



EMTECH  
Emerging Technologies Institute



# REPORTE-02- GLORIA SALVADOR

Introducción al análisis de datos

Salvador Gloria  
sal\_glo\_a@yahoo.fr

# Proyecto 2

## Opción 1

In [269]:

```

# IMPORTAR LIBRERIAS
import csv
import pandas as pd
df = pd.read_csv('synergy_logistics_database.csv')

#CAMPOS DATAFRAME
df = pd.DataFrame(df,columns=['direction','origin','destination','year','product','transport_mode','total_value'])

#FUSIÓN DE CAMPOS (origin y destination)
df["origin-destination"] = df["origin"] + "-" + df["destination"]

#GENERA NUEVA BASE CSV
df.to_csv('synergy_logistics_Rutas.csv')

#CARGA DE CSV
df = pd.read_csv('synergy_logistics_Rutas.csv')
df
df1=df.groupby(['origin-destination'])['total_value'].sum()
#IMPRESIÓN DE DATAFRAME
print(df)

```

	Unnamed: 0	direction	origin	destination	year	product	\
0	0	Exports	Japan	China	2015	Cars	
1	1	Exports	Japan	China	2015	Cars	
2	2	Exports	Japan	China	2015	Cars	
3	3	Exports	Japan	China	2015	Cars	
4	4	Exports	Japan	China	2015	Cars	
...	...	...	...	...	...	...	
19051	19051	Imports	Japan	Singapore	2020	Gas turbines	
19052	19052	Imports	Malaysia	Singapore	2020	Gas turbines	
19053	19053	Imports	Malaysia	Singapore	2020	Gas turbines	
19054	19054	Imports	Malaysia	Singapore	2020	Gas turbines	
19055	19055	Imports	Malaysia	Singapore	2020	Gas turbines	

	transport_mode	total_value	origin-destination
0	Sea	33000000	Japan-China
1	Sea	16000000	Japan-China
2	Sea	29000000	Japan-China
3	Sea	14000000	Japan-China
4	Sea	17000000	Japan-China
...	...	...	...
19051	Sea	1000000	Japan-Singapore

REPORTE-02-GLORIA SALVADOR

19052	Sea	2000000	Malaysia-Singapore
19053	Sea	33000000	Malaysia-Singapore
19054	Sea	13000000	Malaysia-Singapore
19055	Sea	30000000	Malaysia-Singapore

[19056 rows x 9 columns]

In [270]:

```
#AGRUPACIÓN DATAFRAME POR ORIGEN & DESTINO CON TIPO DE DIRECCIÓN
```

```
print(df1)
```

```
origin-destination
```

```
Australia-Brazil          172000000
```

```
Australia-Japan          920000000
```

```
Australia-Mexico         84000000
```

```
Australia-Philippines    344000000
```

```
Australia-Singapore      493000000
```

```
...
```

```
United Kingdom-Ireland   584123000
```

```
United Kingdom-Italy     77000
```

```
United Kingdom-Spain     1378025000
```

```
United Kingdom-USA       111000
```

```
Vietnam-United Arab Emirates 540000000
```

```
Name: total_value, Length: 172, dtype: int64
```

In [271]:

```
#CREACION DE LA FIGURA
```

```
#LIBRERIAS PARA GRAFICAR
```

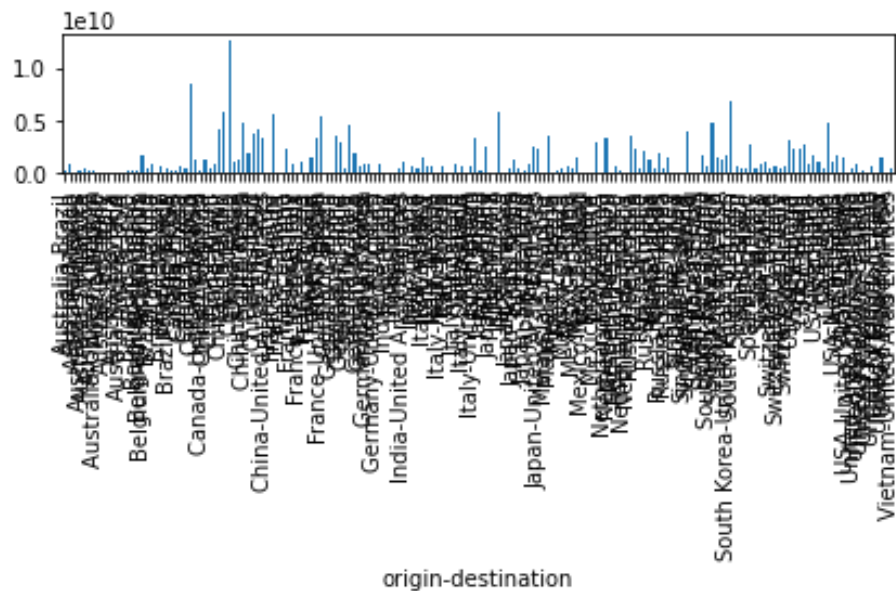
```
import matplotlib.pyplot as plt
```

```
import numpy as np
```

```
%matplotlib inline
```

```
df1.plot(kind="bar")
```

```
plt.tight_layout()
```



## Opción 2

In [272]:

```
# 02
import csv
#definir lista vacia para agregar los valores
df2 = pd.read_csv('synergy_logistics_Rutas.csv')
df2
```

Out[272]:

	Unnamed: 0	direction	origin	destination	year	product	transport_mode	total_value	origin-destination
	0	Exports	Japan	China	2015	Cars	Sea	33000000	Japan-China
	1	Exports	Japan	China	2015	Cars	Sea	16000000	Japan-China
	2	Exports	Japan	China	2015	Cars	Sea	29000000	Japan-China
	3	Exports	Japan	China	2015	Cars	Sea	14000000	Japan-China
	4	Exports	Japan	China	2015	Cars	Sea	17000000	Japan-China
	...	...	...	...	...	...	...	...	...
	19051	Imports	Japan	Singapore	2020	Gas turbines	Sea	1000000	Japan-Singapore
	19052	Imports	Malaysia	Singapore	2020	Gas turbines	Sea	2000000	Malaysia-Singapore
	19053	Imports	Malaysia	Singapore	2020	Gas turbines	Sea	33000000	Malaysia-Singapore
	19054	Imports	Malaysia	Singapore	2020	Gas turbines	Sea	13000000	Malaysia-Singapore
	19055	Imports	Malaysia	Singapore	2020	Gas turbines	Sea	30000000	Malaysia-Singapore

19056 rows x 9 columns

In [273]:

```
# CREAR ESTRUCTURA DE LECTURA
# RECORRER DF E IMPRIMIR POR TIPO DE TRANSPORTE
FromTo = {}
for index, row in df2.iterrows():
```

```

if row["transport_mode"] in FromTo:
    FromTo[row["transport_mode"]] = FromTo.get(row["transport_mode"]) + row["total_value"]
else:
    FromTo[row["transport_mode"]] = row["total_value"]

```

```
print(FromTo)
```

```
{'Sea': 100530622000, 'Air': 38262147000, 'Rail': 43628043000, 'Road': 33270486000}
```

In [274]:

```
#GROUP BY
```

```
print(df2.groupby(['transport_mode'])['total_value'].sum())
```

```
df3 = df2.groupby(['transport_mode'])['total_value'].sum()
```

```
transport_mode
```

```
Air      38262147000
```

```
Rail     43628043000
```

```
Road     33270486000
```

```
Sea      100530622000
```

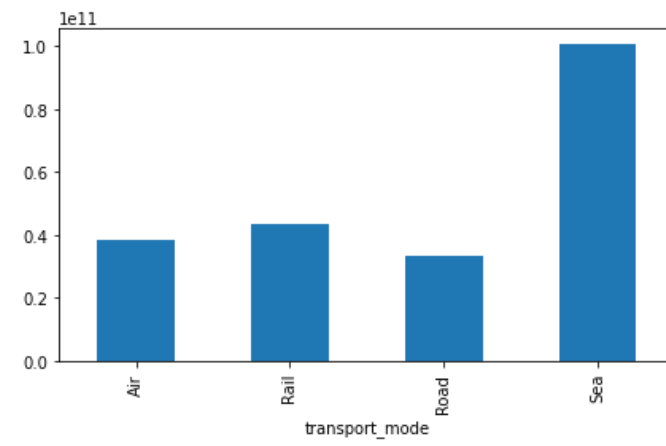
```
Name: total_value, dtype: int64
```

In [275]:

```
#GRAFICA DE TRANSPORTE
```

```
df3.plot(kind="bar")
```

```
plt.tight_layout()
```



## OPCIÓN 3

In [276]:

```
#CREAR GROUP BY PARA ORIGEN POR PAIS
df4=df.groupby(['origin'])['total_value'].sum()
```

In [208]:

```
#LEER CSV
totales = []
with open("synergy_logistics_Rutas.csv","r") as archivo_csv:
    lector = csv.reader(archivo_csv)
#ITERAR LOS DATOS
    for linea in lector:
        if linea[7] == "total_value":
#EXCLUIR TITULO DE CABECERA
            continue
        totales.append(int(linea[7]))
#IMPRIMIR EL TOTAL DE VENTA Y EL TOTAL POR PAIS
print("El total de las exp es:", sum(totales))
print(df4)
```

El total de las exp es: 215691298000

origin	
Australia	2570000000
Austria	1155000
Belgium	2588000000
Brazil	2763000000
Canada	11253000000
China	45210046000
France	19930332000
Germany	15593233000
India	2626000000
Italy	6634684000
Japan	20042976000
Malaysia	3560000000
Mexico	6040755000
Netherlands	4120369000
Russia	14074000000
Singapore	4017684000
South Korea	18510146000
Spain	6419000000
Switzerland	2154000000
USA	23646306000
United Arab Emirates	371000000
United Kingdom	3025612000
Vietnam	540000000
Name: total_value, dtype: int64	

In [277]:

```
#CREAR GROUP BY PARA OBTENER % PARTICIPACIÓN
```

REPORTE-02-GLORIA SALVADOR

```
df5=df.groupby(['origin'])['total_value'].sum()/sum(totales)
df5
```

Out[277]:

origin	
Australia	0.011915
Austria	0.000005
Belgium	0.011999
Brazil	0.012810
Canada	0.052172
China	0.209605
France	0.092402
Germany	0.072294
India	0.012175
Italy	0.030760
Japan	0.092924
Malaysia	0.016505
Mexico	0.028006
Netherlands	0.019103
Russia	0.065251
Singapore	0.018627
South Korea	0.085818
Spain	0.029760
Switzerland	0.009986
USA	0.109630
United Arab Emirates	0.001720
United Kingdom	0.014028
Vietnam	0.002504
Name: total_value, dtype: float64	

In [278]:

```
#IMPRIME paRTICIPACION
df6 = df5.sort_values()
#CREA NUEVO DF
df6.to_csv('synergy_logistics_origen_total_value.csv')
#NOMBRAR COLUMNAS
df6 = pd.read_csv('synergy_logistics_origen_total_value.csv')
df6.rename(columns={'origin': 'Pais', '': 'total_value'}, inplace=True)
df6.sort_values('total_value',ascending=False)
```

Out[278]:

	Pais	total_value
22	China	0.209605
21	USA	0.109630

## REPORTE-02-GLORIA SALVADOR

	<b>Pais</b>	<b>total_value</b>
<b>20</b>	Japan	0.092924
<b>19</b>	France	0.092402
<b>18</b>	South Korea	0.085818
<b>17</b>	Germany	0.072294
<b>16</b>	Russia	0.065251
<b>15</b>	Canada	0.052172
<b>14</b>	Italy	0.030760
<b>13</b>	Spain	0.029760
<b>12</b>	Mexico	0.028006
<b>11</b>	Netherlands	0.019103
<b>10</b>	Singapore	0.018627
<b>9</b>	Malaysia	0.016505
<b>8</b>	United Kingdom	0.014028
<b>7</b>	Brazil	0.012810
<b>6</b>	India	0.012175
<b>5</b>	Belgium	0.011999
<b>4</b>	Australia	0.011915
<b>3</b>	Switzerland	0.009986
<b>2</b>	Vietnam	0.002504



	Pais	total_value
1	United Arab Emirates	0.001720
0	Austria	0.000005

In [279]:

```
df6.plot(kind="bar")
```

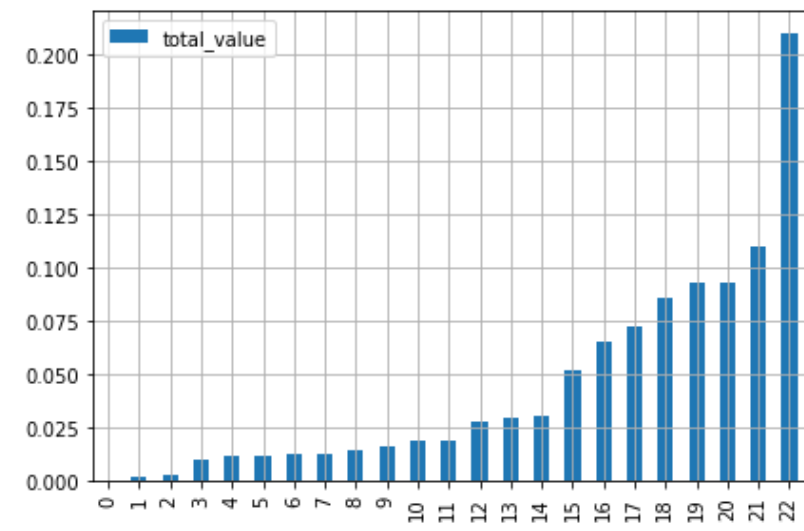
```
plt.tight_layout()
```

```
plt.grid(True)
```

```
plt.legend()
```

Out[279]:

<matplotlib.legend.Legend at 0x16c9c7d8bb0>



In [ ]:

## Resultados de Estrategia Operativa.

### Elaborado con Jupyter Notebook

El análisis de información y toma de decisiones va más allá de convertir datos y/o graficar, se deberá tener un contexto total de la empresa ya que para estas 3 opciones no se tiene información financiera; gastos, costos, ebitda, inversión, todos estos importantes para evaluar break even, ROI, productividad e incrementales, así que difícilmente se puede tomar una decisión y contar una historia con la misma. Sin embargo, en consideración que se tiene que dar una respuesta, comparto los resultados.

### Opción 1.

Se deberá continuar por lo menos con las 10 rutas con mayor ingreso

China-Mexico \$12,494,000,000

Canada-Mexico \$8,450,000,000

South Korea-Vietnam \$6,877,007,000

China-Japan \$5,891,000,000

Japan-Mexico \$5,829,000,000

France-Belgium \$5,538,069,000

France-United Kingdom \$5,427,000,000

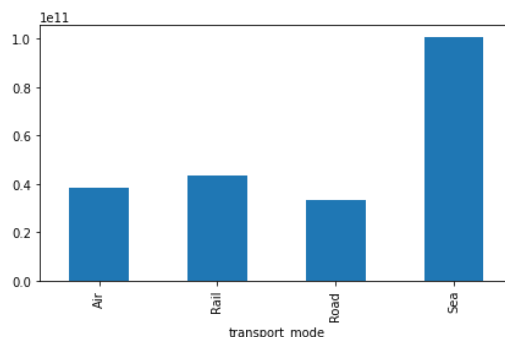
China-South Korea \$4,790,000,000

South Korea-Japan \$4,741,000,000

USA-Mexico \$4,710,000,000

### Opción 2.

EL medio de transporte más usado es "Sea", pero; se deberá invertir en Air y Road, para reducir los tiempos de entrega ya que los tiempos de espera con en el resto son muy amplios, pensemos en "experiencia al cliente"



### Opción 3.

Listado de países que generan  $\cong 80\%$ . Se deberá buscar y cumplir con los tratados y normas internacionales.

Out[278]:

	Pais	total_value
22	China	0.209605
21	USA	0.109630
20	Japan	0.092924
19	France	0.092402
18	South Korea	0.085818
17	Germany	0.072294
16	Russia	0.065251
15	Canada	0.052172
14	Italy	0.030760