

Dataset Documentation

The dataset we've chosen for analyzing is Ticket price of flights.

Size: The dataset has 11 columns and 582 rows including the label row. The target columns the last column and rest of the columns are the features.

Feature Description: The columns are:

- **Airline** (Name of the airline used for traveling)
- **Date of Journey** (Date at which a person traveled YYYY-MM-DD)
- **Source** (Starting location of flight)
- **Destination** (Ending location of flight)
- **Route** (This contains information on starting and ending location of the journey in the standard format used by airlines)
- **Dep_time** (Starting time of flight from starting location)
- **Arrival_time** (Arrival time of flight at destination)
- **Duration** (Duration of flight in hours/minutes)
- **Total_Stops** (Number of total stops flight took before landing at the destination.)
- **Additional_Info** (Shown any additional information about a flight)
- **Price** (Price of the flight)

Citations: Below are the websites from where the data was collected.

<https://www.yatra.com/>

https://www.makemytrip.com/international-flights/dhaka-india-cheap-airtickets.html?fbclid=IwAR1Ndxh3h4iOKgWofu9Zjo1G8A_ZK_rjewFIYzuLIRpdy6T3AVTLkr5ylOw