

LIBRERIAS PANDA

Archivos

ASISTENCIA METODOS...

ASISTENCIA METODOS...

ASISTENCIA METODOS...

Actividad_Big_Data_Ra...

Hueto Nicole tesla (1)...

Hueto Nicole tesla.png

Combined_Flights_201...

Combined_Flights_201...

Combined_Flights_202...

Combined_Flights_202...

Combined_Flights_202...

Documento sin título...

Ejercicio No 2.pdf

Ejercicio No1.pdf

Envio parcial.pdf

Estrategia de emprendi...

Fisica.pdf

Flujograma (1).png

Flujograma.png

GLOSARIO.docx

GUÍA LABORATORIO VL...

graficos RC.xlsx

GruposAuditoriaT1190...

ISO 27000 A CIBERSEG...

Idea_Convocatoria_Dr...

LA INNOVATION EN EL...

17-28 GB de espacio disponible

+ Código + Texto

```
import pandas as pd
flights_file1 = "/content/drive/MyDrive/Combined_Flights_2018.parquet"
flights_file2 = "/content/drive/MyDrive/Combined_Flights_2019.parquet"
flights_file3 = "/content/drive/MyDrive/Combined_Flights_2020.parquet"
flights_file4 = "/content/drive/MyDrive/Combined_Flights_2021.parquet"
flights_file5 = "/content/drive/MyDrive/Combined_Flights_2022.parquet"
df1 = pd.read_parquet(flights_file1)
df2 = pd.read_parquet(flights_file2)
df3 = pd.read_parquet(flights_file3)
df4 = pd.read_parquet(flights_file4)
df5 = pd.read_parquet(flights_file5)

df = df1.concat([df2, df3, df4, df5])

df = df.reset_index()

df_agg = df.groupby(["Airline", "Year"])[["DepDelayMinutes", "ArrDelayMinutes"]].agg(
    ["mean", "sum", "max"])
df_agg = df_agg.reset_index()
df_agg.to_parquet("temp_pandas.parquet")

!ls -l ./temp_pandas.parquet
!cat temp_pandas.parquet

pd.read_parquet("temp_pandas.parquet")
```

	Airline	Year	DepDelayMinutes			ArrDelayMinutes		
			mean	sum	max	mean	sum	max
0	Air Wisconsin Airlines Corp	2019	16.868511	1742281.0	1690.0	17.610384	1811545.0	1707.0
1	Alaska Airlines Inc.	2019	9.836041	2576246.0	1117.0	10.787284	2815643.0	1087.0
2	Allegiant Air	2019	14.678433	1536876.0	1979.0	15.556524	1624770.0	1966.0
3	American Airlines Inc.	2019	14.895515	13814816.0	2315.0	15.251863	14096412.0	2350.0
4	Capital Cargo International	2019	11.525332	1367642.0	1182.0	12.489465	1474806.0	1190.0
5	Comair Inc.	2019	14.427466	4081732.0	1844.0	14.578732	4106304.0	1842.0
6	Communtair Aka Champlain Enterprises, Inc.	2019	30.572619	1683787.0	1388.0	31.969338	1750577.0	1420.0
7	Compass Airlines	2019	14.630234	1369068.0	1767.0	15.090585	1409853.0	1752.0
8	Delta Air Lines Inc.	2019	10.856695	10750245.0	1266.0	10.786294	10657128.0	1304.0

0 s completado a las 19:19

RAM del sistema
2.5 / 12.7 GB



Disco
30.4 / 107.7 GB



	Airline	Year	DepDelayMinutes			ArrDelayMinutes		
			mean	sum	max	mean	sum	max
0	Air Wisconsin Airlines Corp	2019	16.868511	1742281.0	1690.0	17.610384	1811545.0	1707.0
1	Alaska Airlines Inc.	2019	9.836041	2576246.0	1117.0	10.787284	2815643.0	1087.0
2	Allegiant Air	2019	14.678433	1536876.0	1979.0	15.556524	1624770.0	1966.0
3	American Airlines Inc.	2019	14.895515	13814816.0	2315.0	15.251863	14096412.0	2350.0
4	Capital Cargo International	2019	11.525332	1367642.0	1182.0	12.489465	1474806.0	1190.0
5	Comair Inc.	2019	14.427466	4081732.0	1844.0	14.578732	4106304.0	1842.0
6	Communtair Aka Champlain Enterprises, Inc.	2019	30.572619	1683787.0	1388.0	31.969338	1750577.0	1420.0
7	Compass Airlines	2019	14.630234	1369068.0	1767.0	15.090585	1409853.0	1752.0
8	Delta Air Lines Inc.	2019	10.856695	10750245.0	1266.0	10.786294	10657128.0	1304.0
9	Empire Airlines Inc.	2019	8.287515	71024.0	546.0	9.082982	77496.0	540.0
10	Endeavor Air Inc.	2019	14.395421	3645482.0	1506.0	14.636930	3695576.0	1511.0
11	Envoy Air	2019	13.117923	4149527.0	2672.0	14.720387	4633389.0	2649.0
12	ExpressJet Airlines Inc.	2019	21.653172	2787651.0	1839.0	23.432743	3004312.0	1844.0
13	Frontier Airlines Inc.	2019	18.826018	2511259.0	1022.0	18.400065	2448331.0	1020.0
14	GoJet Airlines, LLC d/b/a United Express	2019	21.314252	1653922.0	2976.0	21.517977	1664975.0	2973.0
15	Hawaiian Airlines Inc.	2019	5.036265	421898.0	1536.0	5.938021	496947.0	1507.0
16	Horizon Air	2019	7.615291	910370.0	575.0	8.499155	1010847.0	566.0
17	JetBlue Airways	2019	21.854736	6420069.0	1769.0	21.414981	6268679.0	1756.0
18	Mesa Airlines Inc.	2019	17.443382	3863308.0	2209.0	18.119389	3997880.0	2206.0
19	Peninsula Airways Inc.	2019	27.830157	33591.0	298.0	29.105983	34054.0	313.0
20	Republic Airlines	2019	12.832237	4136433.0	1436.0	14.326528	4599890.0	1449.0
21	SkyWest Airlines Inc.	2019	16.416314	13463873.0	2710.0	16.764599	13684154.0	2695.0
22	Southwest Airlines Co.	2019	11.793784	15692786.0	804.0	10.182261	13517583.0	809.0
23	Spirit Air Lines	2019	14.221530	2864074.0	1434.0	14.108004	2834171.0	1429.0
24	Trans States Airlines	2019	24.023597	1756197.0	1840.0	25.601927	1860492.0	1865.0
25	United Air Lines Inc.	2019	16.467488	10222473.0	1525.0	17.102921	10581697.0	1523.0

Librerías: ipynb

Archivo Editar Ver Insertar Entorno de ejecución Herramientas Ayuda Se han guardado todos los cambios

Archivos

ASISTENCIA METODOS...

ASISTENCIA METODOS...

ASISTENCIA.png

Actividad_Big_Data_Re...

Busto Nicolás Tesla (T)...

Busto Nicolás Tesla.png

Combined_Flights_201...

Combined_Flights_201...

Combined_Flights_202...

Combined_Flights_202...

Combined_Flights_202...

Documento sin título...

Ejercicio No 2.pdf

Ejercicio No1.pdf

Envío parcial.pdf

Estrategia de emprendi...

Fisica.pdf

Flujograma (1).png

Flujograma.png

GLSARIO.docx

GUÍA LABORATORIO VI...

Gráficos RC.xlsx

GrupoAuditoriaT1150...

ISO 27000 A OBREROS...

Idem_Convocatorio_Gr...

1 año de trabajo en PYME

77.28 GB de espacio disponible

+ Código + Texto

```
import pandas as pd

agg_pandas = pd.read_parquet("temp_pandas.parquet")
agg_polars = pd.read_parquet("temp_polars.parquet")
agg_spark = pd.read_parquet("temp_spark.parquet")
agg_dask = pd.read_parquet("temp_dask.parquet")

agg_pandas.shape, agg_polars.shape, agg_spark.shape, agg_dask.shape
((126, 8), (122, 8), (122, 8), (46, 8))

agg_pandas.shape, agg_polars.shape, agg_spark.shape, agg_dask.shape
((126, 8), (122, 8), (122, 8), (46, 8))

agg_pandas.sort_values(["Airline", "Year"]).head()
Airline Year DepDelayMinutes ArrDelayMinutes
0 Air Wisconsin Airlines Corp 2019 16.868511 1742281.0 1690.0 17.616384 1811545.0 1707.0
1 Alaska Airlines Inc. 2019 9.836041 2576246.0 1117.0 10.787284 2815643.0 1087.0
2 Allegiant Air 2019 14.678433 1536876.0 1979.0 15.556324 1524770.0 1966.0
3 American Airlines Inc. 2019 14.895515 13814816.0 2315.0 15.251983 14096412.0 2350.0
4 Capital Cargo International 2019 11.525332 1367642.0 1182.0 12.489465 1474806.0 1190.0

agg_polars.sort_values(["Airline", "Year"]).head()
Airline Year avg_dep_delay sum_dep_delay max_dep_delay avg_arr_delay sum_arr_delay max_arr_delay
98 Air Wisconsin Airlines Corp 2018 16.763460 1686774.0 1296.0 17.881934 1708887.0 1292.0
99 Air Wisconsin Airlines Corp 2019 16.868511 1742281.0 1690.0 17.616384 1811545.0 1707.0
100 Air Wisconsin Airlines Corp 2020 8.553725 433315.0 1460.0 9.982509 452450.0 1428.0
101 Air Wisconsin Airlines Corp 2021 16.553045 1291194.0 1421.0 17.327449 1346682.0 1416.0
93 Air Wisconsin Airlines Corp 2022 13.124801 510581.0 1355.0 13.340409 517261.0 1353.0
```

0 s completado a las 19:19

RAM del sistema
7.1 / 12.7 GB



Disco
30.4 / 107.7 GB



	FlightDate	Airline	Origin	Dest	Cancelled	Diverted	\
0	2020-09-01	Comair Inc.	PHL	DAY	False	False	
1	2020-09-02	Comair Inc.	PHL	DAY	False	False	
2	2020-09-03	Comair Inc.	PHL	DAY	False	False	
3	2020-09-04	Comair Inc.	PHL	DAY	False	False	
4	2020-09-05	Comair Inc.	PHL	DAY	False	False	
...
590537	2022-03-31	Republic Airlines	MSY	EW	False	True	
590538	2022-03-17	Republic Airlines	CLT	EW	True	False	
590539	2022-03-08	Republic Airlines	ALB	ORD	False	False	
590540	2022-03-25	Republic Airlines	EW	PIT	False	True	
590541	2022-03-07	Republic Airlines	EW	RD	False	True	
	CRSDepTime	DepTime	DepDelayMinutes	DepDelay	...	WheelsOff	\
0	1905	1858.0	0.0	-7.0	...	1914.0	
1	1905	1858.0	0.0	-7.0	...	1914.0	
2	1905	1855.0	0.0	-10.0	...	2000.0	
3	1905	1857.0	0.0	-8.0	...	1910.0	
4	1905	1856.0	0.0	-9.0	...	1910.0	
...
590537	1949	2014.0	25.0	25.0	...	2031.0	
590538	1733	1817.0	44.0	44.0	...	NaN	
590539	1700	2318.0	378.0	378.0	...	2337.0	
590540	2129	2322.0	113.0	113.0	...	2347.0	
590541	1154	1148.0	0.0	-6.0	...	1201.0	
	WheelsOn	TaxiIn	CRSArrTime	ArrDelay	ArrDel15	ArrivalDelayGroups	\
0	2030.0	4.0	2056	-22.0	0.0	-2.0	
1	2022.0	5.0	2056	-29.0	0.0	-2.0	
2	2117.0	5.0	2056	26.0	1.0	1.0	
3	2023.0	4.0	2056	-29.0	0.0	-2.0	
4	2022.0	4.0	2056	-30.0	0.0	-2.0	
...
590537	202.0	32.0	2354	NaN	NaN	NaN	
590538	NaN	NaN	1942	NaN	NaN	NaN	
590539	52.0	7.0	1838	381.0	1.0	12.0	
590540	933.0	6.0	2255	NaN	NaN	NaN	
590541	1552.0	4.0	1333	NaN	NaN	NaN	
	ArrTimeBlk	DistanceGroup	DivAirportLandings				
0	2000-2059	2	0.0				
1	2000-2059	2	0.0				
2	2000-2059	2	0.0				
3	2000-2059	2	0.0				
4	2000-2059	2	0.0				
...				
590537	2300-2359	5	1.0				
590538	1900-1959	3	0.0				
590539	1800-1859	3	0.0				
590540	2200-2259	2	1.0				
590541	1300-1359	2	1.0				



4	Allegiant Air	2020	12.825575	1080016.0	1648.0	13.331111	1115734.0	1645.0
5	Allegiant Air	2022	22.688601	1602632.0	1917.0	25.350068	1785963.0	1919.0
6	American Airlines Inc.	2020	7.624477	4084097.0	3890.0	7.861155	4202644.0	3864.0
7	American Airlines Inc.	2022	17.718716	8464195.0	2994.0	17.860139	8499122.0	2977.0
8	Capital Cargo International	2020	7.665063	512969.0	1482.0	8.427212	561522.0	1470.0
9	Capital Cargo International	2022	12.052814	619599.0	1512.0	13.050802	667418.0	1490.0
10	Comair Inc.	2020	10.068723	1798294.0	1919.0	10.686808	1903235.0	1888.0
11	Comair Inc.	2022	16.925615	2212601.0	1607.0	17.623038	2293374.0	1612.0
12	Commatair Aka Champlain Enterprises, Inc.	2020	12.266858	385670.0	1557.0	13.438158	421340.0	1555.0
13	Commatair Aka Champlain Enterprises, Inc.	2022	16.342795	700730.0	1464.0	17.008007	726497.0	1456.0
14	Compass Airlines	2020	8.215550	120670.0	1431.0	8.641498	126693.0	1412.0
15	Delta Air Lines Inc.	2020	5.581694	3083283.0	1195.0	6.209070	3424414.0	1193.0
16	Delta Air Lines Inc.	2022	13.842472	6948367.0	1287.0	13.111550	6565084.0	1285.0
17	Empire Airlines Inc.	2020	6.861561	32613.0	274.0	7.028136	33222.0	272.0
18	Endeavor Air Inc.	2020	5.653603	1156405.0	2579.0	5.877866	1200707.0	2560.0
19	Endeavor Air Inc.	2022	13.284602	1819220.0	1973.0	14.184149	1935952.0	1968.0
20	Envoy Air	2020	6.685444	1339382.0	2248.0	8.417206	1682490.0	2238.0
21	Envoy Air	2022	10.196954	1498942.0	5327.0	10.942435	1603023.0	5324.0
22	ExpressJet Airlines Inc.	2020	6.396004	308249.0	1357.0	7.443095	357797.0	1338.0
23	Frontier Airlines Inc.	2020	7.398021	640062.0	645.0	7.330406	633479.0	638.0
24	Frontier Airlines Inc.	2022	23.560713	1982257.0	1288.0	23.980928	2011856.0	1311.0
25	GoJet Airlines, LLC d/b/a United Express	2020	8.309550	304786.0	1408.0	8.336411	305196.0	1382.0
26	GoJet Airlines, LLC d/b/a United Express	2022	20.259624	656817.0	1220.0	21.749690	702254.0	1199.0
27	Hawaiian Airlines Inc.	2020	3.592000	137491.0	1484.0	4.092661	156487.0	1481.0
28	Hawaiian Airlines Inc.	2022	8.890739	372193.0	1847.0	9.069343	379289.0	1805.0
29	Horizon Air	2020	4.849894	445133.0	775.0	5.351232	489932.0	758.0
30	Horizon Air	2022	8.893327	492717.0	803.0	9.589591	529729.0	803.0
31	JetBlue Airways	2020	8.213825	1118838.0	1135.0	8.503231	1155334.0	1153.0

```
[33] agg_pandas.sort_values(['Airline', 'Year']).head()
```

	Airline	Year	DepDelayMinutes			ArrDelayMinutes		
			mean	sum	max	mean	sum	max
0	Air Wisconsin Airlines Corp	2019	16.868511	1742281.0	1690.0	17.610384	1811545.0	1707.0
1	Alaska Airlines Inc.	2019	9.836041	2576246.0	1117.0	10.787284	2815643.0	1087.0
2	Allegiant Air	2019	14.678433	1536876.0	1979.0	15.556524	1624770.0	1966.0
3	American Airlines Inc.	2019	14.895515	13814816.0	2315.0	15.251863	14096412.0	2350.0
4	Capital Cargo International	2019	11.525332	1367642.0	1182.0	12.489465	1474806.0	1190.0

```
[34] agg_polars.sort_values(['Airline', 'Year']).head()
```

	Airline	Year	avg_dep_delay	sum_dep_delay	max_dep_delay	avg_arr_delay	sum_arr_delay	max_arr_delay
98	Air Wisconsin Airlines Corp	2018	16.753459	1606774.0	1296.0	17.881934	1708887.0	1292.0
103	Air Wisconsin Airlines Corp	2019	16.868511	1742281.0	1690.0	17.610384	1811545.0	1707.0
104	Air Wisconsin Airlines Corp	2020	8.583725	433315.0	1460.0	8.982529	452450.0	1439.0
110	Air Wisconsin Airlines Corp	2021	16.553045	1290194.0	1421.0	17.327440	1346602.0	1416.0
93	Air Wisconsin Airlines Corp	2022	13.124801	510581.0	1355.0	13.340409	517261.0	1353.0

```
agg_dask.sort_values(['Airline', 'Year']).head()
```

	Airline	Year	DepDelayMinutes			ArrDelayMinutes		
			mean	sum	max	mean	sum	max
0	Air Wisconsin Airlines Corp	2020	8.583725	433315.0	1460.0	8.982529	452450.0	1439.0
1	Air Wisconsin Airlines Corp	2022	13.124801	510581.0	1355.0	13.340409	517261.0	1353.0
2	Alaska Airlines Inc.	2020	5.818328	772930.0	823.0	6.365082	843157.0	788.0
3	Alaska Airlines Inc.	2022	10.153994	1278134.0	915.0	11.026280	1382905.0	908.0
4	Allegiant Air	2020	12.825575	1080016.0	1648.0	13.331111	1115734.0	1645.0