REPORT

Airlines Data Analysis Project

Business Problem :

Incorporating a diverse array of aircraft, ranging from small business jets to medium-sized aircraft, our company has consistently delivered premium air transportation services to our clientele over numerous years. Our paramount objective is to ensure a secure, comfortable, and streamlined travel experience for our passengers.

Presently, we find ourselves confronted with a series of challenges emanating from various quarters. These encompass stringent environmental regulations, escalating flight levies, elevated interest rates, surging fuel costs, and a constrained labor market, which has precipitated augmented labor expenses. Consequently, the organization's profitability stands subject to mounting pressures, necessitating a comprehensive approach for resolution.

In response to this imperative, the organization is proactively pursuing a comprehensive analysis of its database. The primary objective of this endeavor is to uncover strategic pathways that will effectively augment the occupancy rate. By doing so, the company aims to catalyze a notable upswing in the average profit realized per seat, thereby addressing the core challenge and fortifying its financial standing.

Main Challenges :

- **Stricter Environmental Regulations –**

  The aviation industry is grappling with an escalating need to curtail its carbon emissions, driving the emergence of stricter environmental regulations. Consequently, the sector is confronting a landscape characterized by heightened legal standards. These regulations not only impose elevated operational expenditures but also impose limitations on growth prospects.

- **Escalated flight taxes –**

  In a global drive to address environmental concerns and bolster revenue streams, governments worldwide are imposing more substantial taxation on aircraft operations. This has a dual effect of elevating the cost of air travel while concurrently dampening demand due to increased financial constraints.

- **Labor Market –**

  The constricting labor market has led to a surge in labor expenses, primarily attributed to a scarcity of skilled personnel within the aviation sector. This scarcity not only amplifies the costs associated with labor but also contributes to elevated turnover rates within the industry.

Objectives :

- **Increase in Occupancy Rate –**

  Elevating the occupancy rate stands as a pivotal strategy, offering the potential to amplify the average profit earned per seat. This measure assumes paramount importance in light of the prevailing challenges, as it effectively serves to counterbalance their impact on the operational landscape.

- **Improve in the Pricing Strategy –**

  It is imperative to formulate a dynamic pricing strategy that not only remains attuned to evolving market dynamics but also aligns seamlessly with shifting customer inclinations. This strategic refinement will serve to not only allure new patrons but also cultivate lasting customer relationships, ultimately driving customer retention.

- **Enhance Customer Experience –**

  Our paramount priority lies in cultivating a seamlessly convenient journey for our esteemed clientele, spanning the entire spectrum from booking to arrival. In an industry marked by intense rivalry, this strategic emphasis on distinction through exceptional service is not only imperative but also instrumental in fostering unwavering customer loyalty.

The ultimate objective of this project is to pinpoint prospects for heightening the occupancy rate on flights exhibiting lower performance metrics. This strategic pursuit holds the potential to culminate in a substantial elevation of the airline's overall profitability.

Analysis & Findings :

1. Fundamental data analysis offers valuable insights into various aspects. These encompass the assessment of aircraft equipped with more than 100 seats, discerning trends in ticket bookings and corresponding revenue generation across time, and the calculation of average fares concerning distinct fare conditions for individual aircraft.These revelations serve as the bedrock for formulating strategies aimed at heightening occupancy rates and fine-tuning pricing strategies tailored to each specific aircraft type.

   In this context, reference to "Table 1" provides a visual representation of aircraft boasting more than 100 seats, accompanied by the precise count of available seating capacity.

	aircraft_code	num_seats
0	773	402
1	763	222
2	321	170
3	320	140
4	733	130
5	319	116

TABLE – 1

2. A Line Chart is visualized to gain a more profound insight into the trajectory of ticket bookings and the corresponding revenue. Upon a thorough analysis of this graphical representation, intriguing patterns come to light.
The period spanning from June 22nd to July 7th showcases a gradual escalation in ticket reservations. Subsequently, from July 8th to August, a stable plateau is discernible. Within this period, a prominent zenith in ticket bookings materializes, marked by a remarkable single-day peak.

```
Tickets = pd.read_sql_query(f"""SELECT *
                           FROM tickets
                           INNER JOIN bookings
                           ON tickets.book_ref=bookings.book_ref;""", connection)
```

The above SQL query retrieves data from the "tickets" and "bookings" tables by performing an inner join on the "book_ref" column. It fetches all columns from both tables for the joined rows and stores the result in the 'Tickets' DataFrame.

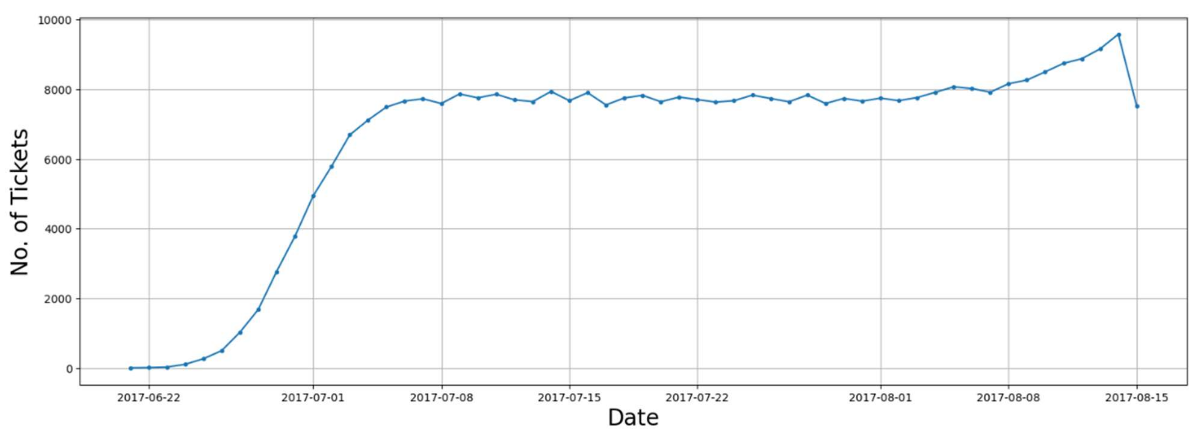

FIGURE – 1

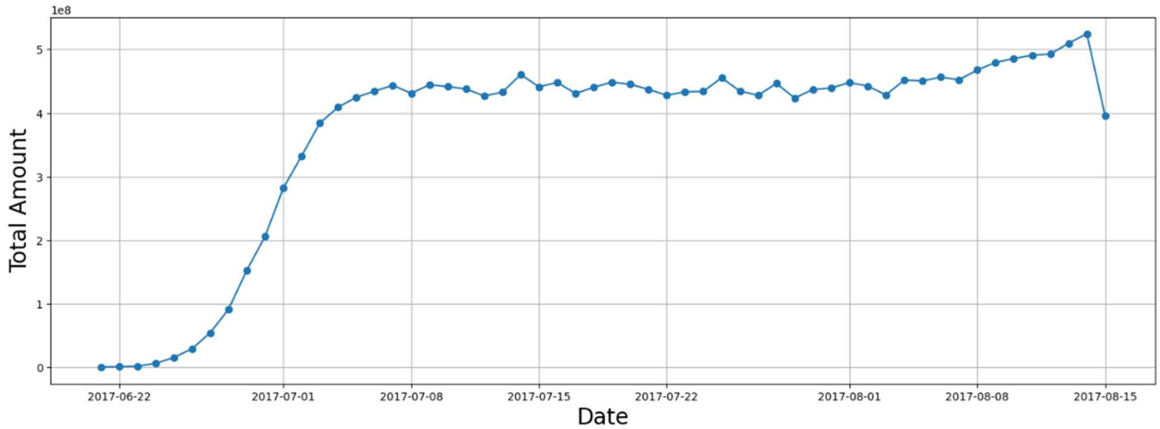

FIGURE – 2

3. Upon the successful completion of the calculations regarding average costs associated with distinct fare categories for each aircraft, a Bar Graph is plotted for a visual comparative representation of the data. Illustrated in Figure 3, the graph delineates information pertaining to three fare types: Business, Economy, and Comfort.
   A noteworthy point is that the exclusive availability of the comfort class pertains solely to the 773 aircraft. Conversely, the CN1 and CR2 planes exclusively accommodate economy class passengers.

   When comparing prices on different aircraft, a clear pattern emerges: business class fares consistently exceed those of economy class. This trend remains consistent across all aircraft, regardless of specific fare conditions.

```
df = pd.read_sql_query(f"""SELECT fare_conditions, aircraft_code, AVG(amount) as avg_amount
                         FROM ticket_flights
                         JOIN flights
                         ON ticket_flights.flight_id=flights.flight_id
                         GROUP BY aircraft_code, fare_conditions""", connection)
```

The above SQL query fetches data from the 'ticket_flights' and 'flights' tables, calculates the average amount for different fare conditions on different aircraft, and stores the results in a pandas DataFrame called 'df'.

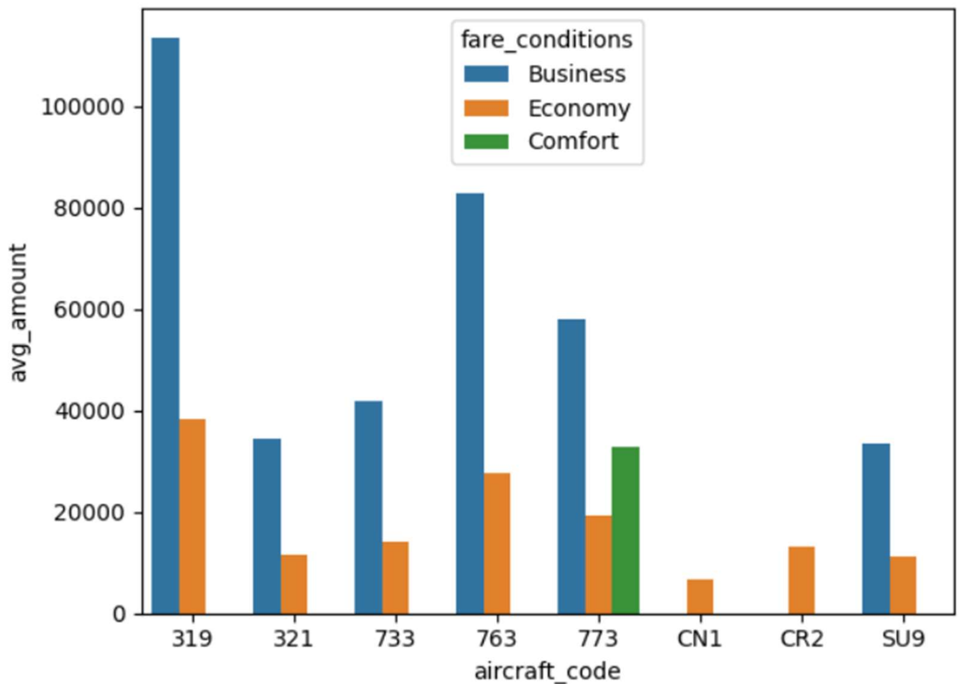

FIGURE – 3

4. **Analysis of the Occupancy Rate** :

   Airlines must conduct thorough revenue analysis to optimize profitability. Key metrics like annual overall income and average revenue per ticket per aircraft hold significance. Such insights guide airlines in identifying high-income aircraft types and routes, aiding operational adjustments. This study also unveils avenues for price optimization and resource allocation to lucrative routes.

```
pd.read_sql_query(f"""SELECT aircraft_code, total_revenue, ticket_count, total_revenue/ticket_count as avg_revenue_per_ticket
                      FROM
                        ( SELECT aircraft_code, COUNT(*) as ticket_count, SUM(amount) as total_revenue
                          FROM ticket_flights
                          JOIN flights
                          ON ticket_flights.flight_id=flights.flight_id
                          GROUP BY aircraft_code )""", connection)
```

The above SQL query retrieves data from the database related to aircraft revenue and ticket information.

It calculates the average revenue per ticket for each aircraft using a subquery and presents the results as a DataFrame using the pandas library.

In the table below (Table 2), one can observe total revenue, tickets, and average revenue per ticket across aircraft.

Remarkably, aircraft SU9 leads in total revenue, potentially due to its competitively priced business and economy classes.

Conversely, CN1 records lower total revenue, possibly stemming from its sole economy class offering, reflecting potential service limitations.

	aircraft_code	total_revenue	ticket_count	avg_revenue_per_ticket
0	319	2706163100	52853	51201
1	321	1638164100	107129	15291
2	733	1426552100	86102	16568
3	763	4371277100	124774	35033
4	773	3431205500	144376	23765
5	CN1	96373800	14672	6568
6	CR2	1982760500	150122	13207
7	SU9	5114484700	365698	13985

TABLE – 2

The average occupancy per aircraft stands as a pivotal metric of significance. For airlines, this metric serves as a gauge to evaluate seat utilization efficiency, and subsequently, to unearth avenues for augmenting occupancy rates. Elevated occupancy rates bear the potential to elevate both revenue and profitability, while concurrently mitigating operational overhead attributed to vacant seats.

A synergy of factors, including pricing strategies, airline schedules, and passenger satisfaction, can sway these occupancy rates.

The below SQL query fetches information from two subqueries ('a' and 'b') that calculate the number of booked seats for each flight on each aircraft and the total number of seats on each aircraft, respectively.

It then joins the results based on the aircraft code and calculates the occupancy rate by dividing the average booked seats by the total number of seats. The final result is stored in the occupancy_rate DataFrame.

This code helps to analyze and understand how efficiently the seats are being utilized on different aircraft.

```
occupancy_rate = pd.read_sql_query(f"""SELECT a.aircraft_code, AVG(a.seats_count) as booked_seats, b.num_seats,
                                          AVG(a.seats_count)/b.num_seats as occupancy_rate
                                   FROM (
                                          SELECT aircraft_code, flights.flight_id, COUNT(*) as seats_count
                                          FROM boarding_passes
                                          INNER JOIN flights
                                          ON boarding_passes.flight_id=flights.flight_id
                                          GROUP BY aircraft_code, flights.flight_id
                                        ) as a

                                   INNER JOIN

                                        (
                                          SELECT aircraft_code, COUNT(*) as num_seats
                                          FROM seats
                                          GROUP BY aircraft_code
                                        ) as b

                                   ON a.aircraft_code = b.aircraft_code
                                   GROUP BY a.aircraft_code""", connection)
occupancy_rate
```

Table 3 below shows how many seats on average are booked compared to the total available seats for each aircraft.

This helps us calculate the occupancy rate, which is the proportion of booked seats to total seats.

A higher occupancy rate indicates that the aircraft's seats are well-used, with fewer empty seats, a sign of strong bookings.

	aircraft_code	booked_seats	num_seats	occupancy_rate
0	319	53.583181	116	0.461924
1	321	88.809231	170	0.522407
2	733	80.255462	130	0.617350
3	763	113.937294	222	0.513231
4	773	264.925806	402	0.659019
5	CN1	6.004431	12	0.500369
6	CR2	21.482847	50	0.429657
7	SU9	56.812113	97	0.585692

TABLE – 3

Airlines have the opportunity to evaluate the potential enhancement of their annual revenue by increasing the occupancy rate of all aircraft by 10%. This analysis delves into the prospective advantages of elevating occupancy rates. The insights drawn from this research serve as a crucial aid for airlines in assessing the financial implications of such a strategy.

Through optimizing pricing strategies and refining operational facets, airlines stand to simultaneously augment occupancy rates, generate higher revenue, and offer an elevated level of value and service to their customers.

occupancy_rate['Increased occupancy rate'] = occupancy_rate['occupancy_rate'] + occupancy_rate['occupancy_rate']*0.1

The above SQL query calculates the increased occupancy rate by adding 10% of the original occupancy rates to the original occupancy rates.

The resulting values are stored in the newly created "Increased occupancy rate" column within the 'occupancy_rate' DataFrame.

This column now contains the projected occupancy rates after increasing the rates by 10%. Refer to Table 4 below.

	aircraft_code	booked_seats	num_seats	occupancy_rate	Increased occupancy rate
0	319	53.583181	116	0.461924	0.508116
1	321	88.809231	170	0.522407	0.574648
2	733	80.255462	130	0.617350	0.679085
3	763	113.937294	222	0.513231	0.564554
4	773	264.925806	402	0.659019	0.724921
5	CN1	6.004431	12	0.500369	0.550406
6	CR2	21.482847	50	0.429657	0.472623
7	SU9	56.812113	97	0.585692	0.644261

TABLE – 4

```
occupancy_rate['Increased Total Annual Turnover'] =
(total_revenue['total_revenue']/occupancy_rate['occupancy_rate'])*occupancy_rate['I
ncreased occupancy rate']
```

The above SQL query calculates the "Increased Total Annual Turnover" by taking the average revenue per passenger (current total revenue divided by the current occupancy rate) and then multiplying it by the "Increased occupancy rate" to project the potential total revenue that would be generated if the occupancy rate were increased by the specified percentage.

The results are stored in the newly added column in the 'occupancy_rate' DataFrame.

The accompanying Table 5 illustrates the incremental increase in total revenue resulting from a 10% upsurge in occupancy rates.

This depiction underscores the gradual growth in revenue and highlights the necessity for airlines to prioritize their pricing strategies in their pursuit of sustainable success.

	aircraft_code	booked_seats	num_seats	occupancy_rate	Increased occupancy rate	Increased Total Annual Turnover
0	319	53.58318098720292	116	0.46192397402761143	0.5081163714303726	2976779410.0
1	321	88.80923076923077	170	0.5224072398190045	0.574647963800905	1801980510.0
2	733	80.25546218487395	130	0.617349709114415	0.6790846800258565	1569207310.0000002
3	763	113.93729372937294	222	0.5132310528350132	0.5645541581185146	4808404810.0
4	773	264.9258064516129	402	0.659019419033863	0.7249213609372492	3774326050.0
5	CN1	6.004431314623338	12	0.5003692762186115	0.5504062038404727	106011180.00000001
6	CR2	21.48284690220174	50	0.42965693804403476	0.4726226318484382	2181036550.0
7	SU9	56.81211267605634	97	0.5856918832583128	0.644261071584144	5625933169.999999

TABLE – 5

Conclusion :

- Key Revenue Metrics:

  Analyzing revenue data, including total annual revenue, average revenue per ticket, and average occupancy per aircraft, holds paramount importance for airlines aiming to maximize profitability.

- Identifying Improvement Areas:

  These revenue indicators serve as crucial guides for airlines to pinpoint areas that warrant improvement. By assessing these metrics, airlines can refine pricing strategies and adjust route plans to optimize their revenue streams.

- Significance of Higher Occupancy:

  A higher occupancy rate stands as a pivotal factor in enhancing profitability. It enables airlines to maximize revenue while mitigating costs tied to unoccupied seats, thus improving overall financial performance.

- Strategic Pricing:

  Airlines must meticulously review their pricing structures for each aircraft. Pricing that is either too low or excessively high can impact ticket sales negatively. Striking a balance between price and aircraft features is vital.

- Balancing Profit and Service:

  While boosting occupancy rates is critical, airlines must ensure it doesn't compromise passenger satisfaction or safety. Striking a harmony between profitability and maintaining high-quality service, as well as adhering to safety regulations, is essential.

- Long-Term Success:

  Airlines that adopt a data-driven approach to revenue analysis and optimization position themselves for enduring success in an intensely competitive industry. Such a strategy allows them to make informed decisions, tailor offerings, and maximize profitability without compromising on customer satisfaction or safety.