CS-586 Database Management Systems
Graduate Student Project Deliverable – 4

Nida Mariam Sheikh Aslam (nidama)
Snehil Shrivastava (snehils)

Subject Area: Flight Operations Analysis at JFK Airport

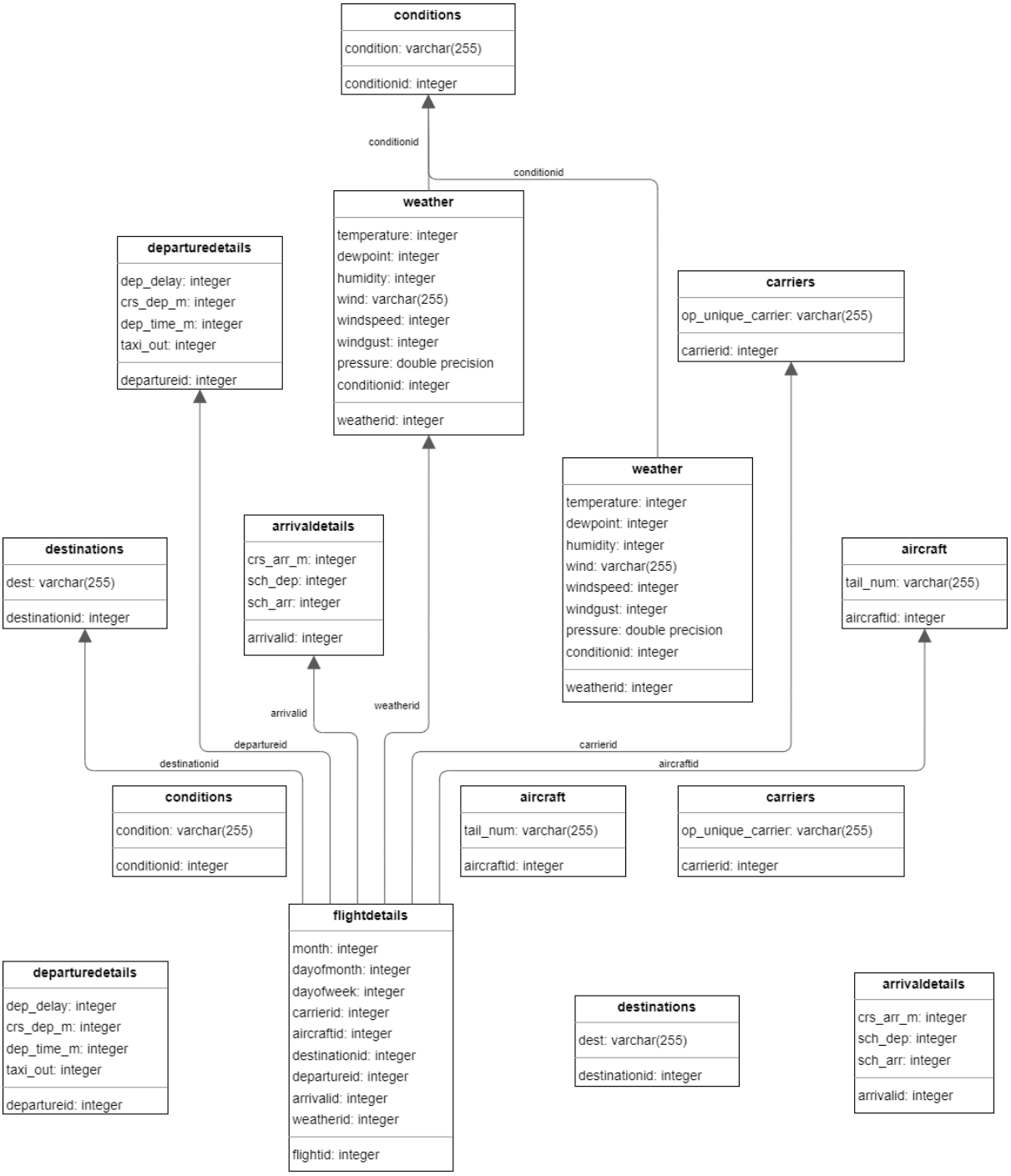
The subject area focuses on the daily operations of flights taking off from John F. Kennedy International Airport (JFK), one of the busiest airports in the United States and a major international air travel hub. JFK Airport serves as a critical point in the global air transportation network, handling flights to and from hundreds of destinations worldwide. The operational efficiency, weather conditions, air traffic control, and scheduling at JFK have a significant impact on flight punctuality, safety, and passenger satisfaction. Analyzing the data related to flights departing from JFK can provide insights into patterns of delays, the influence of weather conditions on flight operations, and the performance of different airlines in managing their schedules. This analysis can help stakeholders, including airlines and airport authorities, to make informed decisions to improving operational efficiency and enhancing the overall travel experience.

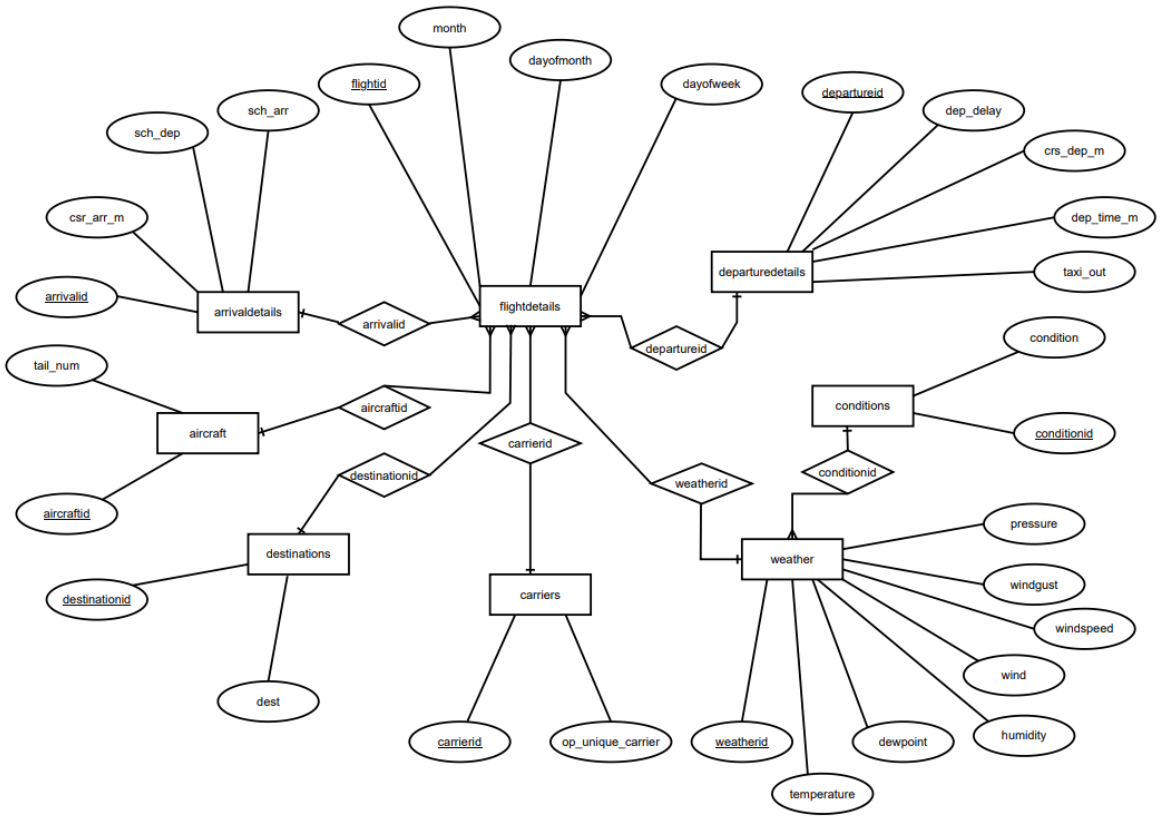
Description of Data Source

The primary source of data for this project will be the JFK Airport Flight Take-off Data available on Kaggle:
[https://www.kaggle.com/datasets/deepankurk/flight-take-off-data-jfk-airport?resource=download]
This dataset includes information on flight departures from JKF airport between Nov 2019-Dec-2020.

The dataset has 28821 rows. In order to ingest the data into a database, the CSV file obtained from Kaggle was imported into a PostgreSQL database using DataGrip. The process involved creating a database schema reflective of the dataset's structure, followed by importing the CSV file directly into the database through DataGrip's import functionality.

ER Diagram:


conditions
condition: varchar(255)
conditionid: integer
conditionid
conditionid
weather
temperature: integer
dewpoint: integer
humidity: integer
wind: varchar(255)
windspeed: integer
windgust: integer
pressure: double precision
conditionid: integer
weatherid: integer
departuredetails
dep_delay: integer
crs_dep_m: integer
dep_time_m: integer
taxi_out: integer
departureid: integer
carriers
op_unique_carrier: varchar(255)
carrierid: integer
weather
temperature: integer
dewpoint: integer
humidity: integer
wind: varchar(255)
windspeed: integer
windgust: integer
pressure: double precision
conditionid: integer
weatherid: integer
destinations
dest: varchar(255)
destinationid: integer
arrivaldetails
crs_arr_m: integer
sch_dep: integer
sch_arr: integer
arrivalid: integer
aircraft
tail_num: varchar(255)
aircraftid: integer
arrivalid
weatherid
departureid
carrierid
destinationid
aircraftid
conditions
condition: varchar(255)
conditionid: integer
aircraft
tail_num: varchar(255)
aircraftid: integer
carriers
op_unique_carrier: varchar(255)
carrierid: integer
flightdetails
month: integer
dayofmonth: integer
dayofweek: integer
carrierid: integer
aircraftid: integer
destinationid: integer
departureid: integer
arrivalid: integer
weatherid: integer
flightid: integer
departuredetails
dep_delay: integer
crs_dep_m: integer
dep_time_m: integer
taxi_out: integer
departureid: integer
destinations
dest: varchar(255)
destinationid: integer
arrivaldetails
crs_arr_m: integer
sch_dep: integer
sch_arr: integer
arrivalid: integer



Relational Schema:

The relational schema defined above organizes data pertaining to flight operations at JFK Airport, structuring the information to facilitate storage, retrieval, and analysis of flight-related data. Below is a detailed description of each table within the schema, highlighting both primary and foreign keys.

1. Carriers Table
   - Description: Stores information about flight carriers.
   - Fields: `CarrierID` (unique identifier for each carrier), `OP_UNIQUE_CARRIER` (carrier code).
   - Primary Key: `CarrierID`
   - Foreign Key(s): None.

2. Aircraft Table
   - Description: Contains details about aircraft.

- Fields: `AircraftID` (unique identifier for each aircraft), `TAIL_NUM` (aircraft tail number).
- Primary Key: `AircraftID`
- Foreign Key(s):None.

3. Destinations Table
   - Description: Lists destination airports.
   - Fields: `DestinationID` (unique identifier for each destination), `DEST` (destination airport code).
   - Primary Key:`DestinationID`
   - Foreign Key(s): None.

4. Conditions Table
   - Description: Catalogs weather conditions.
   - Fields: `ConditionID` (unique identifier for each condition), `Condition` (description of the weather condition).
   - Primary Key: `ConditionID`
   - Foreign Key(s): None.

5. Weather Table
   - Description: Records weather data relevant to flights.
   - Fields: Include `WeatherID`, `Temperature`, `DewPoint`, `Humidity`, `Wind`, `WindSpeed`, `WindGust`, `Pressure`, and `ConditionID` (links to Conditions Table).
   - Primary Key: `WeatherID`
   - Foreign Key(s):`ConditionID` (links to `Conditions.ConditionID`).

6. Departure Details Table
   - Description: Captures details about flight departures.
   - Fields: Include `DepartureID`, `DEP_DELAY`, `CRS_DEP_M`, `DEP_TIME_M`, `TAXI_OUT`.
   - Primary Key: `DepartureID`
   - Foreign Key(s): None.

7. Arrival Details Table
   - Description: Stores information on flight arrivals.
   - Fields: Include `ArrivalID`, `CRS_ARR_M`, `sch_dep`, `sch_arr`.
   - Primary Key: `ArrivalID`
   - Foreign Key(s): None.

8. Flight Details Table
   - Description: The central table that aggregates details from all other tables to represent a complete flight record.
   - Fields: Include `FlightID`, `Month`, `DayOfMonth`, `DayOfWeek`, `CarrierID`, `AircraftID`, `DestinationID`, `DepartureID`, `ArrivalID`, `WeatherID`.
   - Primary Key: `FlightID`
   - Foreign Key(s): `CarrierID` (links to `Carriers.CarrierID`), `AircraftID` (links to `Aircraft.AircraftID`), `DestinationID` (links to `Destinations.DestinationID`), `DepartureID` (links to `DepartureDetails.DepartureID`), `ArrivalID` (links to `ArrivalDetails.ArrivalID`), `WeatherID` (links to `Weather.WeatherID`).

<u>Table Create Statements:</u>

```
--Create the schema JFK Flights Schema
DROP SCHEMA jfk_flights_schema CASCADE;
CREATE SCHEMA jfk_flights_schema;

-- Carriers Table
CREATE TABLE jfk_flights_schema.Carriers (
    CarrierID SERIAL PRIMARY KEY,
    OP_UNIQUE_CARRIER VARCHAR(255) UNIQUE
);

-- Aircraft Table
CREATE TABLE jfk_flights_schema.Aircraft (
    AircraftID SERIAL PRIMARY KEY,
    TAIL_NUM VARCHAR(255) UNIQUE
);

-- Destinations Table
CREATE TABLE jfk_flights_schema.Destinations (
```

```
    DestinationID SERIAL PRIMARY KEY,
    DEST VARCHAR(255) UNIQUE
);

-- Conditions Table
CREATE TABLE jfk_flights_schema.Conditions (
    ConditionID SERIAL PRIMARY KEY,
    Condition VARCHAR(255) UNIQUE
);

-- Weather Table
CREATE TABLE jfk_flights_schema.Weather (
    WeatherID SERIAL PRIMARY KEY,
    Temperature INT,
    DewPoint INT,
    Humidity INT,
    Wind VARCHAR(255),
    WindSpeed INT,
    WindGust INT,
    Pressure FLOAT,
    ConditionID INT,
    FOREIGN KEY (ConditionID) REFERENCES Conditions(ConditionID)
);

-- Departure Details Table
CREATE TABLE jfk_flights_schema.DepartureDetails (
    DepartureID SERIAL PRIMARY KEY,
    DEP_DELAY INT,
    CRS_DEP_M INT,
    DEP_TIME_M INT,
    TAXI_OUT INT
);

-- Arrival Details Table
CREATE TABLE jfk_flights_schema.ArrivalDetails (
    ArrivalID SERIAL PRIMARY KEY,
    CRS_ARR_M INT,
    sch_dep INT,
    sch_arr INT
);
```

```
--Flight Details Table
CREATE TABLE jfk_flights_schema.FlightDetails (
    FlightID SERIAL PRIMARY KEY,
    Month INT,
    DayOfMonth INT,
    DayOfWeek INT,
    CarrierID INT,
    AircraftID INT,
    DestinationID INT,
    Distance INT,
    DepartureID INT,
    ArrivalID INT,
    WeatherID INT,
    FOREIGN KEY (CarrierID) REFERENCES
     jfk_flights_schema.Carriers(CarrierID),
    FOREIGN KEY (AircraftID) REFERENCES
     jfk_flights_schema.Aircraft(AircraftID),
    FOREIGN KEY (DestinationID) REFERENCES
     jfk_flights_schema.Destinations(DestinationID),
    FOREIGN KEY (DepartureID) REFERENCES
     jfk_flights_schema.DepartureDetails(DepartureID),
    FOREIGN KEY (ArrivalID) REFERENCES
     jfk_flights_schema.ArrivalDetails(ArrivalID),
    FOREIGN KEY (WeatherID) REFERENCES
     jfk_flights_schema.Weather(WeatherID)
);
```

<u>View Created:</u>

The `VwFlightDelaysByWeather` view provides a detailed analysis of how different weather conditions affect flight delays at JFK, including total flights, the number of delayed flights, and the average delay time. It enables quick identification of weather conditions most associated with significant flight delays.

```
CREATE OR REPLACE VIEW jfk_flights_schema.VwFlightDelaysByWeather
AS
SELECT
    con.Condition AS WeatherCondition,
    COUNT(fd.FlightID) AS TotalFlights,
    SUM(CASE WHEN dd.DEP_DELAY > 15 THEN 1 ELSE 0 END) AS
     DelayedFlights,
```

```
    AVG(dd.DEP_DELAY) AS AverageDelay
FROM
    jfk_flights_schema.FlightDetails fd
    JOIN jfk_flights_schema.DepartureDetails dd
          ON fd.DepartureID = dd.DepartureID
    JOIN jfk_flights_schema.Weather w
          ON fd.WeatherID = w.WeatherID
    JOIN jfk_flights_schema.Conditions con
          ON w.ConditionID = con.ConditionID
GROUP BY
    con.Condition
ORDER BY
    AverageDelay DESC;
```

<u>Inserting Data into Tables:</u>

The data insertion process for the `jfk_flights_schema` involved a planned approach to transfer and normalize data from the comprehensive `jfk_airport.m1_final` dataset. The normalization process was done to organize the flight information into a structured and efficient database schema consisting of several interconnected tables. Each table was designed to store specific segments of data, such as carrier information, weather conditions, flight details, and more. Here's an overview of how the data was inserted across the various tables:

- ***Distinct Value Extraction:***
  For each designated table within the schema, a `SELECT DISTINCT` query was executed against the `m1_final` dataset to extract unique values. For tables like `Carriers`, `Conditions`, and `Destinations`, this step ensured that each entry in these tables was unique and that there was no duplication of information.
- ***Handling Conflicts:***
  The `ON CONFLICT DO NOTHING` clause played a vital role in the insertion queries. It was used to gracefully handle attempts to insert duplicate entries into tables with unique constraints, thereby maintaining the integrity of the database without interrupting the data insertion process.
- ***Linking Records with Foreign Keys:***
  Inserting data into the `Weather`, `DepartureDetails`, `ArrivalDetails`, and `FlightDetails` tables required a more complex strategy due to their dependencies on

foreign keys from other tables. Subqueries were employed within the `INSERT` statements to look up and reference the correct IDs from tables like `Carriers`, `Conditions`, and `Destinations`. This ensured that each record in these tables was accurately linked to its corresponding entities across the schema.

- ***Weather Table Specifics:***
  The insertion into the `Weather` table was particularly nuanced. It involved inserting distinct weather records along with a subquery to match each weather condition to its appropriate `ConditionID` from the `Conditions` table. This step was critical to correctly associate weather observations with the predefined conditions.
- ***Comprehensive Flight Details Insertion:***
  The `FlightDetails` table, being the centerpiece, aggregated references to all other tables to provide a holistic view of each flight record. The insertion into this table was the most complex, requiring multiple subqueries to fetch and associate the correct IDs for carriers, aircraft, destinations, weather, departure, and arrival details based on the original `m1_final` data. This multi-step insertion ensured that the `FlightDetails` table accurately reflected the comprehensive and relational nature of the flight data.

The following are the insert statements:

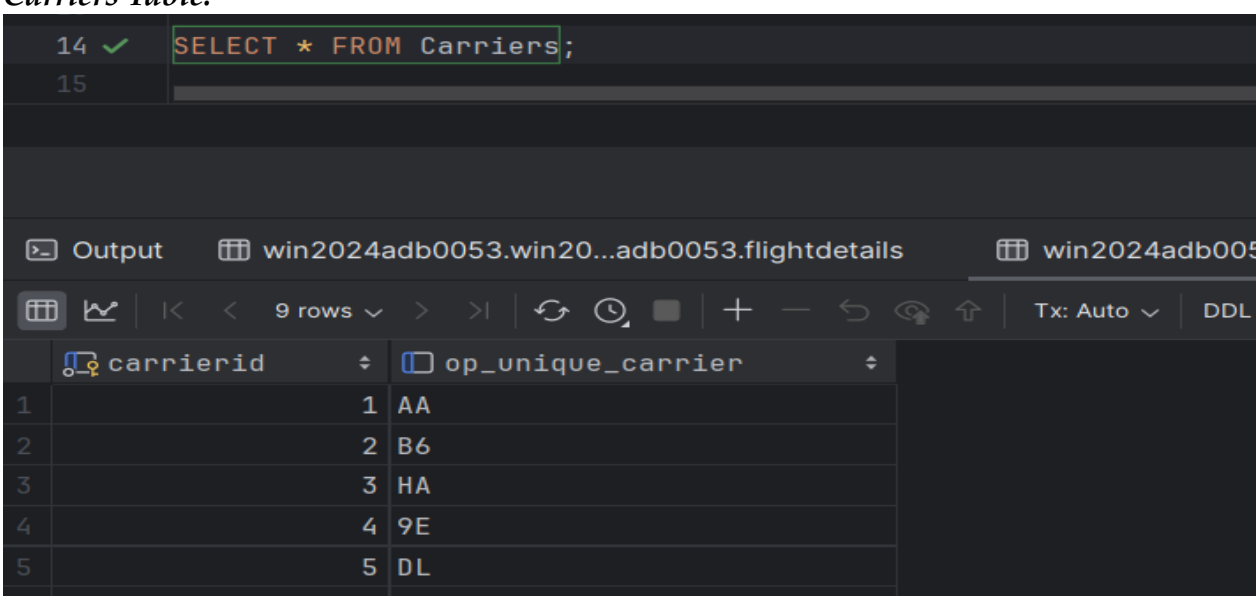
```
--Inserting data into Carriers Table
INSERT INTO jfk_flights_schema.Carriers (OP_UNIQUE_CARRIER)
SELECT DISTINCT op_unique_carrier
FROM jfk_airport.m1_final
ON CONFLICT (OP_UNIQUE_CARRIER) DO NOTHING;

--Inserting data into Conditions Table
INSERT INTO jfk_flights_schema.Conditions (Condition)
SELECT DISTINCT condition
FROM jfk_airport.m1_final
ON CONFLICT (Condition) DO NOTHING;

--Inserting data into Weather Table
INSERT INTO jfk_flights_schema.Weather
     (Temperature, DewPoint,Humidity, Wind, WindSpeed,WindGust,
     Pressure, ConditionID)
SELECT DISTINCT
    temperature, "Dew Point", humidity, wind, "Wind Speed", "Wind
     Gust", pressure,
     (SELECT ConditionID FROM Conditions c
```

```
          WHERE c.Condition =m.condition)
FROM jfk_airport.m1_final m
ON CONFLICT DO NOTHING;
--Inserting data into Departure Details Table
INSERT   INTO   jfk_flights_schema.DepartureDetails   (DEP_DELAY,
CRS_DEP_M, DEP_TIME_M, TAXI_OUT)
SELECT DISTINCT
    dep_delay, crs_dep_m, dep_time_m, taxi_out
FROM jfk_airport.m1_final;

--Inserting data into Arrival Details Table
INSERT    INTO    jfk_flights_schema.ArrivalDetails    (CRS_ARR_M,
sch_dep, sch_arr)
SELECT DISTINCT
    crs_arr_m, sch_dep, sch_arr
FROM jfk_airport.m1_final;

--Inserting data into Destinations Table
INSERT INTO jfk_flights_schema.Destinations (DEST)
SELECT DISTINCT
    dest
FROM jfk_airport.m1_final;

--Inserting data into Aircraft Table
INSERT INTO jfk_flights_schema.Aircraft (TAIL_NUM)
SELECT DISTINCT
    tail_num
FROM jfk_airport.m1_final;

--Inserting data into Flight Details Table
INSERT INTO jfk_flights_schema.FlightDetails (Month, DayOfMonth,
DayOfWeek,   CarrierID,   AircraftID,   DestinationID,   Distance,
DepartureID, ArrivalID, WeatherID)
SELECT
    month, day_of_month, day_of_week,
    (SELECT CarrierID FROM jfk_flights_schema.Carriers c
          WHERE c.OP_UNIQUE_CARRIER = m.op_unique_carrier),
    (SELECT AircraftID FROM jfk_flights_schema.Aircraft x
          WHERE x.TAIL_NUM = m.tail_num),
    (SELECT DestinationID FROM jfk_flights_schema.Destinations y
          WHERE y.DEST = m.dest),
```

```
    distance,
   (SELECT DepartureID FROM jfk_flights_schema.DepartureDetails d
        WHERE d.DEP_DELAY = m.dep_delay
          AND d.CRS_DEP_M = m.crs_dep_m
          AND d.DEP_TIME_M = m.dep_time_m
          AND d.TAXI_OUT = m.taxi_out),
    (SELECT ArrivalID FROM jfk_flights_schema.ArrivalDetails a
        WHERE a.CRS_ARR_M = m.crs_arr_m
          AND a.sch_dep = m.sch_dep
          AND a.sch_arr = m.sch_arr),
    (SELECT WeatherID FROM jfk_flights_schema.Weather w
        WHERE w.Temperature = m.temperature
          AND w.DewPoint = m."Dew Point"
          AND w.Humidity = m.humidity
          AND w.Wind = m.wind AND w.WindSpeed = m."Wind Speed"
          AND w.WindGust = m."Wind Gust"
          AND w.Pressure = m.pressure
          AND ConditionID = (SELECT ConditionID
                             FROM   jfk_flights_schema.Conditions   c
                             WHERE c.Condition = m.condition))
FROM jfk_airport.m1_final m ON CONFLICT DO NOTHING;
```

<u>Listing of 5 rows from each of the tables and row counts of all tables:</u>

Carriers Table:



	carrierid	op_unique_carrier
1	1	AA
2	2	B6
3	3	HA
4	4	9E
5	5	DL

```
SELECT count(*) FROM Carriers;
```

Output | win2024adb0053.win20...adb0053.flightdetails | count(*):bigint

1 row

count
9

Destinations Table:

```
SELECT * FROM jfk_flights_schema.Destinations;
```

Output | win2024adb0053.jfk_f...s_schema.destinations

65 rows | Tx: Auto | DDL

destinationid	dest
1	LGB
2	PSP
3	DEN
4	IND
5	ORF

```
SELECT count(*) FROM jfk_flights_schema.Destinations;
```

Output | count(*):bigint

1 row

count
65

Aircraft Table:

```
SELECT * FROM jfk_flights_schema.Aircraft;
```

Output | win2024adb0053.jfk_f...ights_schema.aircraft

1-500 of 501+ | Tx: Auto | DDL

	aircraftid	tail_num
1	1	N157UW
2	2	N889NN
3	3	N344NW
4	4	N602LR
5	5	N952AT

Tx: Auto | Playground

```
SELECT count(*) FROM jfk_flights_schema.Aircraft;
```

Output | count(*):bigint

1 row

	count
1	2074

Conditions Table:

```
SELECT * FROM Conditions;
```

4adb0053.win20...adb0053.flightdetails | win2024adb0053.win2024adb0053.conditions

24 rows | Tx: Auto | DDL

	conditionid	condition
1	1	Mostly Cloudy / Windy
2	2	Wintry Mix
3	3	Fog
4	4	Light Rain
5	5	Light Drizzle

```
SELECT count(*) FROM Conditions;
```

Output | win2024adb0053.win20...adb0053.flightdetails | count(*):bigint

1 row | CSV

count
24

Weather Table:

```
SELECT * FROM Weather;
```

Output | win2024adb0053.win20...adb0053.flightdetails | win2024adb0053.win2024adb0053.weather

1-500 of 1,819 | Tx: Auto | DDL | CSV

	weatherid	temperature	dewpoint	humidity	wind	windspeed	windgust
1	1	22	-1	37	NNW	10	0
2	2	22	12	66	NNE	5	0
3	3	22	14	72	WNW	20	33
4	4	23	-2	34	NNW	12	0
5	5	23	-1	35	NNW	9	0

```
SELECT * FROM Weather;

SELECT count(*) FROM Weather;
```

Output | win2024adb0053.win20...adb0053.flightdetails | count(*):bigint

1 row

count
1819

Departure Details Table:

```
114 ✓ SELECT * FROM DepartureDetails;
115
```

ces

Output | win2024adb0053.win20...adb0053.flightdetails | win2024adb0053.win20...0053.departuredetails

1-500 of 501+ | Tx: Auto | DDL | CSV

	departureid	dep_delay	crs_dep_m	dep_time_m	taxi_out
1	1	-6	700	694	21
2	2	-4	659	655	16
3	3	0	475	475	22
4	4	62	1259	1321	14
5	5	34	820	854	15

```
116 ✓ SELECT count(*) FROM DepartureDetails;
117
```

ces

Output | win2024adb0053.win20...adb0053.flightdetails | count(*):bigint

1 row

	count
1	23039

Arrival Details:

```
123 ✓ SELECT * FROM ArrivalDetails;
124
```

ces

Output | win2024adb0053.win20...adb0053.flightdetails | win2024adb0053.win20...db0053.arrivaldetails

1-500 of 501+ | Tx: Auto | DDL | CSV

	arrivalid	crs_arr_m	sch_dep	sch_arr
1	1	1314	35	30
2	2	959	25	25
3	3	732	50	19
4	4	588	44	13
5	5	10	22	18

```
SELECT count(*) FROM DepartureDetails;
```

Output | win2024adb0053.win20...adb0053.flightdetails | count(*):bigint

1 row

count
23039

Flight Details Table:

```
SELECT * FROM jfk_flights_schema.FlightDetails;
```

Output | win2024adb0053.jfk_f..._schema.flightdetails 2 | win2024adb0053.jfk_f..._schema.flightdetails

1-500 of 501+ | Tx: Auto | DDL | CSV

	flightid	month	dayofmonth	dayofweek	carrierid	departureid	arrivalid
1	1	11	1	5	2	5603	225
2	2	11	1	5	2	14624	38
3	3	11	1	5	2	16394	1616
4	4	11	1	5	2	4664	1430
5	5	11	1	5	5	2613	137

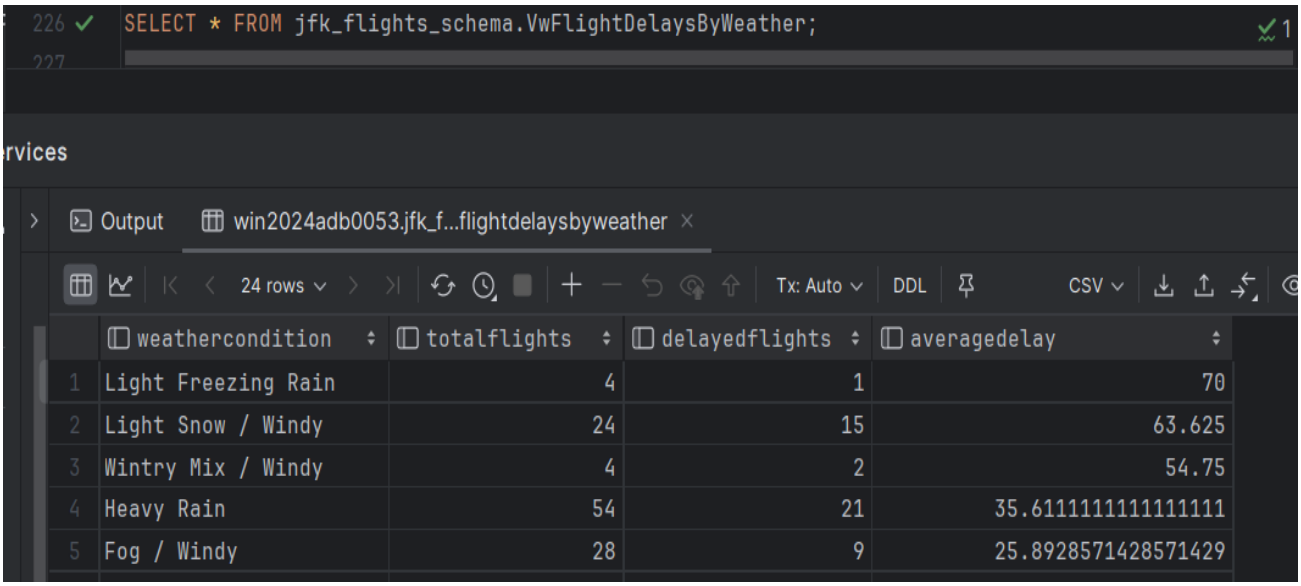
```
SELECT count(*) FROM jfk_flights_schema.FlightDetails;
```

Output | count(*):bigint

1 row

count
27095

View VwFlightDelaysByWeather



	weathercondition	totalflights	delayedflights	averagedelay
1	Light Freezing Rain	4	1	70
2	Light Snow / Windy	24	15	63.625
3	Wintry Mix / Windy	4	2	54.75
4	Heavy Rain	54	21	35.6111111111111111
5	Fog / Windy	28	9	25.8928571428571429

<u>The 20 Questions and their answer queies:</u>

Actual Question 1: How many flights departed from JFK to each destination between Nov 2019-Dec-2020?
Reason to modify: As we didn't have a year column in our dataset.
Modified Question 1: For a the month of November, how many flights departed from JFK to each destination, including the carriers and the average delay in departure for those flights?
Query:

```
SELECT
    d.DEST AS Destination,
    c.OP_UNIQUE_CARRIER AS Carrier,
    COUNT(fd.FlightID) AS NumberOfFlights,
    ROUND(AVG(dd.DEP_DELAY), 4) AS AverageDepartureDelay
FROM
      jfk_flights_schema.FlightDetails fd
JOIN
      jfk_flights_schema.Destinations  d  ON  fd.DestinationID  =
d.DestinationID
JOIN
     jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
JOIN
```

```
     jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
WHERE
     fd.Month = 11
GROUP BY
     d.DEST, c.OP_UNIQUE_CARRIER
ORDER BY
      NumberOfFlights DESC, Destination, Carrier;
```

136 rows

	destination	carrier	numberofflights	averagedeparturedelay
1	LAX	B6	325	-0.7231
2	LAX	AA	279	2.4516
3	LAX	DL	240	2.2083
4	FLL	B6	218	4.7385
5	ATL	DL	211	3.8389
6	SFO	DL	180	2.0278
7	MCO	B6	178	4.8539
8	CLT	AA	170	3.2294
9	BUF	B6	169	4.0414
10	BOS	B6	160	10.9688
11	MIA	AA	160	2.1125
12	SFO	B6	158	4.6456

Question 2: What are the average departure delays and total flight count for each airline operating out of JFK, considering all destinations?

Query:

```
SELECT
    c.OP_UNIQUE_CARRIER AS Airline,
    COUNT(fd.FlightID) AS TotalFlights,
    ROUND(AVG(dd.DEP_DELAY), 4) AS AverageDepartureDelay
FROM
    jfk_flights_schema.FlightDetails fd
JOIN
    jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
JOIN
    jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
    jfk_flights_schema.Destinations d ON fd.DestinationID =
d.DestinationID
```

```
GROUP BY
    c.OP_UNIQUE_CARRIER
ORDER BY
    AverageDepartureDelay ASC;
```

9 rows

	airline	totalflights	averagedeparturedelay
1	HA	84	0.9048
2	YX	870	3.9437
3	MQ	984	5.0742
4	AS	1050	5.3629
5	9E	4193	5.378
6	DL	7195	5.4802
7	AA	3377	5.5357
8	B6	9052	8.0896
9	OO	290	17.1586

Question 3: What flights were delayed by more than 60 minutes and by how much?
Modified Question 3: What flights were delayed by more than 60 minutes, by how much, and what were the carrier, aircraft type, destination, and weather conditions at the time of departure?
Query:

```
SELECT
    fd.FlightID,
    c.OP_UNIQUE_CARRIER AS Carrier,
    a.TAIL_NUM AS Aircraft,
    d.DEST AS Destination,
    dd.DEP_DELAY AS DelayMinutes,
    con.Condition AS WeatherCondition
FROM
    jfk_flights_schema.FlightDetails fd
JOIN
    jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
    jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
JOIN
```

```
    jfk_flights_schema.Aircraft a ON fd.AircraftID = a.AircraftID
JOIN
    jfk_flights_schema.Destinations d ON fd.DestinationID =
d.DestinationID
JOIN
    jfk_flights_schema.Weather w ON fd.WeatherID = w.WeatherID
JOIN
    jfk_flights_schema.Conditions con ON w.ConditionID =
con.ConditionID
WHERE
    dd.DEP_DELAY > 60
ORDER BY
    dd.DEP_DELAY DESC;
```

1-500 of 501+

	flightid	carrier	aircraft	destination	delayminutes	weathercondition
1	9627	OO	N286SY	ORD	1276	Fair / Windy
2	9650	DL	N709TW	SFO	1199	Partly Cloudy / Windy
3	9635	DL	N351NW	SAT	1173	Partly Cloudy / Windy
4	16098	YX	N867RW	PIT	1148	Mostly Cloudy
5	12894	AA	N786AN	MIA	1048	Cloudy
6	14071	B6	N203JB	BUF	830	Fair
7	3932	DL	N356NW	SAT	805	Mostly Cloudy
8	16385	B6	N958JB	PSE	782	Mostly Cloudy
9	22806	AA	N192UW	PHX	747	Partly Cloudy / Windy

Question 4: Are there any noticeable trends in departure delay based on time of day?
Query:

```
SELECT
    TimeOfDay,
    ROUND(AVG(AverageDelay), 4) AS AverageDelay
FROM (
    SELECT
        DEP_DELAY AS AverageDelay,
        CASE
            WHEN DEP_TIME_M >= 300 AND DEP_TIME_M < 720 THEN
'Morning'
            WHEN DEP_TIME_M >= 720 AND DEP_TIME_M < 1020 THEN
'Afternoon'
            WHEN DEP_TIME_M >= 1020 AND DEP_TIME_M < 1260 THEN
'Evening'
            ELSE 'Night'
        END AS TimeOfDay
```

```
    FROM jfk_flights_schema.DepartureDetails
) AS SubQuery
GROUP BY TimeOfDay
ORDER BY
    CASE TimeOfDay
        WHEN 'Morning' THEN 1
        WHEN 'Afternoon' THEN 2
        WHEN 'Evening' THEN 3
        ELSE 4
    END;
```

4 rows

	timeofday	averagedelay
1	Morning	2.6627
2	Afternoon	6.881
3	Evening	10.7037
4	Night	26.4964

Question 5: Which destinations have the highest incidence of departure delays?

Query:

```
SELECT
    d.DEST AS Destination,
    COUNT(*) AS TotalFlights,
    SUM(CASE WHEN dd.DEP_DELAY > 0 THEN 1 ELSE 0 END) AS
DelayedFlights,
    ROUND((CAST(SUM(CASE WHEN dd.DEP_DELAY > 0 THEN 1 ELSE 0 END)
AS NUMERIC) / CAST(COUNT(*) AS NUMERIC)) * 100, 2) AS
DelayPercentage
FROM
    jfk_flights_schema.FlightDetails fd
JOIN
    jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
    jfk_flights_schema.Destinations d ON fd.DestinationID =
d.DestinationID
```

```
GROUP BY
    d.DEST
ORDER BY
    DelayPercentage DESC, DelayedFlights DESC
LIMIT 10;
```

10 rows

	destination	totalflights	delayedflights	delaypercentage
1	JAC	4	2	50
2	SMF	77	33	42.86
3	ORH	86	35	40.7
4	ONT	77	31	40.26
5	PSE	73	29	39.73
6	IAH	238	93	39.08

Question 6: Which destination received the most flights from JFK in November 2019?

Query:

```
SELECT
    d.DEST AS Destination,
    COUNT(fd.FlightID) AS NumberOfFlights,
    c.OP_UNIQUE_CARRIER AS Carrier,
    ARRAY_AGG(DISTINCT a.TAIL_NUM) AS AircraftTailNumbers
FROM
    jfk_flights_schema.FlightDetails fd
JOIN
    jfk_flights_schema.Destinations d ON fd.DestinationID =
d.DestinationID
JOIN
    jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
JOIN
    jfk_flights_schema.Aircraft a ON fd.AircraftID = a.AircraftID
WHERE
    fd.Month = 11
GROUP BY
    d.DEST, c.OP_UNIQUE_CARRIER
ORDER BY
    NumberOfFlights DESC, Destination, Carrier
```

```
LIMIT 1;
```

	destination	numberofflights	carrier	aircrafttailnumbers
1	LAX	325	B6	{N923JB,N929JB,N934JB,N935JB,N937JB,N942JB,N943JT,N944JT,N945JT,N946JL,N947JB,N9

Question 7: Which day of the week experiences the highest average departure delay?

Query:

```
SELECT
    fd.DayOfWeek AS DayOfWeek,
    CASE
        WHEN fd.DayOfWeek = 1 THEN 'Monday'
        WHEN fd.DayOfWeek = 2 THEN 'Tuesday'
        WHEN fd.DayOfWeek = 3 THEN 'Wednesday'
        WHEN fd.DayOfWeek = 4 THEN 'Thursday'
        WHEN fd.DayOfWeek = 5 THEN 'Friday'
        WHEN fd.DayOfWeek = 6 THEN 'Saturday'
        WHEN fd.DayOfWeek = 7 THEN 'Sunday'
    END AS DayName,
    ROUND(AVG(dd.DEP_DELAY),4) AS AverageDepartureDelay,
    COUNT(fd.FlightID) AS NumberOfFlights
FROM
    jfk_flights_schema.DepartureDetails dd
JOIN
    jfk_flights_schema.FlightDetails  fd  ON  dd.DepartureID  =
fd.DepartureID
GROUP BY
    DayOfWeek
ORDER BY
    AverageDepartureDelay DESC
LIMIT 1;
```

1 row

	dayofweek	dayname	averagedepartureddelay	numberofflights
1	6	Saturday	11.0953	3358

Question 8: Which aircraft models are most frequently delayed?

Query:

```
SELECT
    c.OP_UNIQUE_CARRIER AS CarrierName,
    COUNT(*) AS DelayedFlights
FROM
    jfk_flights_schema.FlightDetails fd
JOIN
    jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
    jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
WHERE
    dd.DEP_DELAY > 0
GROUP BY
    c.OP_UNIQUE_CARRIER
ORDER BY
    DelayedFlights DESC;
```

9 rows

	carriername	delayedflights
1	B6	2786
2	DL	2076
3	9E	853
4	AA	838
5	AS	302
6	MQ	234
7	YX	147
8	OO	80
9	HA	16

Question 9: Are longer flights (in terms of distance) more susceptible to departure delays?

Query:

```
SELECT
  DistanceCategory,
  ROUND(AVG(AverageDepartureDelay),4) AS AverageDepartureDelay
FROM (
  SELECT
    CASE
      WHEN fd.Distance < 1000 THEN 'Short-haul'
      WHEN fd.Distance >= 1000 AND fd.Distance < 3000 THEN 'Medium-
haul'
      WHEN fd.Distance >= 3000 THEN 'Long-haul'
      ELSE 'Undefined'
    END AS DistanceCategory,
    dd.DEP_DELAY AS AverageDepartureDelay
  FROM
    jfk_flights_schema.FlightDetails fd
  JOIN
    jfk_flights_schema.DepartureDetails  dd  ON  fd.DepartureID  =
dd.DepartureID
) AS sub
GROUP BY
  DistanceCategory
ORDER BY
  CASE DistanceCategory
    WHEN 'Short-haul' THEN 1
    WHEN 'Medium-haul' THEN 2
    WHEN 'Long-haul' THEN 3
    ELSE 4
  END;
```

3 rows

	distancecategory	averagedeparturedelay
1	Short-haul	6.9772
2	Medium-haul	5.9226
3	Long-haul	2.0938

Question 10: How do specific holidays or events impact flight operations?

Query:

```
SELECT
  CASE
    WHEN fd.month = 11 AND fd.DayOfMonth = 24 THEN 'Thanksgiving'
    WHEN fd.month = 12 AND fd.DayOfMonth = 25 THEN 'Christmas'
    WHEN fd.month = 12 AND fd.DayOfMonth = 31 THEN 'New Year''s
Eve'
    WHEN fd.month = 1 AND fd.DayOfMonth = 1 THEN 'New Year''s Day'
    ELSE 'Other'
  END AS Holiday,
  COUNT(fd.FlightID) AS NumberOfFlights,
  ROUND(AVG(dd.DEP_DELAY), 4) AS AverageDepartureDelay
FROM
  jfk_flights_schema.FlightDetails fd
JOIN
  jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
WHERE
  (fd.month = 11 AND fd.DayOfMonth = 24)
  OR (fd.month = 12 AND fd.DayOfMonth = 25)
  OR (fd.month = 12 AND fd.DayOfMonth = 31)
  OR (fd.month = 1 AND fd.DayOfMonth = 1)
GROUP BY
  Holiday
ORDER BY
  Holiday;
```

4 rows

	holiday	numberofflights	averagedeparturedelay
1	Christmas	280	4.9607
2	New Year's Day	293	4.6928
3	New Year's Eve	282	6.5993
4	Thanksgiving	323	1.517

Question 11: Is there a difference in departure delays between weekdays and weekends?

Query:

```
SELECT
  CASE
    WHEN fd.DayOfWeek BETWEEN 1 AND 5 THEN 'Weekday'
    ELSE 'Weekend'
  END AS WeekdayOrWeekend,
  c.OP_UNIQUE_CARRIER AS Carrier,
  ROUND(AVG(dd.DEP_DELAY), 4) AS AverageDepartureDelay
FROM
  jfk_flights_schema.FlightDetails fd
JOIN
  jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
  jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
GROUP BY
  WeekdayOrWeekend, Carrier
ORDER BY
 AverageDepartureDelay DESC;
```

Output | Result 57

18 rows

	weekdayorweekend	carrier	averagedepartureddelay
1	Weekday	OO	19.2615
2	Weekend	OO	10.7917
3	Weekend	B6	10.5331
4	Weekend	9E	8.3674
5	Weekday	B6	7.1307
6	Weekend	DL	7.0178
7	Weekend	AS	6.8571
8	Weekend	MQ	6.6996
9	Weekday	AA	5.6163
10	Weekend	AA	5.3277

Question 12: On average how many flights are departing and arriving to JFK?

Query:

```
WITH DailyFlightCounts AS (
  SELECT
    fd.Month,
    fd.DayOfMonth,
    'Departure' AS FlightType,
    COUNT(*) AS TotalFlights
  FROM jfk_flights_schema.FlightDetails fd
  JOIN jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
  GROUP BY fd.Month, fd.DayOfMonth

  UNION ALL

  SELECT
    fd.Month,
    fd.DayOfMonth,
    'Arrival' AS FlightType,
    COUNT(*) AS TotalFlights
  FROM jfk_flights_schema.FlightDetails fd
  JOIN  jfk_flights_schema.ArrivalDetails  ad  ON  fd.ArrivalID  =
ad.ArrivalID
  GROUP BY fd.Month, fd.DayOfMonth
),
AverageDailyFlights AS (
  SELECT
    FlightType,
   ROUND(AVG(TotalFlights),4 )AS AverageFlightsPerDay
  FROM DailyFlightCounts
  GROUP BY FlightType
)
SELECT * FROM AverageDailyFlights;
```

2 rows

	flighttype	averageflightsperday
1	Arrival	294.5109
2	Departure	294.5109

Question 13: Do the number of flights departing and arriving to JFK vary based on holidays?

Query:

```
WITH HolidayFlights AS (
  SELECT
    fd.FlightID,
    CASE
      WHEN fd.month = 11 AND fd.DayOfMonth = 24 THEN 'Thanksgiving'
      WHEN fd.month = 12 AND fd.DayOfMonth = 25 THEN 'Christmas'
      WHEN fd.month = 12 AND fd.DayOfMonth = 31 THEN 'New Year''s
Eve'
      WHEN fd.month = 1 AND fd.DayOfMonth = 1 THEN 'New Year''s
Day'
      ELSE 'Other'
    END AS Holiday,
    'Departure' AS FlightType
  FROM
    jfk_flights_schema.FlightDetails fd
    JOIN jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
  WHERE
    (fd.month = 11 AND fd.DayOfMonth = 24)
    OR (fd.month = 12 AND fd.DayOfMonth = 25)
    OR (fd.month = 12 AND fd.DayOfMonth = 31)
    OR (fd.month = 1 AND fd.DayOfMonth = 1)
  UNION ALL
  SELECT
    fd.FlightID,
    CASE
      WHEN fd.month = 11 AND fd.DayOfMonth = 24 THEN 'Thanksgiving'
      WHEN fd.month = 12 AND fd.DayOfMonth = 25 THEN 'Christmas'
      WHEN fd.month = 12 AND fd.DayOfMonth = 31 THEN 'New Year''s
Eve'
      WHEN fd.month = 1 AND fd.DayOfMonth = 1 THEN 'New Year''s
Day'
      ELSE 'Other'
    END AS Holiday,
    'Arrival' AS FlightType
  FROM
    jfk_flights_schema.FlightDetails fd
```

```
    JOIN jfk_flights_schema.ArrivalDetails ad ON fd.ArrivalID =
ad.ArrivalID
  WHERE
    (fd.month = 11 AND fd.DayOfMonth = 24)
    OR (fd.month = 12 AND fd.DayOfMonth = 25)
    OR (fd.month = 12 AND fd.DayOfMonth = 31)
    OR (fd.month = 1 AND fd.DayOfMonth = 1)
)
SELECT
  Holiday,
  COUNT(CASE WHEN FlightType = 'Departure' THEN 1 END) AS
Departures,
  COUNT(CASE WHEN FlightType = 'Arrival' THEN 1 END) AS Arrivals
FROM
  HolidayFlights
GROUP BY
  Holiday
ORDER BY
  Holiday;
```

4 rows

	holiday	departures	arrivals
1	Christmas	280	280
2	New Year's Day	293	293
3	New Year's Eve	282	282
4	Thanksgiving	323	323

Actual Question 14: How do scheduled departure and arrival times compare to actual times across different carriers?

Reason to Modify: As we didn't have actual arrival time column in our dataset

Modified Question 14: How do scheduled departure times compare to actual departure times across different carriers?

Query:

```
SELECT
  c.OP_UNIQUE_CARRIER AS Carrier,
  dd.CRS_DEP_M AS ScheduledDepartureTime,
  dd.DEP_TIME_M AS ActualDepartureTime,
```

```
  (dd.DEP_TIME_M - dd.CRS_DEP_M) AS DepartureDelay
FROM
  jfk_flights_schema.FlightDetails fd
JOIN
  jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
JOIN
  jfk_flights_schema.DepartureDetails  dd  ON  fd.DepartureID  =
dd.DepartureID
ORDER BY
  DepartureDelay DESC;
```

1-500 of 501+

	carrier	scheduleddepartureetime	actualdepartureetime	departuredelay
1	AS	560	1181	621
2	B6	643	1228	585
3	AS	690	1210	520
4	YX	930	1363	433
5	AA	665	1091	426
6	B6	959	1379	420
7	OO	810	1224	414
8	DL	485	875	390
9	9E	474	863	389
10	AA	455	838	383

Question 15: What is the overall on-time departure rate for flights departing from JFK?

Query:

```
SELECT
  c.OP_UNIQUE_CARRIER AS Carrier,
  d.Dest AS Destination,
  w.Temperature,
  w.WindSpeed,
  cond.Condition,
  a.tail_Num AS AircraftTailNumber,
  ROUND(
    (COUNT(CASE WHEN dd.DEP_DELAY <= 15 THEN 1 END) * 100.0) /
COUNT(fd.FlightID),
    2
  ) AS OnTimeDepartureRate
```

```
FROM
  jfk_flights_schema.FlightDetails fd
JOIN
  jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
  jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
JOIN
  jfk_flights_schema.Destinations d ON fd.DestinationID =
d.DestinationID
JOIN
  jfk_flights_schema.Weather w ON fd.WeatherID = w.WeatherID
JOIN
  jfk_flights_schema.Conditions cond ON w.ConditionID =
cond.ConditionID
JOIN
  jfk_flights_schema.aircraft a ON fd.AircraftID = a.AircraftID
GROUP BY
  c.OP_UNIQUE_CARRIER, d.Dest, w.Temperature, w.WindSpeed,
cond.Condition, a.tail_num
ORDER BY
 OnTimeDepartureRate DESC;
```

1-500 of 501+ CSV

	carrier	destination	temperature	windspeed	condition	aircrafttailnumber	ontimedepartureate
1	DL	LAX	42	8	Partly Cloudy	N175DN	100
2	9E	BNA	27	21	Partly Cloudy / Windy	N279PQ	100
3	9E	BNA	28	12	Partly Cloudy	N933XJ	100
4	9E	BNA	29	3	Fair	N926XJ	100
5	9E	BNA	29	6	Mostly Cloudy	N132EV	100
6	9E	BNA	29	14	Fair	N195PQ	100
7	9E	BNA	29	17	Fair	N607LR	100
8	B6	TPA	34	20	Fair	N645JB	100
9	B6	TPA	35	0	Mostly Cloudy	N508JL	100
10	B6	TPA	35	0	Mostly Cloudy	N566JB	100

Question 16: For each type of wind condition, list the average departure delay and total number of affected flights.

Query:

```
SELECT
  wc.Condition AS WeatherCondition,
  w.Wind AS WindCondition,
```

```
  ROUND(AVG(dd.DEP_DELAY), 2) AS AverageDepartureDelay,
  COUNT(fd.FlightID) AS TotalAffectedFlights
FROM
  jfk_flights_schema.FlightDetails fd
JOIN
  jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
  jfk_flights_schema.Weather w ON fd.WeatherID = w.WeatherID
JOIN
  jfk_flights_schema.Conditions wc ON w.ConditionID =
wc.ConditionID
GROUP BY
  wc.Condition, w.Wind
ORDER BY
  AverageDepartureDelay DESC;
```

191 rows

	weathercondition	windcondition	averagedepartureddelay	totalaffectedflights
1	Light Snow	WNW	421	1
2	Light Freezing Rain	ENE	281	1
3	Light Snow / Windy	N	193.67	3
4	Light Snow / Windy	NE	166	1
5	Light Snow / Windy	WNW	138	1
6	Light Rain / Windy	NE	126	1
7	Cloudy / Windy	WSW	98.76	29
8	Light Snow / Windy	NNW	64	7
9	Wintry Mix / Windy	N	54.75	4
10	Light Snow	NNW	46.83	6

Actual Question 17: Between November 2019 and December 2020, which month has the highest number of delays?

Reason to modify: As we didn't have year column in our dataset.

Modified Question 17: Which month ,carriers and destinations experienced the highest number of departure delays from JFK, and what are the top instances of these delays?

Query:

```
SELECT
  fd.Month,
  c.OP_UNIQUE_CARRIER AS Carrier,
  d.Dest AS Destination,
```

```
  COUNT(*) AS TotalDelays
FROM
  jfk_flights_schema.FlightDetails fd
JOIN
  jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
  jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
JOIN
  jfk_flights_schema.Destinations d ON fd.DestinationID =
d.DestinationID
WHERE
  dd.DEP_DELAY > 0
GROUP BY
  fd.Month, c.OP_UNIQUE_CARRIER, d.Dest
ORDER BY
  TotalDelays DESC
LIMIT 10;
```

10 rows

	month	carrier	destination	totaldelays
1	12	B6	FLL	123
2	12	B6	LAX	101
3	12	DL	ATL	91
4	12	B6	MCO	90
5	12	AA	MIA	88
6	12	B6	SJU	87
7	11	B6	FLL	72
8	12	DL	LAX	72
9	12	DL	SFO	71
10	11	DL	ATL	69

Question 18: Are there any significant changes in departure delay patterns over the course of the dataset period (Nov 2019-Dec 2020)?

Query:

```
SELECT
  fd.Month,
  c.OP_UNIQUE_CARRIER AS Carrier,
  ROUND(AVG(dd.DEP_DELAY), 2) AS AverageDepartureDelay,
  COUNT(*) AS TotalFlights,
```

```
  SUM(CASE WHEN dd.DEP_DELAY > 0 THEN 1 ELSE 0 END) AS
TotalDelayedFlights,
  ROUND((SUM(CASE WHEN dd.DEP_DELAY > 0 THEN 1 ELSE 0 END)::DECIMAL
/ COUNT(*)) * 100, 2) AS DelayPercentage
FROM
  jfk_flights_schema.FlightDetails fd
JOIN
  jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
  jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
GROUP BY
  fd.Month, c.OP_UNIQUE_CARRIER
ORDER BY
  fd.Month, c.OP_UNIQUE_CARRIER, AverageDepartureDelay DESC;
```

26 rows

	month	carrier	averagedeparturedelay	totalflights	totaldelayedflights	delaypercentage
1	1	9E	4.96	1423	227	15.95
2	1	AA	4.71	1159	250	21.57
3	1	AS	3.94	333	98	29.43
4	1	B6	4.22	2888	729	25.24
5	1	DL	4.02	2473	610	24.67
6	1	HA	0.67	27	3	11.11
7	1	MQ	4.86	285	65	22.81
8	1	YX	2.39	344	47	13.66
9	11	9E	1.77	1426	238	16.69
10	11	AA	2.94	1117	253	22.65
11	11	AS	2.58	370	72	19.46

Question 19: How do wind speed and direction affect flight operations?

Query:

```
SELECT
  w.Wind AS WindDirection,
  w.WindSpeed,
  CAST(AVG(dd.DEP_DELAY) AS DECIMAL(10, 4)) AS
AverageDepartureDelay,
  COUNT(*) AS TotalFlights,
  SUM(CASE WHEN dd.DEP_DELAY > 0 THEN 1 ELSE 0 END) AS
TotalDelayedFlights,
```

```
  CAST((SUM(CASE WHEN dd.DEP_DELAY > 0 THEN 1 ELSE 0 END)::FLOAT
/ COUNT(*)) * 100 AS DECIMAL(10, 4)) AS DelayPercentage
FROM
  jfk_flights_schema.FlightDetails fd
JOIN
  jfk_flights_schema.DepartureDetails dd ON fd.DepartureID =
dd.DepartureID
JOIN
  jfk_flights_schema.Weather w ON fd.WeatherID = w.WeatherID
GROUP BY
  w.Wind, w.WindSpeed
ORDER BY
  AverageDepartureDelay DESC;
```

293 rows CSV

	winddirection	windspeed	averagedepartureddelay	totalflights	totaldelayedflights	delaypercentage
1	WNW	36	167.0000	1	1	100.0000
2	NE	25	126.8750	8	8	100.0000
3	N	24	103.5000	2	2	100.0000
4	NNE	21	102.3333	12	11	91.6667
5	WSW	31	98.7586	29	24	82.7586
6	N	23	77.0000	6	3	50.0000
7	SW	18	73.2000	5	4	80.0000
8	WSW	22	61.3750	8	7	87.5000
9	WSW	20	56.1778	45	31	68.8889
10	W	5	51.9000	10	7	70.0000

Question 20: What percentage of flights depart early in the morning (before 6 AM) and their punctuality?

Query:

```
SELECT
  c.OP_UNIQUE_CARRIER AS Carrier,
  d.Dest AS Destination,
  CAST(CAST(SUM(CASE WHEN FLOOR(dd.DEP_TIME_M / 100) < 6 THEN 1
ELSE 0 END) AS FLOAT) / NULLIF(COUNT(*), 0) * 100 * 10000 AS
INTEGER) / 10000.0 AS EarlyMorningFlightsPercentage,

  CAST(CAST(SUM(CASE WHEN FLOOR(dd.DEP_TIME_M / 100) < 6 AND
dd.DEP_DELAY <= 0 THEN 1 ELSE 0 END) AS FLOAT) / NULLIF(SUM(CASE
WHEN FLOOR(dd.DEP_TIME_M / 100) < 6 THEN 1 ELSE 0 END), 0) * 100
* 10000 AS INTEGER) / 10000.0 AS PunctualityRate
FROM
  jfk_flights_schema.DepartureDetails dd
```

```
JOIN
  jfk_flights_schema.FlightDetails fd ON dd.DepartureID =
fd.DepartureID
JOIN
  jfk_flights_schema.Carriers c ON fd.CarrierID = c.CarrierID
JOIN
  jfk_flights_schema.Destinations d ON fd.DestinationID =
d.DestinationID
GROUP BY
  c.OP_UNIQUE_CARRIER, d.Dest
HAVING
  SUM(CASE WHEN FLOOR(dd.DEP_TIME_M / 100) < 6 THEN 1 ELSE 0 END)
> 0 AND
  SUM(CASE WHEN FLOOR(dd.DEP_TIME_M / 100) < 6 AND dd.DEP_DELAY <=
0 THEN 1 ELSE 0 END) > 0
ORDER BY
  EarlyMorningFlightsPercentage DESC, PunctualityRate DESC;
```

124 rows

	carrier	destination	earlymorningflightspercentage	punctualityrate
1	9E	MSY	100	100
2	DL	SRQ	100	100
3	YX	MSY	100	100
4	DL	STT	100	84.0909
5	YX	DTW	100	78.5714
6	AS	LAS	100	50
7	AA	DCA	98.75	87.3418
8	DL	SJC	98.4127	87.0968
9	YX	SRQ	92.3077	75
10	AS	SJC	78.5714	81.8182