✈️ Flight Fare Analysis Using SQL

Project Objective:

The objective of this project is to analyse flight fare data by applying SQL techniques, perform data cleaning and exploratory data analysis (EDA) using SQL, and then build a predictive model using machine learning to estimate flight ticket prices based on various travel-related features.

About data:

- Airline: The name of the airline (e.g., IndiGo, Air India, SpiceJet). Different airlines have different pricing strategies.
- Date_of_Journey: The date on which the passenger is scheduled to fly. Can be split into day, month, and year for analysis.
- Source: The city from which the flight departs (e.g., Delhi, Mumbai).
- Destination: The city where the flight is landing.
- Route: The route taken by the flight, sometimes with stops (e.g., DEL → BOM → BLR). This can give information about layovers.
- Arrival_Time: Time at which the flight arrives at the destination. Similar to departure time, can be transformed.
- Dep_Time: Departure time of the flight. Can be broken into hour and minute or part of day (Morning, Afternoon, Evening).
- Duration: Total duration of the flight from source to destination.
- Total_Stops: Number of stops (e.g., non-stop, 1 stop, 2 stops). Usually, more stops mean cheaper flights.
- Additional_Info: Any extra info (e.g., 'No info', 'In-flight meal not included', etc.). Might affect pricing.
- Price: The target variable. This is what you're trying to predict using the other features.

Data Wrangling:

- **First five rows:**

```
SELECT * FROM flight
LIMIT 5;
```

MyUnknownColumn	Airline	Date_of_Journey	Source	Destination	Route	Dep_Time	Arrival_Time	Duration	Total_Stops	Additional_Info	Price
0	IndiGo	24/03/2019	Banglore	New Delhi	BLR â†' DEL	22:20	01:10 22 Mar	2h 50m	non-stop	No info	3897
1	Air India	1/05/2019	Kolkata	Banglore	CCU â†' IXR â†' BBI â†' BLR	05:50	13:15	7h 25m	2 stops	No info	7662
2	Jet Airways	9/06/2019	Delhi	Cochin	DEL â†' LKO â†' BOM â†' COK	09:25	04:25 10 Jun	19h	2 stops	No info	13882
3	IndiGo	12/05/2019	Kolkata	Banglore	CCU â†' NAG â†' BLR	18:05	23:30	5h 25m	1 stop	No info	6218
4	IndiGo	01/03/2019	Banglore	New Delhi	BLR â†' NAG â†' DEL	16:50	21:35	4h 45m	1 stop	No info	13302

New Delhi

- **Last five rows:**

```
SELECT * FROM flight
ORDER BY RAND(Airline)
LIMIT 5;
```

MyUnknownColumn	Airline	Date_of_Journey	Source	Destination	Route	Dep_Time	Arrival_Time	Duration	Total_Stops	Additional_Info	Price
10315	Jet Airways	18/05/2019	Kolkata	Banglore	CCU â†' BOM â†' BLR	14:05	22:35	8h 30m	1 stop	In-flight meal not included	10844
10316	IndiGo	18/03/2019	Banglore	New Delhi	BLR â†' HYD â†' DEL	12:10	17:05	4h 55m	1 stop	No info	4410
10317	IndiGo	21/05/2019	Mumbai	Hyderabad	BOM â†' HYD	02:35	04:05	1h 30m	non-stop	No info	2754
10318	Jet Airways	6/04/2019	Banglore	Delhi	BLR â†' DEL	19:50	22:50	3h	non-stop	No info	7229
10319	IndiGo	09/03/2019	Banglore	New Delhi	BLR â†' DEL	23:30	02:20 10 Mar	2h 50m	non-stop	No info	5694

- **Checking Null values in each column:**

```
SELECT 'Airline', COUNT(*) FROM flight WHERE Airline IS NULL
UNION ALL
SELECT 'Date_of_Journey', COUNT(*) FROM flight WHERE Date_of_Journey IS NULL
UNION ALL
SELECT 'Source', COUNT(*) FROM flight WHERE Source IS NULL
UNION ALL
SELECT 'Destination', COUNT(*) FROM flight WHERE Destination IS NULL
UNION ALL
SELECT 'Route', COUNT(*) FROM flight WHERE Route IS NULL
UNION ALL
SELECT 'Dep_Time', COUNT(*) FROM flight WHERE Dep_Time IS NULL
UNION ALL
SELECT 'Arrival_Time', COUNT(*) FROM flight WHERE Arrival_Time IS NULL
UNION ALL
SELECT 'Duration', COUNT(*) FROM flight WHERE Duration IS NULL
UNION ALL
SELECT 'Total_Stops', COUNT(*) FROM flight WHERE Total_Stops IS NULL
UNION ALL
SELECT 'Additional_Info', COUNT(*) FROM flight WHERE Additional_Info IS NULL
UNION ALL
SELECT 'Price', COUNT(*) FROM flight WHERE Price IS NULL;
```

Airline	missing_value
Airline	0
Date_of_Journey	0
Source	0
Destination	0
Route	0
Dep_Time	0
Arrival_Time	0
Duration	0
Total_Stops	0
Additional_Info	0
Price	0

- **Checking duplicate values:**

```
SELECT count(*) AS duplicate_rows
FROM flight
GROUP BY
```

```
MyUnknownColumn, Airline, Date_of_Journey, Source, Destination,
Route, Dep_Time, Arrival_Time, Duration, Total_Stops,
Additional_Info, Price
HAVING COUNT(*) > 1;
```

0 duplicate rows are there.

- **Data cleaning and feature transformation:**

- Convert `Date_of_Journey` from text to date format:

```
ALTER TABLE project.flight ADD COLUMN DATE_OF_JOURNEY_2 DATE;
UPDATE project.flight
SET DATE_OF_JOURNEY_2 = STR_TO_DATE(Date_of_Journey,'%d/%m/%Y');
```

- **Create `DEPARTURE_TIME` column by adding `DATE_OF_JOURNEY_2` and `Dep_Time`:**

```
ALTER TABLE project.test ADD COLUMN DEPARTURE_TIME DATETIME;

UPDATE project.test
SET DEPARTURE_TIME = STR_TO_DATE(CONCAT(DATE_OF_JOURNEY_2,' ',Dep_Time),'%Y-%m-%d
%H:%i');
```

- Create `DURATIONMIN` column by Converting `Duration` which was in Hours , in to minutes:

```
ALTER TABLE project.flight ADD COLUMN DURATIONMIN INT;
UPDATE project.flight
SET DURATIONMIN = CAST(REPLACE((CASE WHEN SUBSTRING_INDEX(Duration,' ',1) =
SUBSTRING_INDEX(Duration,' ',-1) THEN 0
ELSE SUBSTRING_INDEX(Duration,' ',1) END), 'h', '') AS UNSIGNED)*60 +
CAST(REPLACE((CASE WHEN SUBSTRING_INDEX(Duration,' ',-1) = SUBSTRING_INDEX(Duration,' ',1)
THEN 0
ELSE SUBSTRING_INDEX(Duration,' ',-1) END), 'm', '') AS UNSIGNED);
```

- Create `ARRIVAL_TIME` column by combining `DEPARTURE_TIME` and `DURATIONMIN `:

```
ALTER TABLE project.flight ADD COLUMN ARRIVAL_TIME DATETIME;
UPDATE project.flight
SET ARRIVAL_TIME = DATE_ADD(DEPARTURE_TIME, INTERVAL DURATIONMIN MINUTE);
```

- - **Drop unwanted columns:**

```
ALTER TABLE flight_dataset
```

```
DROP COLUMN Route,
DROP COLUMN Additional_Info,
DROP COLUMN MyUnknownColumn,
DROP COLUMN Date_of_Journey;
```

Airline	Source	Destination	Dep_Time	Duration	Total_Stops	Price	DATE_OF_JOURNEY_2	DEPARTURE_TIME	DURATIONMIN	ARRIVAL_TIME
IndiGo	Banglore	New Delhi	22:20	2h 50m	non-stop	3897	2019-03-24	2019-03-24 22:20:00	170	2019-03-25 01:10:00
Air India	Kolkata	Banglore	05:50	7h 25m	2 stops	7662	2019-05-01	2019-05-01 05:50:00	445	2019-05-01 13:15:00
Jet Airways	Delhi	Cochin	09:25	19h	2 stops	13882	2019-06-09	2019-06-09 09:25:00	0	2019-06-09 09:25:00
IndiGo	Kolkata	Banglore	18:05	5h 25m	1 stop	6218	2019-05-12	2019-05-12 18:05:00	325	2019-05-12 23:30:00
IndiGo	Banglore	New Delhi	16:50	4h 45m	1 stop	13302	2019-03-01	2019-03-01 16:50:00	285	2019-03-01 21:35:00
SpiceJet	Kolkata	Banglore	09:00	2h 25m	non-stop	3873	2019-06-24	2019-06-24 09:00:00	145	2019-06-24 11:25:00
Jet Airways	Banglore	New Delhi	18:55	15h 30m	1 stop	11087	2019-03-12	2019-03-12 18:55:00	930	2019-03-13 10:25:00
Jet Airways	Banglore	New Delhi	08:00	21h 5m	1 stop	22270	2019-03-01	2019-03-01 08:00:00	1265	2019-03-02 05:05:00
Jet Airways	Banglore	New Delhi	08:55	25h 30m	1 stop	11087	2019-03-12	2019-03-12 08:55:00	1530	2019-03-13 10:25:00
Multiple carriers	Delhi	Cochin	11:25	7h 50m	1 stop	8625	2019-05-27	2019-05-27 11:25:00	470	2019-05-27 19:15:00
Air India	Delhi	Cochin	09:45	13h 15m	1 stop	8907	2019-06-01	2019-06-01 09:45:00	795	2019-06-01 23:00:00
IndiGo	Kolkata	Banglore	20:20	2h 35m	non-stop	4174	2019-04-18	2019-04-18 20:20:00	155	2019-04-18 22:55:00
Air India	Chennai	Kolkata	11:40	2h 15m	non-stop	4667	2019-06-24	2019-06-24 11:40:00	135	2019-06-24 13:55:00
Jet Airways	Kolkata	Banglore	21:10	12h 10m	1 stop	9663	2019-05-09	2019-05-09 21:10:00	730	2019-05-10 09:20:00
IndiGo	Kolkata	Banglore	17:15	2h 35m	non-stop	4804	2019-04-24	2019-04-24 17:15:00	155	2019-04-24 19:50:00
Air India	Delhi	Cochin	16:40	26h 35m	2 stops	14011	2019-03-03	2019-03-03 16:40:00	1595	2019-03-04 19:15:00
SpiceJet	Delhi	Cochin	08:45	4h 30m	1 stop	5830	2019-04-15	2019-04-15 08:45:00	270	2019-04-15 13:15:00
Jet Airways	Delhi	Cochin	14:00	22h 35m	1 stop	10262	2019-06-12	2019-06-12 14:00:00	1355	2019-06-13 12:35:00
Air India	Delhi	Cochin	20:15	23h	2 stops	13381	2019-06-12	2019-06-12 20:15:00	0	2019-06-12 20:15:00
Jet Airways	Delhi	Cochin	16:00	20h 35m	1 stop	12898	2019-05-27	2019-05-27 16:00:00	1235	2019-05-28 12:35:00
GoAir	Delhi	Cochin	14:10	5h 10m	1 stop	19495	2019-03-06	2019-03-06 14:10:00	310	2019-03-06 19:20:00

Exploratory Data Analysis:

Q1 find the month with most number of flights:

```
SELECT MONTHNAME(DATE_OF_JOURNEY_2) MONTH_NAME ,COUNT(*) NO_OF_FLIGHTS

FROM flight_dataset

GROUP BY MONTHNAME(DATE_OF_JOURNEY_2)

ORDER BY NO_OF_FLIGHTS DESC;
```

MONTH_NAME	NO_OF_FLIGHTS
May	3466
June	3414
March	2724
April	1079

Insights: May , june have most number of Flights

Reason :

- **SummerVacation :** In India, schools and colleges usually have summer holidays during May and June, leading to a surge in family travel.
- **TourismSeason :**Many people plan holidays during these months to hill stations or tourist destinations to escape the heat.
- **Favorable whether for travel :**It's typically before the monsoon starts, so flight operations are less likely to be disrupted by rain.

Q2: Which weekday has the most expensive flights?

```
SELECT DAYNAME(DATE_OF_JOURNEY_2) MONTH_NAME ,ROUND(AVG(Price)) PRICE

FROM flight_dataset

GROUP BY DAYNAME(DATE_OF_JOURNEY_2)

ORDER BY PRICE DESC;
```

MONTH_NAME	PRICE
Friday	9718
Sunday	9526
Wednesday	9278
Tuesday	9026
Saturday	8973
Thursday	8931
Monday	8500

Insights: Fridays have the most expensive flights, followed by Sundays, Wednesdays, Tuesdays, Saturdays, Thursdays, and Mondays

Reason :

- Many people travel on Friday evenings for short weekend trips, causing a spike in demand. High demand ~ High price
- Sunday : People returning from weekend trips → high evening demand.
- Wednesday:Popular day for business travel due to midweek meetings/conferences.
- Monday : Many travellers avoid early Monday flights, bcz of their busy schedule , low demand~ low price

Q3: Find the list of all flights that depart between 10 AM and 2 PM from Bangalore to Delhi.

```
SELECT * FROM flight_dataset

WHERE Source='Banglore' AND Destination='Delhi' AND Dep_Time>'10:00' AND
Dep_Time<'14:00';
```

Airline	Source	Destination	Dep_Time	Duration	Total_Stops	Price	DATE_OF_JOURNEY_2	DEPARTURE_TIME	DURATIONMIN	ARRIVAL_TIME
Air Asia	Banglore	Delhi	11:10	2h 45m	non-stop	3383	2019-05-06	2019-05-06 11:10:00	165	2019-05-06 13:55:00
GoAir	Banglore	Delhi	11:40	2h 50m	non-stop	3898	2019-04-03	2019-04-03 11:40:00	170	2019-04-03 14:30:00
GoAir	Banglore	Delhi	11:40	2h 50m	non-stop	3898	2019-06-27	2019-06-27 11:40:00	170	2019-06-27 14:30:00
Air Asia	Banglore	Delhi	11:10	2h 45m	non-stop	3383	2019-06-12	2019-06-12 11:10:00	165	2019-06-12 13:55:00
Jet Airways	Banglore	Delhi	11:10	2h 55m	non-stop	4030	2019-05-24	2019-05-24 11:10:00	175	2019-05-24 14:05:00
Vistara	Banglore	Delhi	11:30	2h 50m	non-stop	4668	2019-06-27	2019-06-27 11:30:00	170	2019-06-27 14:20:00
IndiGo	Banglore	Delhi	10:10	2h 50m	non-stop	3943	2019-05-18	2019-05-18 10:10:00	170	2019-05-18 13:00:00
GoAir	Banglore	Delhi	11:40	2h 50m	non-stop	4937	2019-04-18	2019-04-18 11:40:00	170	2019-04-18 14:30:00
GoAir	Banglore	Delhi	11:40	2h 50m	non-stop	3398	2019-05-24	2019-05-24 11:40:00	170	2019-05-24 14:30:00
Air India	Banglore	Delhi	13:20	2h 50m	non-stop	6961	2019-06-09	2019-06-09 13:20:00	170	2019-06-09 16:10:00
Jet Airways	Banglore	Delhi	11:10	2h 55m	non-stop	7229	2019-04-21	2019-04-21 11:10:00	175	2019-04-21 14:05:00
Air India	Banglore	Delhi	13:20	2h 50m	non-stop	5228	2019-04-27	2019-04-27 13:20:00	170	2019-04-27 16:10:00
Air Asia	Banglore	Delhi	11:10	2h 45m	non-stop	3383	2019-04-24	2019-04-24 11:10:00	165	2019-04-24 13:55:00
GoAir	Banglore	Delhi	11:40	2h 50m	non-stop	3898	2019-05-03	2019-05-03 11:40:00	170	2019-05-03 14:30:00
Air Asia	Banglore	Delhi	11:10	2h 45m	non-stop	4483	2019-04-18	2019-04-18 11:10:00	165	2019-04-18 13:55:00
IndiGo	Banglore	Delhi	13:00	2h 50m	non-stop	3943	2019-06-24	2019-06-24 13:00:00	170	2019-06-24 15:50:00
IndiGo	Banglore	Delhi	10:10	2h 50m	non-stop	4423	2019-04-09	2019-04-09 10:10:00	170	2019-04-09 13:00:00
Air India	Banglore	Delhi	13:20	2h 50m	non-stop	6961	2019-06-15	2019-06-15 13:20:00	170	2019-06-15 16:10:00

Q4: Find the number of IndiGo flights for each month.

```
SELECT MONTHNAME(DATE_OF_JOURNEY_2) MONTH_NAME ,COUNT(*) NO_OF_FLIGHTS

FROM flight_dataset

WHERE Airline='IndiGo'

GROUP BY MONTHNAME(DATE_OF_JOURNEY_2)

ORDER BY NO_OF_FLIGHTS DESC;
```

MONTH_NAME	NO_OF_FLIGHTS
June	611
March	567
May	562
April	313

Insights :June has the highest number of IndiGo flight bookings, followed by March, May, and April.

Reason : Because of Summer Holidays and Festivals like holi.

Q5: Find the Number of all flights that depart between 10 AM and 2 PM from Bangalore to Delhi.

```
SELECT Airline,count(*) Number_of_Flights FROM flight_dataset
```

```
WHERE Source='Banglore' AND Destination='Delhi' AND Dep_Time>'10:00' AND
Dep_Time<'14:00'

Group by Airline
```

Airline	Number_of_Flights
Air Asia	23
GoAir	18
Jet Airways	46
Vistara	21
IndiGo	52
Air India	24
Vistara Premium economy	1

Insights : IndiGo has the highest number of flight bookings on the Bangalore to Delhi route.

Reason : IndiGo offers the lowest ticket fares compared to other airlines on this route, making it the most preferred choice for travellers.

Q6: Which source-destination pair has the highest travel cost?

```
Select Concat(Source,"-",Destination) result , Round(Avg(Price)) Traveling_Cost From
flight_dataset

Group By Source,Destination

Order by Traveling_Cost Desc
```

result	Traveling_Cost
Banglore-New Delhi	11918
Delhi-Cochin	10539
Kolkata-Banglore	9158
Banglore-Delhi	5144
Mumbai-Hyderabad	5060
Chennai-Kolkata	4790

Insights : Bangalore to New delhi have highest travelling cost followed by delhi cochin and cheapest is chennai kolkata

Reason :The longer distance between Bangalore and Delhi, combined with the frequent travel of corporate and government employees, leads to higher demand, which in turn results in higher ticket prices.

Q7: Calculate the arrival time for all flights by adding the duration to the departure time.

```
SELECT ARRIVAL_TIME

FROM flight_dataset
```

ARRIVAL_TIME
2019-03-25 01:10:00
2019-05-01 13:15:00
2019-06-09 09:25:00
2019-05-12 23:30:00
2019-03-01 21:35:00
2019-06-24 11:25:00
2019-03-13 10:25:00
2019-03-02 05:05:00
2019-03-13 10:25:00
2019-05-27 19:15:00
2019-06-01 23:00:00
2019-04-18 22:55:00

Q8: Find the airline and their corresponding price, and sort the result in decreasing order of price.

```
Select Airline ,Round(Avg(Price)) Avg_Price From flight_dataset

Group by Airline

Order by Avg_Price desc
```

Airline	Avg_Price
Jet Airways Business	58359
Jet Airways	11644
Multiple carriers Premium economy	11419
Multiple carriers	10903
Air India	9611
Vistara Premium economy	8962
Vistara	7796
GoAir	5861
IndiGo	5674
Air Asia	5590
SpiceJet	4338
Trujet	4140

Jet Airways (Business Class) has the highest ticket cost compared to other airlines.They likely offer better facilities, premium services, and a more comfortable travel experience, justifying the higher price.

Q9: Find the number of flights that travel on multiple dates (i.e., flights where the departure date is not equal to the arrival date).

```
SELECT COUNT(*) NO_OF_FLIGHTS

FROM flight_dataset

WHERE DATE(DEPARTURE_TIME) != DATE(ARRIVAL_TIME);
```

NO_OF_FLIGHTS
3982

Q10: Calculate the average flight duration between all city pairs.

```
SELECT
Source,Destination,TIME_FORMAT(SEC_TO_TIME(ROUND(AVG(DURATIONMIN))*60),'%Hh:%im')
AVG_TIME

FROM flight_dataset

GROUP BY Source,Destination;
```

Source	Destination	AVG_TIME
Banglore	New Delhi	10h:14m
Kolkata	Banglore	11h:39m
Delhi	Cochin	11h:52m
Chennai	Kolkata	02h:20m
Banglore	Delhi	02h:19m
Mumbai	Hyderabad	02h:56m

Highest price is between bangalore to new delhi : because of longer distance between them.

Q11: Find all the non-stop flights (0 stops) that departed before midnight but arrived at their destination after midnight.

```
SELECT * FROM flight_dataset

WHERE TIME(DEPARTURE_TIME)>'23:59:59' AND TIME(ARRIVAL_TIME)< '23:59:59';
```

0 Flights rows are there.

Q12: Find the quarter-wise number of flights for each airline.

```
SELECT Airline,QUARTER(DATE_OF_JOURNEY_2) QUATER,COUNT(*) NO_OF_FLIGHTS FROM
flight_dataset

GROUP BY Airline,QUARTER(DATE_OF_JOURNEY_2)

ORDER BY NO_OF_FLIGHTS DESC;
```

Airline	QUATER	NO_OF_FLIGHTS
Jet Airways	2	2952
IndiGo	2	1486
Air India	2	1222
Jet Airways	1	897
Multiple carriers	2	883
SpiceJet	2	639
IndiGo	1	567
Air India	1	530
Vistara	2	363
Multiple carriers	1	313
Air Asia	2	259
SpiceJet	1	170

Q13: Find the longest flight between cities in India in terms of duration (time).

```
SELECT Source,Destination,TIME_FORMAT(SEC_TO_TIME(AVG(DURATIONMIN)*60),'%hh:%im')
AVG_DURATION FROM flight_dataset

GROUP BY Source,Destination

ORDER BY AVG_DURATION DESC;
```

Airline	QUATER	NO_OF_FLIGHTS
Jet Airways	2	2952
IndiGo	2	1486
Air India	2	1222
Jet Airways	1	897
Multiple carriers	2	883
SpiceJet	2	639
IndiGo	1	567
Air India	1	530
Vistara	2	363
Multiple carriers	1	313
Air Asia	2	259
SpiceJet	1	179
GoAir	2	154
Vistara	1	116
Air Asia	1	60
GoAir	1	40
Multiple carrier...	1	13
Jet Airways Bu...	1	6

Q14: Find the average flight duration for flights with 1 stop and for flights with more than 1 stop.

```
WITH TEMP AS (

SELECT *,(CASE WHEN Total_Stops='1 stop' THEN '1' ELSE '1+' END) NUMBER_OF_STOPS FROM
flight_dataset

WHERE Total_Stops!='non-stops'

)

SELECT
NUMBER_OF_STOPS,TIME_FORMAT(SEC_TO_TIME(ROUND(AVG(DURATIONMIN))*60),'%Hh:%im'
) AVG_TIME FROM TEMP

GROUP BY NUMBER_OF_STOPS

ORDER BY AVG_TIME DESC
```

NUMBER_OF_STOPS	AVG_TIME
1	11h:42m
1+	07h:17m

Q15: Find all city pairs where the average flight duration is greater than 3 hours.

```
SELECT Source,Destination FROM flight_dataset

GROUP BY Source,Destination

Having Avg(DURATIONMIN)>180;
```

Source	Destination
Banglore	New Delhi
Kolkata	Banglore
Delhi	Cochin

Q16: Does the number of stops affect the fare price?

```
Select Total_Stops , Round(avg(Price)) Cost From flight_dataset

Where Total_Stops in ('non-stop','1 stop','2 stops','3 stops','4 stops')

Group by Total_Stops

Order By Cost
```

Total_Stops	Cost
non-stop	5025
1 stop	10594
2 stops	12716
3 stops	13112
4 stops	17686

There is a positive correlation between number of stops and ticket price which is obvious.More stops often mean longer travel duration, additional airport charges, and higher operational costs, all of which contribute to higher fares.

Q17: Which part of the day (Morning, Afternoon, Evening, or Midnight) is the cheapest for flying?

```
With Temp as (

  Select (Case when Hour(Dep_Time)  > 6 and Hour(Dep_Time)<=12 Then 'Morning'

  When Hour(Dep_Time) >12 and Hour(Dep_Time)<=16 Then 'Afternoon'
```

```
WHen  Hour(Dep_Time)>16  and Hour(Dep_Time)<=19 Then 'Evening'
When Hour(Dep_Time)>19  and Hour(Dep_Time)<=24 Then 'Night'
ELse 'Midnight'
End) Day_time, Price From flight_dataset)
Select Day_time, Round(Avg(Price)) Cost From Temp
Group By Day_time
Order by Cost
```

Day_time	Cost
Midnight	8639
Night	8884
Evening	8991
Morning	9286
Afternoon	9407

Midnight is the cheapest time to travel by flight.During midnight hours, the demand is low, leading to lower ticket prices as part of airlines' dynamic pricing strategy.

Q18: Is there any correlation between price and flight duration?

```
WITH
X AS (
  SELECT AVG(Price) AS avg_price FROM flight_dataset
),
Y AS (
  SELECT AVG(DURATIONMIN) AS avg_duration FROM flight_dataset
)
```

```
SELECT
  SUM((Price - (SELECT avg_price FROM X)) * (DURATIONMIN - (SELECT avg_duration FROM Y))) /
  SQRT(
    SUM(POWER(Price - (SELECT avg_price FROM X), 2)) *
    SUM(POWER(DURATIONMIN - (SELECT avg_duration FROM Y), 2))
  ) AS correlation
FROM
  flight_dataset;
```

correlation
0.4369804621197976

There is a positive relationship between ticket price and flight duration .As the flight duration increases, the ticket price also tends to increase, indicating that longer flights generally cost more due to factors like fuel usage, crew hours, and operational expenses.

Q19: Create a weekday vs time-grid table showing the frequency of flights from Bangalore to Delhi — similar to a pivot table.

```
-- STEP 1 : FILTER THE DATA FROM BANGLORE TO DELHI
    SELECT * FROM flight_dataset
    WHERE Source='Banglore' AND Destination='Delhi';

-- STEP 2 : NOW GROUP BY ON DAYNAME
    SELECT DAYOFWEEK(DATE_OF_JOURNEY_2) DAY_NUM,DAYNAME(DATE_OF_JOURNEY_2)
    WEEKDAYS FROM flight_dataset
    WHERE Source='Banglore' AND Destination='Delhi'
    GROUP BY DAYOFWEEK(DATE_OF_JOURNEY_2),DAYNAME(DATE_OF_JOURNEY_2)
```

```
ORDER BY DAYOFWEEK(DATE_OF_JOURNEY_2) ASC;


-- STEP 3 CREATE TIME RANGE LIKE 0-6AM,6-12PM,12-18PM,18-00AM AND FOR EACH COMBO CAL AVG(PRICE)

    SELECT DAYOFWEEK(DATE_OF_JOURNEY_2) DAY_NUM,DAYNAME(DATE_OF_JOURNEY_2) WEEKDAYS,

    ROUND(AVG(CASE WHEN HOUR(DEPARTURE_TIME) BETWEEN 0 AND 5 THEN Price ELSE NULL END)) '12AM-6AM',

    ROUND(AVG(CASE WHEN HOUR(DEPARTURE_TIME) BETWEEN 6 AND 11 THEN Price ELSE NULL END )) '6AM-12PM', -- BETWEEN DONO NUMBER KO INCLUDE KARTA HAI


    ROUND(AVG(CASE WHEN HOUR(DEPARTURE_TIME) BETWEEN 12 AND 17 THEN Price ELSE NULL END)) '12PM-18PM',

    ROUND(AVG(CASE WHEN HOUR(DEPARTURE_TIME) BETWEEN 18 AND 23 THEN Price ELSE NULL END)) '18PM-12AM'

     FROM flight_dataset

    WHERE Source='Banglore' AND Destination='Delhi'

    GROUP BY DAYOFWEEK(DATE_OF_JOURNEY_2),DAYNAME(DATE_OF_JOURNEY_2)

    ORDER BY DAYOFWEEK(DATE_OF_JOURNEY_2) ASC;
```

DAY_NUM	WEEKDAYS	12AM-6AM	6AM-12PM	12PM-18PM	18PM-12AM
1	Sunday	15	53	25	38
2	Monday	28	98	48	78
3	Tuesday	18	50	16	41
4	Wednesday	24	93	36	63
5	Thursday	20	60	24	50
6	Friday	20	69	24	55
7	Saturday	20	94	39	66