

# Proyecto 2.

## INTRODUCCIÓN AL ANÁLISIS DE DATOS

---

5 MARZO

---

*EMTCH*

Creado por: Quintero Fernández Raúl



## Objetivo.

Poner en práctica las herramientas esenciales de Python para el análisis de datos a partir de la clasificación y manejo de los mismos, mediante la creación y el uso de archivos y estructuras de datos, implementación y creación de funciones, uso de módulos importados y destructuring.

## Descripción del 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.

## Análisis del caso

### 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? ¿porqué?

Python se realiza de la siguiente manera:

```
1 import pandas as pd
2 sldb=pd.read_csv("synergy_logistics_database.csv")
3 rutas_ventas=sldb.groupby(["origin","destination", "transport_mode"]).sum()
4 rutas_ventas
5
6 orden_rutas_ventas=pd.DataFrame(rutas_ventas)
7 orden_rutas_ventas=orden_rutas_ventas.sort_values(by=['total_value'],
8 ascending=False)
9
10 top10_rutas_ventas=orden_rutas_ventas[0:10]
11 t=top10_rutas_ventas["total_value"]
12 top10_rutas_ventas=t.reset_index()
13 top10_rutas_ventas
```

## Resultado de Opción 1.

*¿le conviene implementar esa estrategia?  
¿porqué?*

- Para esto calculamos el porcentaje de la suma del top 10 rutas con respecto a las ventas totales

	origin	destination	transport_mode	total_value
0	China	Mexico	Air	12250000000
1	Canada	Mexico	Rail	8450000000
2	South Korea	Vietnam	Sea	6877007000
3	France	United Kingdom	Sea	5427000000
4	South Korea	Japan	Sea	4741000000
5	China	South Korea	Rail	4535000000
6	USA	Mexico	Rail	4370000000
7	France	Belgium	Road	4257000000
8	China	Germany	Air	4090000000
9	China	Japan	Air	4018000000

```
#Ventas totales
1 rutas_ventas['total_value'].sum()
2
3 #Ventastop 10
4 top10_rutas_ventas['total_value'].sum()
5
6 #Porcentaje
7 porcentaje=(top10_rutas_ventas['total_value'].sum()/
8             rutas_ventas['total_value'].sum())*100
9 porcentaje
```

Para calcular el porcentaje realizamos lo siguiente:

4. Calculamos las ventas totales.
5. Realizamos lo mismo el punto anterior pero solo del top 10.
- 6.

## CONCLUSIÓN

27.360865990986806

*El top 10 de las rutas de comercio es sólo del 27.36% por lo que no es conveniente enfocarnos sólo en estas.*

## 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 medio de transporte que podrían reducir?

Python se realiza de la siguiente manera:

```
1 tipos_transporte_sum=sldb.groupby("transport_mode").sum()
2 transporte_sum=tipos_transporte_sum.sort_values(by=['total_value'], ascending=False)
3
4 #Total medios de transporte
5 total_transporte = transporte_sum['total_value'].sum()
6 total_transporte
7
8 transporte_sum["Procentaje %"] = (transporte_sum['total_value']/total_transporte)*100
9 transporte_sum
10
11 #Eliminar filas innecesarias
12 transporte_sum.drop(["register_id", "year"], axis=1)
13
14 #Top 3 medios de transporte
15 top3_medios_de_transporte=transporte_sum.drop(["register_id", "year"], axis=1)[0:3]
16 top3_medios_de_transporte
```

## Resultado de Opción 2.

**Los métodos de transporte con mejores ventas son: mar, ferrocarril, aire. Por aire, es donde obtenemos casi el 50% de las ganancias.**

**¿Cuál es medio de transporte que podrían reducir?**

	total_value	Procentaje %
transport_mode		
Sea	100530622000	46.608566
Rail	43628043000	20.227076
Air	38262147000	17.739310

```
#Columnas innecesarias
1 min_transporte=transporte_sum.drop(["register_id",
2                                     "year"], axis=1)
3
4 #Valor mínimo del medio de transporte
5 min_medios_de_transporte=min_transporte.iloc[-1]
6
7 #Formato de tabla
8 min_medios_de_transporte.reset_index()
```

**Para calcular el valor mínimo realizamos lo siguiente:**

1. Eliminamos las columnas innecesarias.
2. Extraemos el valor mínimo
3. Le damos otra vez formato de tabla.

## CONCLUSIÓN

	register_id	year	total_value	Procentaje %
transport_mode				
Road	23256105	5241426	33270486000	15.425048

*Sólo queda el método de transporte por tierra o carretera el porcentaje de este con respecto a las ventas es de 15.42% así que es una buena cantidad, por lo que se recomendaría no eliminar algún tipo de transporte.*

### 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?

Python se realiza de la siguiente manera:

```
1 paises=sldb.groupby("origin").sum().drop(["register_id","year"], axis=1)
2
3 #Porcentaje
4 paises['Porcentaje']=(paises['total_value']/total_ventas)*100
5 ordenamineto=paises.sort_values(by=['total_value'], ascending=False)
6 ordenamineto["sum_acumulada"]=ordenamineto.cumsum()["Porcentaje"]
7
8 paises_ventas_80 = ordenamineto[ordenamineto["sum_acumulada"]<80]
9 paises_ventas_80
```

### Resultado de Opción 3.

**Los 8 países de lista logran entre si el 78%, de las ventas totales. Este es un gran resultado, pero para este caso solo es casi el 80% así que perderíamos el 20% de las ganancias, por lo que se recomienda trabajar con los países que mínimo acumulen el 90% de nuestras ventas totales**

**¿Cuántos países se necesitan para tener este acumulado?**  
**R: 13 países**

```
1 paises_ventas_90 =
2 ordenamineto[ordenamineto["sum_acumulada"]<91]
3 paises_ventas_90.count()
```

```
total_value      12
Porcentaje        12
sum_acumulada     12
dtype: int64
```

	total_value	Porcentaje	sum_acumulada
origin			
China	45210046000	20.960533	20.960533
USA	23646306000	10.963032	31.923565
Japan	20042976000	9.292436	41.216001
France	19930332000	9.240211	50.456213
South Korea	18510146000	8.581777	59.037990
Germany	15593233000	7.229421	66.267411
Russia	14074000000	6.525066	72.792477
Canada	11253000000	5.217178	78.009656

**Para calcular el valor mínimo realizamos lo siguiente:**

1. Extraemos los países que cumplen la condición de acumular menos o igual del 90%.
2. Contamos los países que cumplen esta condición.

### Conclusión: sugerencia a Synergy logistics

Al analizar las 3 opciones sugerimos enfocarnos en una variante de la **opción 3**: trabajar en los 12 de los 23 países que hay en total, los cuales logran acumular el 90% de nuestras ventas.

	total_value	Porcentaje	sum_acumulada
origin			
China	45210046000	20.960533	20.960533
USA	23646306000	10.963032	31.923565
Japan	20042976000	9.292436	41.216001
France	19930332000	9.240211	50.456213
South Korea	18510146000	8.581777	59.037990
Germany	15593233000	7.229421	66.267411
Russia	14074000000	6.525066	72.792477
Canada	11253000000	5.217178	78.009656
Italy	6634684000	3.076009	81.085665
Spain	6419000000	2.976013	84.061677
Mexico	6040755000	2.800648	86.862326
Netherlands	4120369000	1.910308	88.772634
Singapore	4017684000	1.862701	90.635335