

# Datos Sintéticos - Aeropuertos Combustibles - Ingresos Gastos

May 1, 2024

## 1 Aeropuertos y Combustibles - Ingresos y Gastos

## 2 Generación de Datos Sintéticos

Creado por:

- V. D. Betancourt

### 2.1 Descripción

Este proyecto permite generar **datos sintéticos** para los **Ingresos y Gastos** correspondientes a **Aeropuertos y Combustibles**.

### 2.2 Generación de Datos Sintéticos

```
[1]: import pandas as pd
import numpy as np
from datetime import datetime, timedelta

# Función para generar fechas mensuales
def generar_fechas(inicio, fin):
    return pd.date_range(start=inicio, end=fin, freq='M')

# Crear las fechas para el DataFrame
fechas = generar_fechas('2023-01-31', '2024-03-31')

# Información de las combinaciones División-Concepto-Subconcepto
data_info = [
    ("Aeropuertos", "Ingresos Cobrados", "Arrendamiento"),
    ("Aeropuertos", "Ingresos Cobrados", "Servicios Aeroportuarios"),
    ("Aeropuertos", "Ingresos Cobrados", "Tarifa de Uso Aeroportuario (TUA)"),
    ("Aeropuertos", "Gastos", "Servicios Personales"),
    ("Aeropuertos", "Gastos", "Materiales"),
    ("Aeropuertos", "Gastos", "Servicios Generales"),
    ("Combustibles", "Ingresos Cobrados", "Almacenamiento"),
    ("Combustibles", "Ingresos Cobrados", "Expendio"),
    ("Combustibles", "Gastos", "Servicios Personales"),
    ("Combustibles", "Gastos", "Materiales"),
    ("Combustibles", "Gastos", "Servicios Generales")
]
```

```

]

# Crear el DataFrame
df = pd.DataFrame({
    "Fecha": np.repeat(fechas, len(data_info)),
    "División": np.tile([x[0] for x in data_info], len(fechas)),
    "Concepto": np.tile([x[1] for x in data_info], len(fechas)),
    "Subconcepto": np.tile([x[2] for x in data_info], len(fechas))
})

# Diccionario para nombres de aeropuertos
nombres_aeropuertos = {
    'A1': 'Primer_Aeropuerto',
    'A2': 'Segundo_Aeropuerto',
    'A3': 'Tercer_Aeropuerto',
    'A4': 'Cuarto_Aeropuerto',
    'A5': 'Quinto_Aeropuerto',
    'A6': 'Sexto_Aeropuerto',
    'A7': 'Séptimo_Aeropuerto',
    'A8': 'Octavo_Aeropuerto',
    'A9': 'Noveno_Aeropuerto',
    'A10': 'Décimo_Aeropuerto',
    'A11': 'Undécimo_Aeropuerto'
}

# Generar datos sintéticos y renombrar columnas
for key, nombre in nombres_aeropuertos.items():
    df[nombre] = np.random.poisson(lam=100000, size=len(df))

# Eliminar las columnas de aeropuertos con etiquetas originales
df.drop(columns=[f'A{i}' for i in range(1, 12)], inplace=True, errors='ignore')

# Guardar el DataFrame como CSV
df.to_csv('datos_sinteticos_aeropuertos.csv', index=False)

df.head(30)

```

```

[1]:
      Fecha      División      Concepto \
0  2023-01-31  Aeropuertos  Ingresos Cobrados
1  2023-01-31  Aeropuertos  Ingresos Cobrados
2  2023-01-31  Aeropuertos  Ingresos Cobrados
3  2023-01-31  Aeropuertos      Gastos
4  2023-01-31  Aeropuertos      Gastos
5  2023-01-31  Aeropuertos      Gastos
6  2023-01-31  Combustibles  Ingresos Cobrados
7  2023-01-31  Combustibles  Ingresos Cobrados
8  2023-01-31  Combustibles      Gastos

```

9	2023-01-31	Combustibles	Gastos
10	2023-01-31	Combustibles	Gastos
11	2023-02-28	Aeropuertos	Ingresos Cobrados
12	2023-02-28	Aeropuertos	Ingresos Cobrados
13	2023-02-28	Aeropuertos	Ingresos Cobrados
14	2023-02-28	Aeropuertos	Gastos
15	2023-02-28	Aeropuertos	Gastos
16	2023-02-28	Aeropuertos	Gastos
17	2023-02-28	Combustibles	Ingresos Cobrados
18	2023-02-28	Combustibles	Ingresos Cobrados
19	2023-02-28	Combustibles	Gastos
20	2023-02-28	Combustibles	Gastos
21	2023-02-28	Combustibles	Gastos
22	2023-03-31	Aeropuertos	Ingresos Cobrados
23	2023-03-31	Aeropuertos	Ingresos Cobrados
24	2023-03-31	Aeropuertos	Ingresos Cobrados
25	2023-03-31	Aeropuertos	Gastos
26	2023-03-31	Aeropuertos	Gastos
27	2023-03-31	Aeropuertos	Gastos
28	2023-03-31	Combustibles	Ingresos Cobrados
29	2023-03-31	Combustibles	Ingresos Cobrados

	Subconcepto	Primer_Aeropuerto	Segundo_Aeropuerto \
0	Arrendamiento	99846	99950
1	Servicios Aeroportuarios	100101	100222
2	Tarifa de Uso Aeroportuario (TUA)	100362	100247
3	Servicios Personales	99549	99811
4	Materiales	100041	99912
5	Servicios Generales	99594	99990
6	Almacenamiento	99754	100206
7	Expendio	99724	100043
8	Servicios Personales	100090	100419
9	Materiales	100343	99803
10	Servicios Generales	100307	100128
11	Arrendamiento	100412	100484
12	Servicios Aeroportuarios	99848	99672
13	Tarifa de Uso Aeroportuario (TUA)	100060	99789
14	Servicios Personales	100057	99532
15	Materiales	99568	99832
16	Servicios Generales	99699	99597
17	Almacenamiento	100124	99922
18	Expendio	100119	99573
19	Servicios Personales	99690	99734
20	Materiales	99526	99378
21	Servicios Generales	99788	99531
22	Arrendamiento	99744	100893
23	Servicios Aeroportuarios	100155	100015

24	Tarifa de Uso Aeroportuario (TUA)	100108	99914
25	Servicios Personales	99938	100355
26	Materiales	100227	99734
27	Servicios Generales	99727	99591
28	Almacenamiento	99303	100028
29	Expendio	100140	99642

	Tercer_Aeropuerto	Cuarto_Aeropuerto	Quinto_Aeropuerto	Sexto_Aeropuerto \
0	100156	99982	99948	99939
1	100312	100077	100333	99578
2	99697	100544	99974	100185
3	99888	99983	99817	99956
4	99983	99441	100118	99716
5	99920	100138	100297	100062
6	99681	100387	100133	100065
7	99706	99876	100931	100191
8	99549	99893	100107	99945
9	99291	100809	100118	99944
10	99684	99924	99075	100164
11	99747	100006	100271	100202
12	100277	100571	100486	100023
13	100337	100473	100225	99475
14	99654	100095	99693	100224
15	100272	100119	99260	99768
16	100525	99731	100016	99659
17	100280	99732	99522	99788
18	99528	100584	99772	99730
19	100249	99638	100390	99834
20	100395	99748	100163	99823
21	100379	99918	99618	100367
22	99899	100426	99953	99515
23	99751	99416	99921	99843
24	100131	99929	100088	100269
25	100257	99721	100161	99956
26	100486	100701	100608	100282
27	100157	100334	100033	100437
28	99684	99925	100119	100224
29	100074	100221	99923	99845

	Séptimo_Aeropuerto	Octavo_Aeropuerto	Noveno_Aeropuerto \
0	100289	99339	99748
1	99828	100045	99887
2	100610	99296	100355
3	100680	99726	99992
4	99889	100372	99445
5	99894	100128	99780
6	100600	99935	100053

7	99786	99685	99876
8	100582	99474	100278
9	99979	99389	100163
10	99763	100458	99847
11	100423	99952	99776
12	100136	100027	99970
13	99878	100084	100442
14	100117	99561	100068
15	100270	99531	99364
16	99903	100135	99807
17	99706	99826	100060
18	99973	100039	99980
19	99677	99956	99971
20	99975	99528	99908
21	99834	99561	100045
22	100117	99718	100138
23	100433	100218	100330
24	100099	99539	99624
25	100428	99763	99836
26	99930	99669	100352
27	100098	99983	99907
28	99998	100188	99471
29	100115	99669	100167

	Décimo_Aeropuerto	Undécimo_Aeropuerto
0	99958	100012
1	100450	100236
2	100027	100103
3	100190	99853
4	100632	100518
5	99781	99895
6	99664	100278
7	99711	100281
8	100179	100275
9	99621	99945
10	100059	99851
11	100205	100100
12	100032	99877
13	99904	100057
14	99959	100122
15	99408	100514
16	99698	100299
17	100210	100115
18	99594	100256
19	100143	100359
20	99538	100060
21	99250	100032

22	99609	100009
23	99888	100038
24	99977	100668
25	99823	100225
26	100235	99676
27	100340	99714
28	100673	100404
29	100157	100367

```
[2]: # Filas y Columnas
df.shape
```

```
[2]: (165, 15)
```

```
[3]: # Info
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 165 entries, 0 to 164
Data columns (total 15 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Fecha                 165 non-null   datetime64[ns]
1   División              165 non-null   object
2   Concepto              165 non-null   object
3   Subconcepto           165 non-null   object
4   Primer_Aeropuerto     165 non-null   int64
5   Segundo_Aeropuerto    165 non-null   int64
6   Tercer_Aeropuerto     165 non-null   int64
7   Cuarto_Aeropuerto     165 non-null   int64
8   Quinto_Aeropuerto     165 non-null   int64
9   Sexto_Aeropuerto      165 non-null   int64
10  Séptimo_Aeropuerto    165 non-null   int64
11  Octavo_Aeropuerto     165 non-null   int64
12  Noveno_Aeropuerto     165 non-null   int64
13  Décimo_Aeropuerto     165 non-null   int64
14  Undécimo_Aeropuerto   165 non-null   int64
dtypes: datetime64[ns](1), int64(11), object(3)
memory usage: 19.5+ KB
```

## 2.3 Carga de Datos

```
[4]: import pandas as pd

# Cargar los datos
df = pd.read_csv('datos_sinteticos_aeropuertos.csv')
```

```
[5]: # Asegurar formato en campo Fecha  
df['Fecha'] = pd.to_datetime(df['Fecha'])
```

## 2.4 Fin