address = input('輸入網址: ')

tota = input('輸入商品總數: ')

import pymysql //匯入python連接mysql的函式庫

db = pymysql.connect("140.136.148.218","dbteam","878787","test", charset='utf8')

//連接資料庫

cur = db.cursor()

import requests

from bs4 import BeautifulSoup //匯入抓取網頁資訊工具的函式庫

nxtpg = ('hello')

i=0

total = (int)(tota) //設定好商品數量和頁數

pag = (total-3)/24

page = (int)(pag)

if page == 0:

lastotal = total

if total < 4:

lastotal = 0

if page>0 and (total-3)%24 == 0:

page = page-1

tagname = ('居家') //設定商品分類

res = requests.get(address) //連接到網址

soup = BeautifulSoup(res.text, "html.parser")

for m in range(page):

x=0

nextpage2 = 1

nextpage1 = 0

if page>9:

nextpage2 = 2

nextpage1 = 1

if m>9 and page >19 and m<(page-(page%10)):

nextpage2 = 3

for x in range(24):

if x==0 and m!=0:

res = requests.get("https://tw.buy.yahoo.com/"+nxtpg) //進入下一頁

soup = BeautifulSoup(res.text, "html.parser")

reup = soup.select('.neighbor.yui3-u')[nextpage2]

nxtpg = reup.findAll('a')[0].get('href') //抓取下一頁的網址

if m==0 and x==0:

reup = soup.select('.neighbor.yui3-u')[nextpage1]

nxtpg = reup.findAll('a')[0].get('href')

reup = soup.select('.yui3-u-1-4')[x]

print( reup.select('.yui3-u-1.name')[0].text)

name = reup.select('.yui3-u-1.name')[0].text //抓商品名稱

print("$",reup.select('.yui3-u-1.pd-price span.nums')[0].text)

price = (int)(reup.select('.yui3-u-1.pd-price span.nums')[0].text) //抓價格

print(reup.findAll('a')[0].get('href'))

link = reup.findAll('a')[0].get('href') //抓商品網頁網址

print(reup.findAll('img')[0].get('\_src'))

Plink = reup.findAll('img')[0].get('\_src') //抓商品圖片網址

print()

dbdata = [name, price, link, Plink, tagname]

cur.execute("INSERT INTO yahoo(name, price, link, pic, tag) VALUES(%s, %s, %s, %s, %s)", dbdata) //匯入資料到資料庫

i = i+1 //計算商品總數