Formula for calculating LTV is 52(a)\*t, and t is given as 10. So, “a” should be more to get high LTV.

“a” will be high when sum of expenditure of X customers. So, the customers with high expenditure are the Top LTV Customers.

My assumptions in my code:

“e” is the week number and D is data

“x” is the number of top customers and D is filtered data based on week number

So, in my logic, I calculated the total expenditure of each customer and find the top expenditure customers from the sum list of expenditures

My code is below and also attaching python file:

from datetime import datetime

import numpy as np

import json

with open('data.json') as data\_file:

D = json.load(data\_file)

def Ingest(e, D):

eventFilterData = [x for x in D if datetime.date(datetime.strptime(x["event\_time"].split(":")[0], "%Y-%m-%d")).isocalendar()[1] == e]

return eventFilterData;

def TopXSimpleLTVCustomers(x, D):

customers = [ud for ud in D if ud["type"] == "CUSTOMER"]

sites = [ud for ud in D if ud["type"] == "SITE\_VISIT"]

orders = [ud for ud in D if ud["type"] == "ORDER"]

images = [ud for ud in D if ud["type"] == "IMAGE"]

custKeyList = []

expenList = []

for c in customers:

custKeyList.append(c["key"])

expenList.append(sum([float(cs["total\_amount"].split(" ")[0]) for cs in orders if cs["customer\_id"] == c["key"]]))

expenList = np.array(expenList)

TopXLTVIndex = expenList.argsort()[-1\*x:][::-1]

TopXLTV = [custKeyList[i] for i in TopXLTVIndex]

return TopXLTV

UD = Ingest(1, D)

print(TopXSimpleLTVCustomers(2, UD))

**Here, I have taken e (week number) as 1 and x (top customers) as 2**

**Output will be list of top X LTV customer keys**