

Proyecto-2 Data Science



INTRODUCCIÓN AL ANÁLISIS DE DATOS

José Eduardo Méndez Martínez

Índice:

Introducción.....	3
Caso.....	4
Objetivo.....	4
Definición del código.....	5
Estructura del archivo	4
Código explicado.....	4
Solución al problema.....	12
Conclusión.....	13
Conclusión Personal.....	13

Introducción:

En este ejercicio se pone a prueba la lógica de programación del estudiante para implementar una solución al caso propuesto de la empresa Synergy Logistics, se nos proporciona una lista de datos para trabajar y poner en práctica las herramientas esenciales de Python para realizar el análisis de datos a partir de la clasificación y manejo de los mismos, obtenemos los datos mediante la creación-uso de archivos y nos apoyamos de lo visto: estructuras de datos, implementación-creación de funciones, uso de módulos importados y destructuring.

Caso:

Synergy Logistics es una empresa dedicada a la intermediación de servicios de importación y exportación de diferentes productos. Actualmente la empresa cuenta con una base de datos que refleja las rutas más importantes que opera desde el año 2015, con su respectivo origen y destino, año, producto, modo de transporte y valor total. Su propósito, es que a partir de estos datos se genere un análisis que sirva de la base para la estructuración de su estrategia operativa.

Objetivo:

La Dirección de Synergy Logistics ha solicitado al equipo operativo, realizar una propuesta que permita enfocar las prioridades de la estrategia operativa 2021; para ello, se plantea analizar la viabilidad de 3 opciones de enfoque: rutas de importación y exportación, medio de transporte utilizado y valor total de importaciones y exportaciones.

Considerando que eres el data analyst del equipo, se te solicita que realices un análisis de los siguientes puntos e identifiques cuál(es) de ellos es la mejor opción para la empresa:

Opción 1)

Rutas de importación y exportación. Synergy logistics está considerando la posibilidad de enfocar sus esfuerzos en las 10 rutas más demandadas. Acorde a los flujos de importación y exportación, ¿cuáles son esas 10 rutas? ¿le conviene implementar esa estrategia? ¿por qué?

Opción 2)

Medio de transporte utilizado. ¿Cuáles son los 3 medios de transporte más importantes para Synergy logistics considerando el valor de las importaciones y exportaciones? ¿Cuál es el medio de transporte que podrían reducir?

Opción 3)

Valor total de importaciones y exportaciones. Si Synergy Logistics quisiera enfocarse en los países que le generan el 80% del valor de las exportaciones e importaciones ¿en qué grupo de países debería enfocar sus esfuerzos?

Implementación de Solución

Se nos proporciona una lista para realizar una implementación y evaluar las opciones de solución con el fin de realizar el análisis a las opciones de solución propuestas:

```
# -*- coding: utf-8 -*-

"""
Analisis de Datos en sistema Synergy Logistics
"""

import csv

lista_datos = []
with open("synergy_logistics_database.csv", "r") as archivo: #Se abre archivo csv
    en modo lectura
    lector = csv.DictReader(archivo) #Se utiliza archivo como diccionario
    for registro in lector: # Iteramos para obtener los registros a la lista de
datos
        lista_datos.append(registro)

# 1.- Rutas más demandadas
def rutas_exportacion_importacion(direccion): #Definimos una función para
recibir los parámetros de las Exportaciones e Importaciones
    contador = 0 #La variable contador guardara el número de
veces que aparece la ruta
    rutas_contadas = [] #Creamos dos listas vacías para utilizar a
continuación
    rutas_conteo = []

    for ruta in lista_datos: #Recorremos la informacion en la lista para
extraer los datos para nuestras listas vacias
        if ruta["direction"] == direccion: # validando la dirección = True
            ruta_actual = [ruta["origin"], ruta["destination"]] #Agrega origen y
destino a la ruta

            if ruta_actual not in rutas_contadas: # Creamos condición para obtener
las rutas que no se han contado.
                for ruta_base in lista_datos: # Iteramos para validar que se
sumen las rutas con los datos especificados:
                    if ruta_actual == [ruta_base["origin"],
ruta_base["destination"]] and ruta_base["direction"] == direccion:
                        contador+=1
```

```

        rutas_contadas.append(ruta_actual) # Agregamos la información que
obtuvimos a nuestra tabla final
        rutas_conteo.append([ruta["origin"], ruta["destination"], contador])
        contador = 0

    rutas_conteo.sort(reverse = True, key = lambda x:x[2]) #Ordenamos de mayor a
menor el conteo por ruta
    rutas_conteo=rutas_conteo[0:10] #Mostramos solo las 10 Rutas con más demanda
de nuestra colección
    return rutas_conteo

conteo_exportaciones = rutas_exportacion_importacion("Exports") # Variable que
mandara a llamar la función para guardar los datos de rutas que contengan Exports
conteo_importaciones = rutas_exportacion_importacion("Imports") # Variable que
mandara a llamar la función para guardar los datos de rutas que contengan Imports

print(" Top 10 de rutas con más demanda en Exportaciones: \n ") # Iteramos para
mostrar las 10 rutas con más demanda para exportaciones
for r_exp in conteo_exportaciones:
    print (r_exp)

print(" Top 10 de rutas con más demanda en Importaciones: \n ") # Iteramos para
mostrar las 10 rutas con más demanda para importaciones
for r_imp in conteo_importaciones :
    print (r_imp)

#2.- Valor de medios de transporte

def transporte_valor_exports_import(direccion): # Creamos función que recibira
parámetros Exportación o Importación
    contador = 0 # La variable contador guardara
el número de veces
    total=0 # Variable total guardara la suma
del valor por cada medio de transporte
    transporte_contadas = [] # Creamos dos listas vacías para
utilizar a continuación
    transportes_conteo = []

    for ruta in lista_datos: # Recorremos la informacion en la
lista para extraer los datos necesarios para nuestras listas vacías
        if ruta["direction"] == direccion: # Validación
            ruta_actual = [ruta["transport_mode"]]

```

```

        if ruta_actual not in transporte_contadas: # Condición para nuestra
variable contador, contara el número de veces que se encuentra el medio de
transporte

            for ruta_bd in lista_datos:
                if ruta_actual == [ruta_bd["transport_mode"]] and
ruta_bd["direction"] == direccion:
                    contador +=1
                    total += int(ruta_bd["total_value"]) #Convertimos el dato a
entero para realizar la suma

            transporte_contadas.append(ruta_actual) # Agregamos información
obtenida a nuestra tabla final

            transportes_conteo.append([ruta["transport_mode"], contador,total])
            contador = 0
            total=0

    transportes_conteo.sort(reverse = True, key = lambda x:x[2]) # Ordenamos de
mayor a menor con nuestra expresion lambda el valor total de cada medio de
transporte

    transportes_conteo= transportes_conteo[0:3] # Mostramos solo los 3 principales
medios de transporte considerando su valor

    return transportes_conteo

transporte_exportaciones = transporte_valor_exports_import("Exports") # Variable
que mandara a llamar la función para guardar los datos de transporte que contengan
Exports
transporte_importaciones = transporte_valor_exports_import("Imports") # variable
que mandara a llamar la función para guardar los datos de transporte que contengan
Imports

print(" Top 3 de mejores medios de transporte en exportaciones: \n ") # Iteramos
para mostrar los 3 mejores medios de transporte en Exportaciones
for t_exp in transporte_exportaciones:
    print (t_exp)

print(" Top 3 de mejores medios de transporte en importaciones: \n ") # Iteramos
para mostrar los 3 mejores medios de transporte en Importaciones
for t_imp in transporte_importaciones:
    print (t_imp)

# 3.- Países que generan 80% de valor

def paises_exports_imports(direccion): # Función que obtendra como parámetro
Exportación o Importación

```

```

    contador=0                                # Variable que guardará cuantas veces
aparece el país
    total=0                                    # Variable que guardara valor total
generado por país
    paises_contados=[]                        # Listas vacias
    valor_paises=[]

    for pais in lista_datos:                  # Se recorre lista de datos
        if pais["direction"] == direccion:
            pais_actual= [pais["origin"]] # Guardamos la variable a comparar

            if pais_actual not in paises_contados: #Conteo pais y total
                for pais_b in lista_datos:
                    if pais_actual== [pais_b["origin"]] and pais_b["direction"]==
direccion:
                        contador +=1
                        total += int (pais_b["total_value"]) # Convertimos el dato a
entero para realizar la suma

                    paises_contados.append(pais_actual)
                    valor_paises.append([pais["origin"],total,contador]) # Agregamos la
información obtenida a la tabla final
                    contador=0
                    total=0

    valor_paises.sort(reverse=True, key = lambda x:x[1]) #Ordena de mayor a menor el
valor total de cada país
    return valor_paises

# Creamos otra función para obtener el porcentaje y conocer los países que forman
parte de dicho porcentaje
def porcentaje_paises_exports_imports (lista_paises, porcentaje = 0.8):
    valor=0
    valor_actual=0
    paises=[]
    porcentajes_calculados=[]

    for pais in lista_paises: # Recorremos nuestra lista de datos por pais
        valor += pais[1]      # Sumamos por pais

    for pais in lista_paises: # Recorremos nuestra lista de datos por pais
        valor_actual += pais[1] # Sumamos valor de operación
        porcentaje_actual = round(valor_actual / valor, 3) # Dividimos valor del
país entre el valor total para obtener porcentaje
        paises.append(pais)

```



```

    porcentajes_calculados.append(porcentaje_actual) #Se agrega la información
obtenida a la tabla final

    if porcentaje_actual <= porcentaje: #Condición para cumplir con el
porcentaje solicitado
        continue
    else:
        if porcentaje_actual - porcentaje <= porcentajes_calculados[-2] -
porcentaje: # Porcentaje que se desea
            break
        else:
            paises.pop(-1) #Si se rebasa el porcentaje que se desea, se borrará
el último elemento sumado
            porcentajes_calculados.pop(-1)
            break

    return paises

paises_80_exportaciones =
porcentaje_paises_exports_imports(paises_exports_imports("Exports")) # Variable que
mandara a llamar la funcion de porcentaje para guardar los porcentajes por pais que
contengan Exports
paises_80_importaciones =
porcentaje_paises_exports_imports(paises_exports_imports("Imports")) # Variable que
mandara a llamar la funcion de porcentaje para guardar los porcentajes por pais que
contengan Imports

print(" Países que cumplen con el Porcentaje de exportaciones son: ") # Iteramos
para mostrar los Países que cumplen con el 80% en exportaciones
for p_exp in paises_80_exportaciones:
    print(p_exp)

print(" Países que cumplen con el Porcentaje de importaciones son: ") # Iteramos
para mostrar los Países que cumplen con el 80% en importaciones
for p_imp in paises_80_importaciones:
    print(p_imp)

```

Análisis: Estrategia sugerida para Synergy logistics

Enfocarse en las rutas no es la mejor opción, la frecuencia con la cual las rutas han sido usadas por la empresa puede variar alguna ruta pudo ser muy frecuente durante los primeros 2 años de operación pero luego pudo disminuir en cuanto a su frecuencia de uso e ingresos. La ruta (South Korea-Vietnam) ha sido la más usada con un total de 497 operaciones, sin embargo no es la ruta que ha generado el mayor de los ingresos, la ruta (China-México) siendo la quinta ruta más frecuente supera a la primera con casi el doble de ingresos generados.

En cuanto a los medios de transporte utilizados se obtuvo que los más frecuentes son los marítimos pero considerando el total de ingresos contra el promedio de los mismos es más conveniente y genera más ganancias usar medios Aéreos que Marítimos

Ingresos/frecuencia de uso:

Marítimos: $(100,530,622,000 / 10688) = 9,405,933.9$

Carril: $(43,628,043,000 / 3381) = 12,903,887.3$

Aéreo: $(38,262,147,000 / 2389) = 16,015,967.7$

Carretera: $(33,270,486,000 / 2598) = 12,806,191.6$

Tomar a los países que han generado el 80% de los ingresos de la empresa como prioridad, puede ayudar a realizar mejoras y a aumentar la cartera de productos y servicios por parte de la empresa. Se deben considerar todos los países y el liderazgo de China, México, USA y Japón ya que sirven de ejemplo para aumentar mejoras en los países que vienen por debajo

Conclusiones:

¿Entonces cuál sería la mejor estrategia para Synergy Logistics?

Synergy Logistics debe planificar su estrategia para 2021 con las siguientes bases:

Enfocarse principalmente en los países más relevantes que hacen uso de su servicio y le han generado el 80% de sus ingresos para dar más soporte a las ganancias mejorando precios, formas de envío, productos y sucursales en estos países.

Como mejora:

1.- Se deberían aumentar más servicios a través de los medios aéreos y de carril que de los marítimos ya que generan mayores ingresos y son más eficientes en tiempo.

2.- Trabajar también en un análisis en cuanto a mayor demanda de productos. El año 2019 ha sido el mejor año en cuanto a ingresos de la empresa, considerando los datos de 2019 y 2020, se obtuvo que los productos más distribuidos han sido computadoras y carros.