Airlines Flight Data Analysis (EDA)

Flight Price Analyss (EDA)[¶](https://www.kaggle.com/code/ankurjat/airlines-flights-data-analysis-eda#Flight-Price-Analyss-(EDA))

This notebook performs an in-depth analysis of flight ticket prices using a dataset containing multiple airlines, routes, timings, and class information.  
The aim is to identify patterns affecting ticket prices and derive meaningful insights for travelers and airlines.

Dataset Description

These are the main Features/Columns available in the dataset:

1) **Airline**:  
The name of the airline company is stored in the airline column.  
It is a categorical feature having 6 different airlines.

2) **Flight**:  
Flight stores information regarding the plane's flight code.  
It is a categorical feature.

3) **Source City**:  
City from which the flight takes off.  
It is a categorical feature having 6 unique cities.

4) **Departure Time**:  
This is a derived categorical feature created by grouping time periods into bins.  
It stores information about the departure time and has 6 unique time labels.

5) **Stops**:  
A categorical feature with 3 distinct values that stores the number of stops between the source and destination cities.

6) **Arrival Time**:  
This is a derived categorical feature created by grouping time intervals into bins.  
It has six distinct time labels and keeps information about the arrival time.

7) **Destination City**:  
City where the flight will land.  
It is a categorical feature having 6 unique cities.

8) **Class**:  
A categorical feature that contains information on seat class;  
it has two distinct values: Business and Economy.

9) **Duration**:  
A continuous feature that displays the overall amount of time it takes to travel between cities in hours.

10) **Days Left**:  
This is a derived characteristic that is calculated by subtracting the trip date by the booking date.

11) **Price**:  
Target variable that stores information about the ticket price.

Objective

* Analyze how ticket prices vary based on:
  + Airline
  + Source and destination cities
  + Departure and arrival timings
  + Number of stops
  + Seat class
  + Days left before departure
* Visualize trends and patterns using plots and graphs.
* Draw key insights for travelers on how to minimize travel cost.

Analysis Plan[¶](https://www.kaggle.com/code/ankurjat/airlines-flights-data-analysis-eda#Analysis-Plan)

1. Import and understand the dataset.
2. Perform exploratory data analysis (EDA) with visualizations.
3. Identify relationships between features and price.
4. Derive actionable insights for travelers and airlines.

**Business Questions SQL + POWER BI**

1. **Cheapest Airline for Each Route**  
   → Helps customers and platform pricing team identify budget-friendly airlines.
2. **Price Impact of Stops**  
   → See if non-stop flights really save customers money.
3. **Top 5 Most Expensive Routes**  
   → Focus marketing on these premium routes.
4. **Price Change Trend by Days Left**  
   → Verify last-minute booking surge.
5. **Price Comparison Between Classes**  
   → Show value gap between Economy & Business.
6. **Busiest Departure Times**  
   → Help airports manage crowding.
7. **Fastest Flight Options for Each Route**  
   → Improve recommendations for time-conscious travelers.
8. **Price Ranges per Airline**  
   → Identify high and low-cost carriers.
9. **Booking Window Categories**  
   → Segment price averages for last-minute, medium, and early bookings.
10. **Airlines with Maximum Route Coverage**  
    → Identify the most versatile carriers.

**Overall Insights from Airline Dataset Analysis**

1. **Pricing Patterns**
   * Across all airlines, the **cheapest routes** are consistently *Delhi → Hyderabad* and *Mumbai → Chennai*, while the **most expensive** are *Delhi → Bangalore* and *Mumbai → Kolkata*.
   * **Non-stop flights** are generally cheaper than flights with one or more stops, contradicting the common belief that direct flights are always more expensive.
   * **Air India** offers the highest-priced tickets, while **AirAsia** is the most budget-friendly carrier overall.
2. **High-Value Routes**
   * Premium-priced routes include *Bangalore → Mumbai* (₹40,354), *Delhi → Chennai* (₹35,145), and *Kolkata → Delhi* (₹34,944).
   * These routes present opportunities for targeted marketing, loyalty programs, or premium cabin promotions.
3. **Booking Behavior & Timing**
   * **Early bookings** (11+ days in advance) have an average ticket price around ₹5,680 — significantly lower than **last-minute bookings** (≤3 days), which average ₹13,502.  
     → Price gap exceeds ₹7,800, confirming a strong *last-minute price surge*.
   * **Morning (38,235), Evening (38,121), and Early Morning (38,081)** departures are the busiest slots, requiring higher airport and ground staff allocation.
   * **Night and Late Night** flights have the lowest traffic, which could be leveraged for discounted fares or cargo operations.
4. **Travel Time Insights**
   * Fastest routes include *Mumbai → Hyderabad*, *Bangalore → Mumbai*, and *Bangalore → Hyderabad*.
   * Longest durations are seen in *Delhi → Bangalore*, *Delhi → Chennai*, and *Kolkata → Mumbai*, suggesting possible indirect routings or congested airspace.
5. **Airline Coverage**
   * Most airlines serve **23 unique city pairs**, with the exception of **SpiceJet**, which serves only 20, indicating slightly less market penetration.

**💡 Business Recommendations**

* **Promote Early Booking Campaigns:** Highlight cost savings of ₹7K+ to attract price-sensitive customers.
* **Target Premium Routes:** Use loyalty programs and exclusive offers for Bangalore–Mumbai, Delhi–Chennai, and Kolkata–Delhi travelers.
* **Optimize Scheduling:** Allocate more operational resources during Morning, Evening, and Early Morning peaks; explore cargo/low-fare sales for low-demand night slots.
* **Leverage Non-stop Pricing:** Market direct flights aggressively as both cheaper and faster on select routes.
* **Network Expansion for SpiceJet:** Closing the gap in route coverage could improve competitiveness.

**Short Summary**

**Designed and executed** SQL-based data analysis on airline pricing, routes, and booking trends, delivering 12+ actionable business insights.

**Developed interactive Power BI dashboard** with slicers, KPIs, and visuals to support pricing strategy and route optimization.

**Automated analysis in Python** using pandas and matplotlib, replicating SQL insights for reproducible reporting.

**Improved decision-making** for pricing, marketing, and operations by identifying cheapest/premium routes, busiest times, and booking window savings (~₹7,800).