

Santander - EMTECH (Emerging Technologies Institute)

Proyecto 2 - Introducción al Análisis de Datos - Synergy Logistics

Pérez Ruiz Daniel Michell

Domingo 27 de septiembre 2020.

Github: <https://github.com/danielmichellpr>

Link del proyecto: <https://github.com/danielmichellpr/Emtech-Introduccion-danielmichellpr>

Resumen

En el siguiente proyecto se pone en práctica el introducción al análisis de datos, se realiza el estudio de la empresa **Synergy Logistics**, la cual es una intermediaria dedicada a la importación y exportación de diferentes productos. Esta empresa cuenta con una base de datos desde el 2015 y se busca obtener cuales son las rutas de mayor demanda, es decir, por las que se realizan más importaciones y exportaciones. Al realizar está consigna se obtuve que las dos rutas de exportación que realizan más viajes son South Korea-Vietnam y Netherlands-Belgium, mientras que las dos rutas de importación más utilizadas son Japan-Mexico y China-Japon. Con está consigna se logra identificar cuales son los clientes más estables para la empresa y con los cuales debería reforzar apoyos para seguir creciendo.

Índice

1. Introducción	1
2. Desarrollo del Código – Consigna de Synergy Logistics	2
3. Opción Sugerida	5
4. Conclusión.	6

1. Introducción

La ciencia de datos es una disciplina que a través de métodos, procesos, algoritmos y sistemas se encarga de la captura, procesamiento, análisis y representación de los datos. Ser un científico de datos requiere habilidades de programación, estadística, comunicación y conocimiento de dominio. La empresa **Synergy Logistics** es intermediaria en la exportación e importación de diversos productos desde 2015, hay relaciones entre países que realizan una mayor demanda a la empresa, es por eso que es de interés saber cuáles son las rutas entre esos países, tanto de importación como de exportación. La consigna es obtener estas rutas, para que para 2021 la empresa obtenga mayor demanda. Independientemente de la cantidad monetaria que ingrese cada ruta, es importante tener la confianza de las rutas más utilizadas, para refrendar el compromiso de la empresa con esas rutas, ya que mientras más confianza se tenga en la empresa, existe más posibilidad de que esas rutas inviertan más. Es por eso que hemos elegido esta consigna, sobre la de valor de importaciones y exportaciones o las de medio de transporte utilizado, ya que la confianza es un beneficio que va más allá de lo monetario y sobre el cual las empresas deben preocuparse.

2. Desarrollo del Código – Consigna de Synergy Logistics

Para obtener esta consigna, lo primero que se hizo con la base de datos de Synergy Logistics fue importarla y guardarla en una lista.

```
1 import csv
2
3 lista_datos = []          #Creamos una lista donde se almacenarán los datos
4
5
6 with open('synergy_logistics_database.csv','r') as archivo_csv:
7     #Cargamos la base de datos
8     lector = csv.reader(archivo_csv)
9     for linea in lector:
10         lista_datos.append(linea)    #Almacenamos los datos en la lista creada
```

Creamos una función, donde almacenaremos el conteo de las rutas de exportación e importación. Imprimimos las rutas de exportación e importación más solicitadas desde 2015 a 2020.

```
1 ##### Función para el conteo de rutas #####
2
3 def rutasTodas(direccion):
4     contador = 0
5     rutas_contadas = []
6     conteo_rutas= []
7     for ruta in lista_datos:
8         if ruta[1] == direccion:
9             ruta_actual =[ruta[2], ruta[3]]
10            if ruta_actual not in rutas_contadas:
11                for movimiento in lista_datos:
12                    if ruta_actual == [movimiento[2], movimiento[3]]:
13                        contador +=1
14                        rutas_contadas.append(ruta_actual)
15                        conteo_rutas.append([ruta[2], ruta[3], contador])
16                        contador = 0
17                        conteo_rutas.sort(reverse = True, key = lambda x:x[2])
18    return conteo_rutas
19
20 rutas_exportacion_todas = rutasTodas('Exports')
21 rutas_importacion_todas = rutasTodas('Imports')
22
23
24
25 input('Las cinco rutas más demandas de exportación por año son (presiona Enter)')
26
27 input('Las rutas de exportación más solicitadas desde 2015 a 2020:')
28 for i in rutas_exportacion_todas[0:5]:
29     print('La ruta de origen de ',i[0], 'con destino',i[1], 'tuvo ', i[2], '
30     exportaciones.' )
31
32 print('*****')
33
34 input('Las rutas de importación más solicitadas desde 2015 a 2020: \n')
35 for i in rutas_importacion_todas[0:5]:
```

```

36 print('La ruta de origen de ',i[0], 'con destino ',i[1], 'tuvo ', i[2], '
importaciones. \n' )

```

A continuación, realizamos una nueva función, donde almacenamos dos listas, una de exportación y una de importación. A partir de esto, realizamos un análisis por cada año creando nuevas listas de exportación e importación.

```

1
2 exportacionesfalso = [] #Creamos dos listas provisionales
3 importacionesfalso = []
4
5 #Creamos una función para almacenar listas de exportación e importación
6 def direction(direccion, lista_e_i):
7     for i in lista_datos:
8         if i[1] == direccion:
9             lista_e_i.append(i)
10    return lista_e_i
11
12 exportaciones = direction('Exports',exportacionesfalso)
13 importaciones = direction('Imports',importacionesfalso)

```

El almacenamiento en estas nuevas listas se realiza a partir de dos funciones más.

```

1
2 ##### CREAMOS LISTAS POR ANHO #####
3
4
5 e2015 = []
6 e2016 = []
7 e2017 = []
8 e2018 = []
9 e2019 = []
10 e2020 = []
11 i2015 = []
12 i2016 = []
13 i2017 = []
14 i2018 = []
15 i2019 = []
16 i2020 = []
17
18
19 ## Creamos dos funciones, para almacenar los datos por años, una de exportación y
una de importación
20
21 def anhose(año, lista_años):
22     for i in exportaciones:
23         if i[4] == año:
24             lista_años.append(i)
25     return lista_años
26
27 def anhosi(año, lista_años):
28     for i in importaciones:
29         if i[4] == año:
30             lista_años.append(i)
31     return lista_años

```

Nuevamente aplicamos una función para obtener el conteo de exportaciones e importaciones, ahora por año.

```

1 ##### Función para el conteo de rutas por año #####
2 def rutas(anhoe):
3     contador = 0
4     rutas_contadas = []
5     conteo_rutas= []
6     for ruta in anhoe:
7         ruta_actual =[ruta[2], ruta[3]]
8         if ruta_actual not in rutas_contadas:
9             for movimiento in anhoe:
10                 if ruta_actual == [movimiento[2], movimiento[3]]:
11                     contador +=1
12                 rutas_contadas.append(ruta_actual)
13                 conteo_rutas.append([ruta[2], ruta[3], contador])
14                 contador = 0
15                 conteo_rutas.sort(reverse = True, key = lambda x:x[2])
16
17     return conteo_rutas
18
19 ##### Almacenamos las rutas obtenidas #####
20
21 resnha2015e = rutas(anhose('2015',e2015))
22 resnha2016e = rutas(anhose('2016',e2016))
23 resnha2017e = rutas(anhose('2017',e2017))
24 resnha2018e = rutas(anhose('2018',e2018))
25 resnha2019e = rutas(anhose('2019',e2019))
26 resnha2020e = rutas(anhose('2020',e2020))
27 resnha2015i = rutas(anhosi('2015',i2015))
28 resnha2016i = rutas(anhosi('2016',i2016))
29 resnha2017i = rutas(anhosi('2017',i2017))
30 resnha2018i = rutas(anhosi('2018',i2018))
31 resnha2019i = rutas(anhosi('2019',i2019))
32 resnha2020i = rutas(anhosi('2020',i2020))

```

Finalmente comprobamos si las rutas principales generan ingresos significativos para la empresa, creando dos funciones, la primera donde se almacenan los ingresos de estas rutas y la segunda donde se promedian estos ingresos con el conteo de rutas.

```

1 valoresE_S_V = []
2 valoresE_N_B = []
3 valoresI_J_M = []
4 valoresI_C_J = []
5
6 def valores(direccion, origen, destino, valor):
7     for x in direccion:
8         if (x[2]== origen and x[3]==destino):
9             valor.append(int(x[9]))
10     return valor
11
12 Exp_South_Korea_Vietnam = valores(exportaciones, 'South Korea', 'Vietnam',
13     valoresE_S_V)
14 Exp_Netherlands_Belgium = valores(exportaciones, 'Netherlands', 'Belgium',
15     valoresE_N_B)

```

```

15 Imp_Japan_Mexico = valores(importaciones, 'Japan', 'Mexico', valoresI_J_M)
16 Imp_China_Japan = valores(importaciones, 'China', 'Japan', valoresI_C_J)
17
18
19 def promedio(lista):
20     suma=0
21     for valor in lista:
22         suma+=valor
23     return int(suma/(len(lista)))
24
25 PromedioExp1 = promedio(Exp_South_Korea_Vietnam)
26 PromedioExp2 = promedio(Exp_Netherlands_Belgium)
27 PromedioImp1 = promedio(Imp_Japan_Mexico)
28 PromedioImp2 = promedio(Imp_China_Japan)
29
30 listaPromedio = [['exportación', 'South Korea', 'Vietnam', PromedioExp1],
31                  ['exportación', 'Netherlands', 'Belgium', PromedioExp2],
32                  ['importación', 'Japan', 'Mexico', PromedioImp1],
33                  ['importación', 'China', 'Japan', PromedioImp2]]
34 input('Dado lo anterior, las rutas de South Korea–Vietnam, Netherlands–Belgium son
    las más solicitadas de exportación, mientras que las de Japan–Mexico, China–
    Japan son las más solicitadas de importación.\n ')
35
36 print('Estás rutas tuvieron en promedio las siguientes ganancias por exportación e
    importación:')
37 for i in listaPromedio:
38     print('La ruta de', i[0], 'de origen en', i[1], 'hacia', i[2], 'en promedio tuvo
    ganancias de', i[3], 'por viaje. \n' )

```

3. Opción Sugerida

A través del análisis de datos de la empresa Synergy Logistics se obtuvieron resultados sobre las rutas más demandas, esta opción se eligió por la confianza que se genera en los clientes, además se analiza el aspecto monetario de estas rutas, para saber si la confianza en estas rutas se ve reflejado en las ganancias. En el detalle de las dos rutas más demandadas se tiene:

- 1.- La ruta de exportación de origen en South Korea hacia Vietnam en promedio tuvo ganancias de \$ 78470 por viaje.
- 2.- La ruta de exportación de origen en Netherlands hacia Belgium en promedio tuvo ganancias de \$ 18306 por viaje.
- 3.- La ruta de importación de origen en Japan hacia Mexico en promedio tuvo ganancias de \$ 14563 por viaje.
- 4.- La ruta de importación de origen en China hacia Japan en promedio tuvo ganancias de \$ 109523 por viaje.

Además la ruta de exportación con USA hacía diferentes países es muy solicitada, por lo que USA es una gran opción para continuar, este aspecto se visualiza más en las exportaciones de año con año. USA exporta entre otros con Canada, Belgium, Mexico, Netherlands. Otro país que

igual recibe muchas importaciones es Thailand, aunque no tiene una ruta específica que importe demasiado, tiene diversos países que importan hacia él, como Singapore, Malaysia, China.

4. Conclusión.

La consigna para verificar cuales son las rutas más demandadas es la óptima, ya que al obtener estas rutas es fácilmente verificable saber los ingresos que se reciben. Es importante para cualquier empresa mantener una relación estable con sus clientes y esta consigna prioriza esa situación y notamos cuales son las rutas de mayor frecuencia y a su vez, cual es el ingreso que están dando. Se observa que las cuatro rutas mostradas en la sección anterior cumplen con una cierta cantidad elevada de ingresos para la empresa, por lo que es importante apoyar aún más estas rutas para que se logren mejores resultados en los siguientes años.