └── data.csv

│ ├── FirstName\_LastName2

│ │ └── data.csv

│ └── FirstName\_LastName3

│ └── data.csv

└── 2020-12-31

├── FirstName\_LastName1

│ └── data.csv

└── FirstName\_LastName5

└── data.csv

Например, 30 декабря три покупателя сделали покупки, 31 – два

(папки 2020-12-30 и 2020-12-31 соответственно). Поскольку клиент FirstName\_LastName1 купил товары в оба дня, для него имеется папка в папке для каждой из дат. Для других клиентов – по одной.

Note: данные в задании покрывают другой временной период, имена тоже другие. Подробности, примеры и возможные подсказки можно найти в текстах следующих шагов.

**Задачи**

1. Соберите все данные из папки data в один датафрэйм, имеющий следующие столбцы: колонки из самих файлов (product\_id, quantity), а также имя пользователя (name), и дата этих покупок (date), соответствует названию папки, где лежит папка с пользователем)

import pandas as pd

import os

tree=os.walk(‘data’)

for item in tree

print(item)

tree=os.walk('data')

data\_files=[]

target\_file='data.csv'

for item in tree:

files=item[-1]

if len(files)>0 and target\_file in files:

path=os.path.join(item[0], target\_file)

data\_files.append(path)

data\_files\_df=[]

for path in data\_files:

df=pd.read\_csv(path,index\_col=0)

date,name=path.split('/')[1:3]

df['name']=name

df['date']=date

data\_files\_df.append(df

1. Выясните, какой пользователь купил больше всего товаров. Если их несколько, то перечислите имена через запятую с пробелом и в алфавитном порядке.
2. Найдите топ-10 товаров по числу проданных единиц за всё время и постройте барплот. Сколько было продано единиц товара с product\_id==56?
3. Визуализируйте продажи по дням.
4. Сколько пользователей приобрели какой-либо товар повторно (более 1 раза)? Повтором будем считать покупку товара с одинаковым product\_id, совершенную в разные дни.