Report:

The flight data set analysis focuses on understanding the factors that contribute to flight prices. The data set contains various attributes such as the source, destination, airline, duration, total stops, additional information, and date of the flights. By conducting exploratory data analysis (EDA) and employing visualizations and statistical techniques, we aim to uncover insights and patterns in the data that can aid in predicting flight prices.

The initial steps of the analysis involve importing the necessary libraries and loading the flight data set. Data pre-processing techniques, such as handling missing values and duplicates, are applied to ensure the data's quality. The data is then explored through various visualizations to gain a better understanding of its distribution and relationships between variables.

The EDA reveals interesting findings. The bar plot of 'Additional-Info' against 'Price' highlights that flight prices tend to increase with an increase in priority. Flights with 'No check-in baggage included' have the lowest prices, while business flights have the highest prices. Additionally, the distribution of additional information shows that a large number of flights have no additional information available.

Price analysis is conducted by binning the continuous price variable into five categories. This categorical representation allows for a clearer understanding of the relationship between price and other variables. Bar plots, box plots, and scatter plots are employed to visualize these relationships. For example, the bar plot of price against the day of the week suggests that flight prices are generally higher on the first and sixth days. The box plot of price against the source reveals that flights from Delhi have the highest median price, followed by flights from Kolkata and Bangalore.

Further analysis examines the impact of variables such as airline, duration, total stops, and additional information on flight prices. These relationships are explored using scatter plots, bar plots, and heat-maps. The scatter plot of price against airline indicates that Jet Airways Business has comparatively higher prices than other airlines for certain flights. The box plot of price against the number of stops reveals that flights with more stops tend to have higher median prices. The bar plot of price against duration suggests a positive relationship between flight duration and price.