**Data Analytics on Indian Airflights**

Dataset chosen: DGCA\_FLIGHT\_TRAFFIC

The analysis of this aviation dataset provides valuable insights into airline performance, operational efficiency, and industry trends. By examining various key metrics such as departures, available seat kilometers (ASK), passenger demand, load factors, cargo transport trends, and ton kilometers, we can better understand airline operations and their impact on profitability and efficiency.

**GROUP MEMBERS:**

Sona S U - (Function 10 , 11, 12)

Meghana L - (Function 4 , 5 ,6)

Aruna Sanjana - (Function 13 ,14 ,15)

Noti Gayatri - (Function 1 , 2 , 3)

Dhanya - (Function 7 , 8 , 9)

Here's an outline of what this dataset typically encompasses:

* Month: The reporting period (e.g., January, February, etc.).
* No. Departure (AF): Number of departures (flights) for aircraft.
* Hours (AF): Total hours flown by aircraft.
* Kms (Thousands) (AF): Distance flown by aircraft in thousands of kilometers.
* No Carried (P): Number of passengers carried.
* Km Performed (Millions) (P): Passenger kilometers performed in millions.
* Avail Seats Km (Millions): Available seat kilometers in millions, representing airline capacity.
* PAX Load %: Passenger load factor as a percentage, indicating seat occupancy.
* Freight CC: Freight cargo capacity.
* Mail CC: Mail cargo capacity.
* Total CC: Total cargo capacity (freight + mail).
* PAX TON KMS Performed: Passenger ton kilometers performed.
* Freight TON KM Performed: Freight ton kilometers performed.
* Mail TON KMS Performed: Mail ton kilometers performed.
* Total TON KMS Performed: Total ton kilometers performed (passenger + freight + mail).
* Avail TONNE KMS (Millions): Available ton kilometers in millions, indicating overall cargo capacity.
* Weight Load Factor %: Percentage of total weight capacity utilized.

**Analytics:**

1. **Available Seat Kilometers (ASK) vs Passenger Demand**

* Plot Used: Line Plot (ASK vs Passenger Demand over time)
* Key Observations:
  + ASK (Available Seat Kilometers) represents the total seating capacity multiplied by the distance flown.
  + Passenger demand (Km Performed) represents actual usage by travelers.
  + Ideally, ASK should align with demand for efficient airline operations.
* Interpretation:
  + If ASK is significantly higher than passenger demand, it indicates unused capacity (low load factor), leading to inefficiencies.
  + If ASK and demand move together, it suggests that airlines optimize seat availability based on demand patterns.
  + Sudden divergence between the two lines may indicate seasonal variations, demand surges, or external factors like economic changes.

1. **Flight Hours vs Passengers Carried**

* Plot Used: Scatter Plot (Flight Hours vs Passengers Carried)
* Key Observations:
  + Flight hours (Hours AF) are the total hours of flights operated.
  + Passengers carried (No Carried (P)) indicates the total number of passengers flown.
  + A correlation matrix is provided to understand the relationship between these two metrics.
* Interpretation:
  + A strong positive correlation indicates that higher flight hours generally correspond to more passengers carried, suggesting efficient use of flights.
  + Low or no correlation might point to inefficiencies or mismatches between flight operations and passenger demand.

1. **Freight Trends Over Time**

* Plot Used: Heatmap (Freight Transport Trends)
* Key Observations:
  + Freight TON KM Performed measures the distance freight has been carried.
  + A heatmap aggregates freight transport performance over months and years.
  + Values indicate the volume of freight transported.
* Interpretation:
  + Seasonal patterns in the heatmap highlight peak and off-peak periods for freight transport.
  + An increase in certain months may indicate higher freight demand, while low values could suggest slower periods or inefficiencies.

1. **Flight Kilometers vs Load Factor**

* Plot Used: Scatter Plot (Flight Kilometers vs Load Factor)
* Key Observations:
  + Flight kilometers represent the distance traveled by the aircraft (Kms AF).
  + Load factor (PAX load %) represents how efficiently the seating capacity is used.
  + A scatter plot highlights this relationship.
* Interpretation:
  + A high load factor paired with long flight kilometers suggests an efficient operation.
  + Low load factors on long flights may indicate underutilization, which could be due to overcapacity or demand mismatches.
  + A correlation value helps to assess how strongly these two metrics are related.

1. **Total Ton Kilometers**

* Plot Used: Line Plot (Total Ton-Kilometers Over Time)
* Key Observations:
  + TONNE KMS represents the total ton-kilometers, indicating the weight of the cargo transported over distance.
  + The plot visualizes this performance over time.
* Interpretation:
  + A steady or increasing trend in ton-kilometers suggests that the airline is efficiently utilizing its cargo capacity.
  + Significant drops or fluctuations might indicate operational disruptions or inefficiencies in cargo transport.

1. **Freight Ton Kilometers by Month**

* Plot Used: Line Plot with Highlighted Points for Peak/Low Months
* Key Observations:
  + Freight TON KM Performed tracks the freight movement over months.
  + The plot highlights the months with the highest and lowest freight performance.
* Interpretation:
  + The highest freight ton kilometers indicate periods of maximum cargo transportation demand.
  + The lowest months suggest either low demand or underperformance, which could be further analyzed to uncover contributing factors.

1. **Flight Departures Over Time**

* Plot Used: Line Plot (Flight Departures over time)
* Key Observations:
  + No. Departure (AF) tracks the number of flights departing.
  + The plot shows how this number changes over time.
* Interpretation:
  + Increasing flight departures might indicate growing demand or expanded operations.
  + Decreasing departures could reflect reduced operations, seasonal variations, or fleet limitations.

1. **Passenger Load Factor vs Available Seat Kilometers**

* Plot Used: Regression Plot (Passenger Load Factor vs Available Seat Kilometers)
* Key Observations:
  + Available Seat Kilometers (ASK) are compared against the Passenger Load Factor (PAX load %).
  + The regression line shows the relationship between the two variables.
* Interpretation:
  + A positive slope indicates that as ASK increases, the load factor also tends to rise, suggesting that larger seat capacities are efficiently filled.
  + A lack of correlation could point to inefficient scheduling or capacity planning, where the load factor remains low despite high ASK.

1. **Weight Load Factor Over Time**

* Plot Used: Line Plot (Weight Load Factor Over Time)
* Key Observations:
  + Weight Load Factor (%) reflects the efficiency of weight carried compared to available capacity.
  + The plot shows how this efficiency fluctuates over time.
* Interpretation:
  + An increase in weight load factor over time indicates improved efficiency in filling available cargo space.
  + A decrease suggests that available cargo capacity is underutilized, which may require operational adjustments.

1. **Seasonal Patterns in Cargo**

* Plot Used: Boxplot (Seasonal Pattern in Cargo Transportation)
* Key Observations:
  + Freight TON KM Performed is visualized for each month using a boxplot to observe seasonal patterns.
  + The plot shows the distribution and range of cargo performance per month.
* Interpretation:
  + The boxplot helps identify months with consistently high or low cargo performance.
  + Seasonal peaks and troughs may suggest external factors, such as holidays or economic trends, influencing cargo volumes.

1. **Peak Passenger Traffic Month**

* Plot Used: Printout of Month with Peak Traffic
* Key Observations:
  + The month with the highest number of passengers carried is identified.
  + The printout highlights the month with peak traffic.
* Interpretation:
  + The month with peak passenger traffic shows the period of highest demand, often tied to seasonal or economic factors.
  + Airlines can use this information for resource allocation during peak periods.

1. **Efficiency of Available Tonne-Kilometers**

* Plot Used: Line Plot (Efficiency of Available Tonne-Kilometers)
* Key Observations:
  + Efficiency is calculated by comparing the available tonne-kilometers with the actual performed tonne-kilometers.
  + The plot shows efficiency fluctuations over time.
* Interpretation:
  + An efficiency close to 100% indicates optimal use of cargo capacity.
  + A significant drop suggests inefficient use of available capacity or potential underperformance in freight operations.

1. **Compare Weight Load Factor with Other Factors**

* Plot Used: Line Plot (Weight Load Factor with Other Factors)
* Key Observations:
  + This plot compares the Weight Load Factor with other operational factors (like passenger demand, freight performance) over time.
* Interpretation:
  + The comparison reveals how the weight load factor behaves relative to other operational factors.
  + Analyzing this relationship helps identify potential inefficiencies or areas for improvement in operations.

1. **Compare Cargo vs Passengers Trends**

* Plot Used: Line Plot (Total Cargo vs Passengers over time)
* Key Observations:
  + Total Cargo (CC) and Passengers (No Carried (P)) are compared over time.
  + The plot visualizes trends for both metrics.
* Interpretation:
  + A strong positive correlation indicates that the airline may be balancing passenger and cargo demand effectively.
  + Divergence between the two trends may suggest a shift in focus from one area to another (e.g., more cargo focus during slower passenger periods).

1. **Compare Flight Delays with Load Factor**

* Plot Used: Line Plot (Flight Delays vs Load Factor)
* Key Observations:
  + Flight delays (No. Departure (AF)) are compared with load factor (PAX load %) over time.
  + The plot shows how delays correlate with load factor.
* Interpretation:
  + If delays increase during high load factors, it may suggest operational inefficiencies.
  + A negative or weak correlation could indicate that delays are driven by factors unrelated to the load factor, such as weather or external delays.