EXPERIMENT NO: 3C

Airline Flight Dataset Generation and CSV Export

Aim:

To generate a synthetic dataset of airline flights with random flight details (IDs, airlines, source and destination cities, departure/arrival times, durations, and fares) and export it as a CSV file for analysis or testing purposes.

Algorithm:

1) **Define Data Sources: Lists of airlines and cities.**
2) **Generate Records: For 15 flights:**
3) **Assign a unique Flight_ID.**
4) **Randomly select airline, source, and destination (different cities).**
5) **Generate random departure time.**
6) **Compute random flight duration and arrival time.**
7) **Generate random fare.**
8) **Store Data: Append flight details to a list.**
9) **Create DataFrame: Convert the list into a pandas DataFrame with appropriate column names.**
10) **Export CSV: Save the DataFrame to "airline.flights".**
11) **Output: Print confirmation and show the first few rows.**

Program:

```
[7]: import pandas as pd
import random
from datetime import datetime,timedelta
airlines=["Air India","Indigo","Spicejet","Vistara","GoAir"]
city=["Chennai","Delhi","Mumbai","Bangalore","Hyderabad","Kolkata","Pune"]
data=[]
for i in range(1,16):
    flight_id=f"FL{i:03d}"
    airline=random.choice(airlines)
    source,destination=random.sample(city,2)
    dep_time=datetime(2024,8,1,random.randint(0,23),random.choice([0,15,30,45]))
    duration_hours=random.randint(1,5)
    duration_min=random.choice([0,15,30,45])
    duration=timedelta(hours=duration_hours,minutes=duration_min)
    arr_time=dep_time+duration
    fare=random.randint(3000,16000)
    data.append([
        flight_id,airline,source,destination,
        dep_time.strftime("%Y-%m-%d %H:%M"),
        arr_time.strftime("%Y-%m-%d %H:%M"),
        f"{duration_hours}h{duration_min}m",
        fare
    ])

df=pd.DataFrame(data,columns=[
    "Flight_ID","AirLine","Source","Destination","Departure_time","Arrival_time","Duration","Fare"
])
df.to_csv("airline.flights",index=True)
print("CSV file 'airline_flights.csv' created successfully!")
print(df.head().to_string())
```

```
CSV file 'airline_flights.csv' created successfully!
  Flight_ID    AirLine   Source Destination    Departure_time      Arrival_time Duration   Fare
0     FL001  Air India    Delhi     Kolkata  2024-08-01 00:00  2024-08-01 02:45    2h45m   7339
1     FL002    Vistara  Kolkata       Delhi  2024-08-01 03:15  2024-08-01 06:45    3h30m   5333
2     FL003  Air India  Kolkata   Bangalore  2024-08-01 09:00  2024-08-01 11:30    2h30m   7069
3     FL004   Spicejet     Pune     Kolkata  2024-08-01 22:15  2024-08-02 02:00    3h45m   5142
4     FL005      GoAir    Delhi     Kolkata  2024-08-01 06:00  2024-08-01 08:00     2h0m  13052
```

Result:

A CSV file "airline.flights" containing 15 randomly generated flight records with columns: Flight_ID, AirLine, Source, Destination, Departure_time, Arrival_time, Duration, and Fare. The dataset can be used for analysis, testing, or simulations.