# Get the proper imports

import re

import csv

# data record example

# data record: 1 134230 345893 Principal Software Engineer

# match the pattern to the dataset

pattern = r"(^\d+)\s+(\d+)\s+(\d+)\s+(\d+)\s+(\d+)\s+(\d+)\s+(\d+)\s+(\d+)\s+(\d+)\s+(\d\*\.?\d\*+)\s+(\d\*\.?\d\*+)\s([A-Za-z ]+)"

# had to use the decimal pattern for two of them, \d\*\.?\d\*+

# OPEN THE TEXT FILE

infile = open('starbucks\_drinkMenu.txt')

# SET UP THE OUTPUT CSV FILE

outfile = open('changedStarbucks\_drinkMenu.csv', 'w', newline = '') # otherwise blank lines between

outfile\_writer = csv.writer(outfile)

# add header row to the data in the csv

fieldnames = ["Calories", "Total Fat (mg)", "Trans Fat (mg)", "Saturated Fat (mg)", "Sodium (mg)"

, "Total Carbohydrates (g)", "Cholesterol (mg)", "Dietary Fiber (g)", "Sugars (g)"

, "Protein (g)", "Caffeine (mg)", "Beverage\_description"]

outfile\_writer.writerow(fieldnames)

# Process each line of data from the text file

for data in infile:

datarec = re.match(pattern, data)

print(datarec)

if datarec:

outfile\_writer.writerow([datarec.group(1), datarec.group(2), datarec.group(3), datarec.group(4)

, datarec.group(5), datarec.group(6), datarec.group(7), datarec.group(8)

, datarec.group(9), datarec.group(10), datarec.group(11), datarec.group(12)])

# gather the data from each group and write it

outfile.close()

print('ALL DONE')