Problem Statement

You are working for an airline company looking to enter the United States domestic market. Specifically, the company has decided to start with 5 round trip routes between medium and large US airports. An example of a round trip route is the combination of JFK to ORD and ORD to JFK. The airline company has to acquire 5 new airplanes (one per round trip route) and the upfront cost for each airplane is $90 million. The company's motto is "On time, for you", so punctuality is a big part of its brand image.

You have been tasked with analyzing 1Q2019 data to identify:

1. The 10 busiest round trip routes in terms of number of round trip flights in the quarter. Exclude canceled flights when performing the calculation.
2. The 10 most profitable round trip routes (without considering the upfront airplane cost) in the quarter. Along with the profit, show total revenue, total cost, summary values of other key components and total round trip flights in the quarter for the top 10 most profitable routes. Exclude canceled flights from these calculations.
3. The 5 round trip routes that you recommend to invest in based on any factors that you choose.
4. The number of round trip flights it will take to breakeven on the upfront airplane cost foreach of the 5 round trip routes that you recommend. Print key summary components for these routes.
5. Key Performance Indicators (KPI's) that you recommend tracking in the future to measure the success of the round trip routes that you recommend.

Datasets Information

1. **Flights dataset**
   Contains data about available routes from origin to destination. For occupancy, use the data provided in this dataset.
2. **Tickets dataset**
   Ticket prices data (sample data only as the data is huge). Consider only round trips in your analysis.
3. **Airport Codes dataset**
   Identifies whether an airport is considered medium or large sized. Consider only medium and large airports in your analysis.