DBMS Deliverable 3

Student 1 Name and Odin ID: Nida Mariam Sheikh Aslam – (nidama)

Student 2 Name and Odin ID: Snehil Shrivastava (snehils)

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?
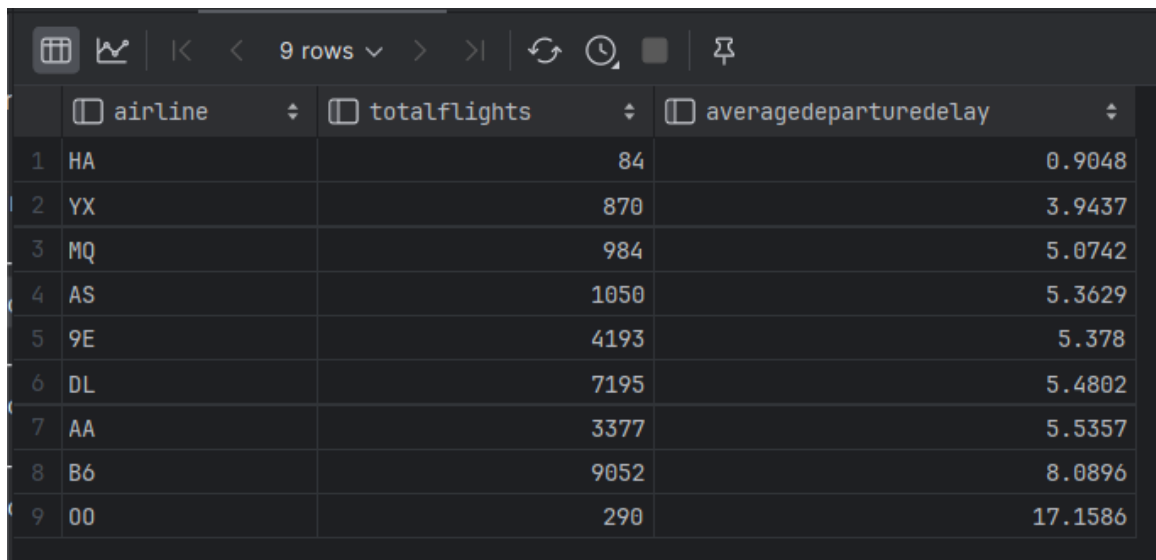
Query:

```
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;
```

	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:

```sql
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;
```

	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?
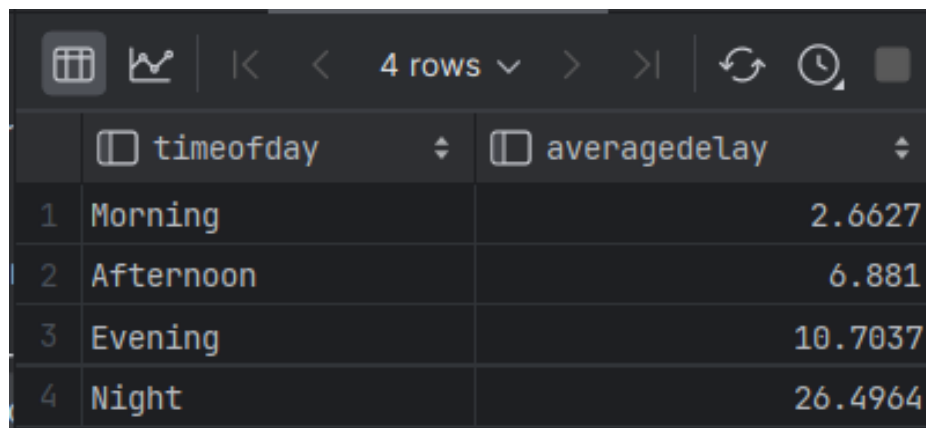
Query:

```sql
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;
```

	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
10	23310	9E	N8969A	BWI	729	Mostly Cloudy
11	14382	B6	N658JB	JAX	627	Cloudy
12	3592	9E	N8877A	ORF	624	Light Rain
13	5282	AS	N292AK	SEA	621	Light Rain
14	15018	B6	N597JB	TPA	585	Mostly Cloudy
15	12209	9E	N928XJ	ORF	585	Fair
16	20040	9E	N835AY	BWI	563	Cloudy
17	22377	9E	N303PQ	BWI	543	Mostly Cloudy
18	4249	9E	N8837B	BWI	541	Mostly Cloudy
19	23282	9E	N336PQ	RDU	526	Mostly Cloudy

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

Query:

```sql
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;
```

	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:
```sql
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;
```

	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
7	DEN	247	92	37.25
8	ATL	743	275	37.01
9	FLL	945	347	36.72
10	SJU	747	274	36.68

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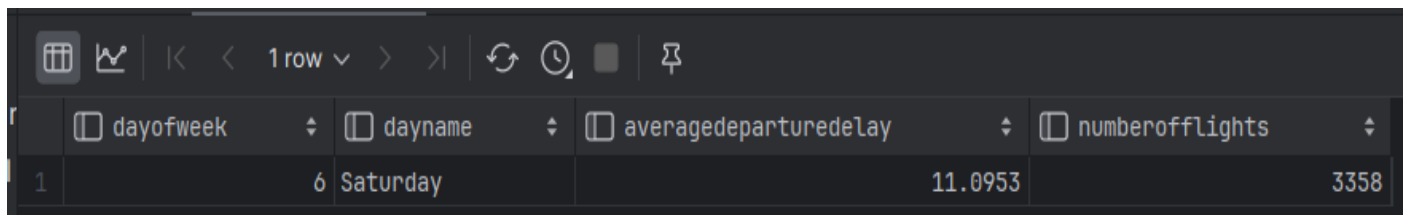
Query:

```sql
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
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:
```sql
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;
```
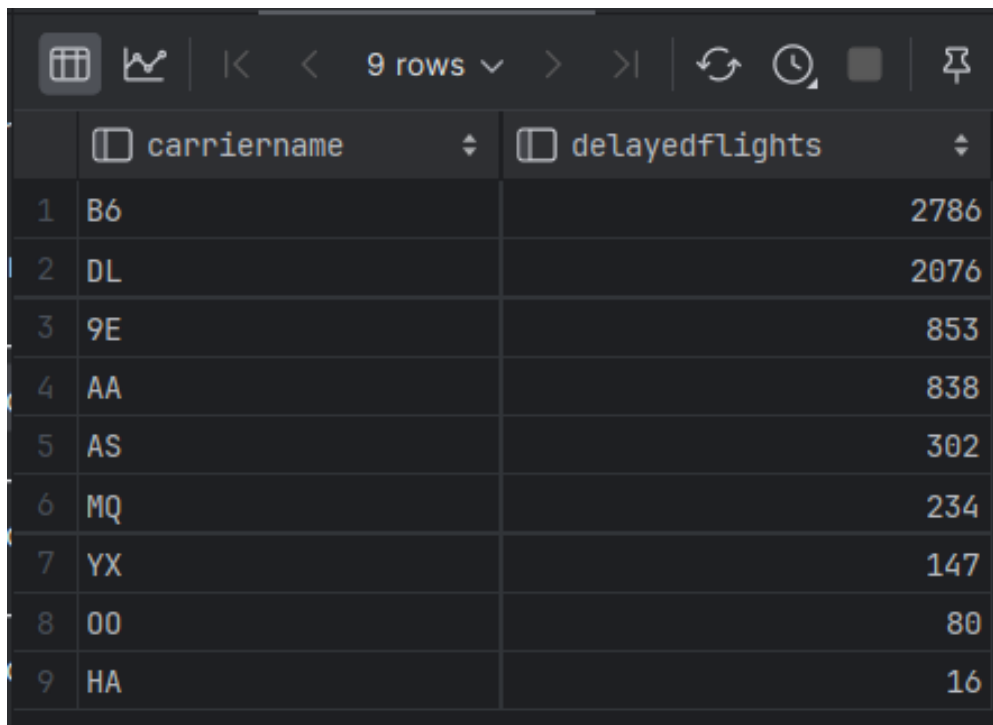
	dayofweek	dayname	averagedeparturedelay	numberofflights
1	6	Saturday	11.0953	3358

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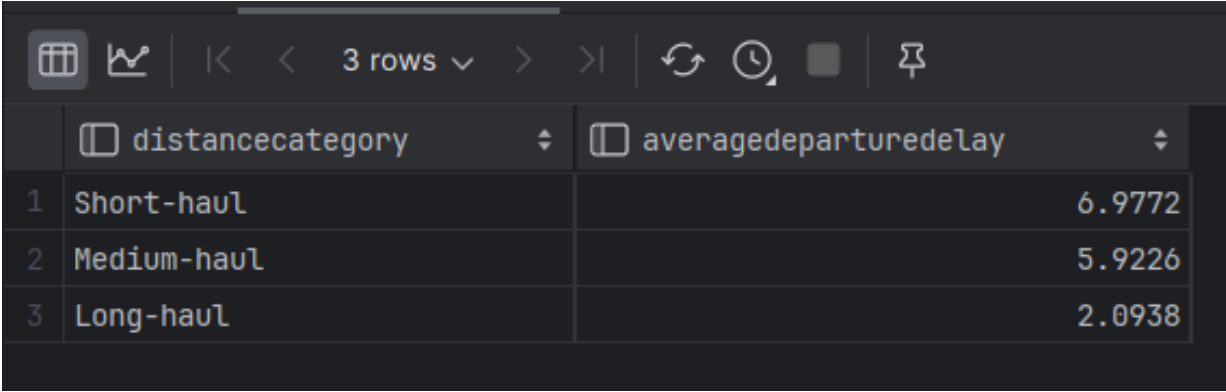
	carriername	delayedflights
1	B6	2786
2	DL	2076
3	9E	853
4	AA	838
5	AS	302
6	MQ	234
7	YX	147
8	OO	80
9	HA	16

9 rows

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

Query:
```sql
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;
```

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?
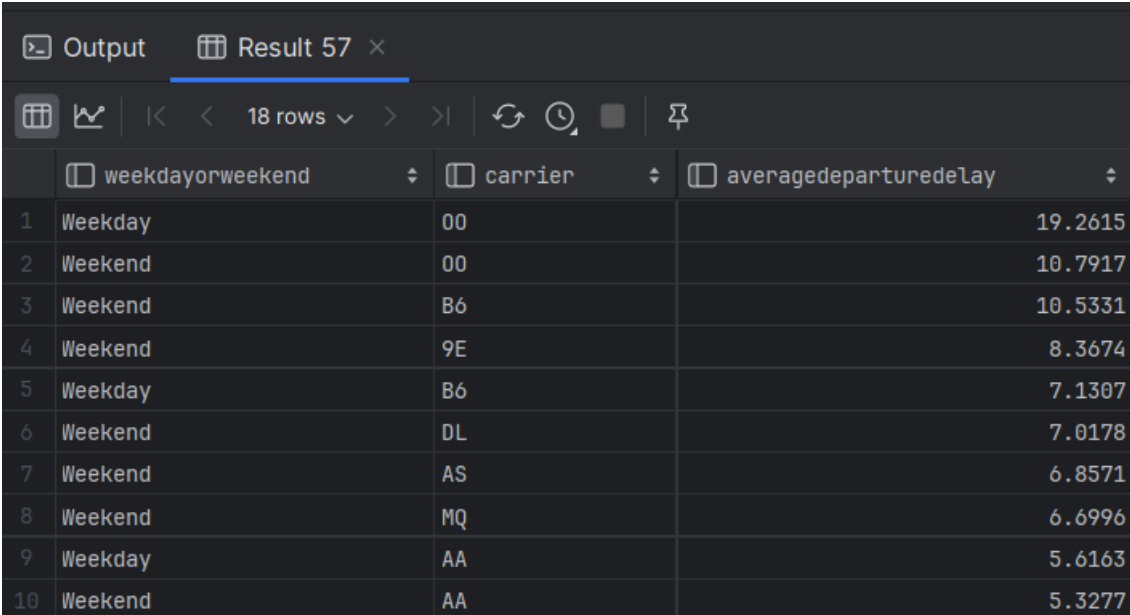
Query:
```sql
SELECT
  CASE
    WHEN fd.month = 11 AND fd.DayOfMonth = 24 THEN 'Thanksgiving'
    WHEN fd.month = 12 AND fd.DayOfMonth = 25 THEN 'Christmas'
    WHEN fd.month = 12 AND fd.DayOfMonth = 31 THEN 'New Year''s Eve'
    WHEN fd.month = 1 AND fd.DayOfMonth = 1 THEN 'New Year''s Day'
    ELSE '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;
```

	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 ˅
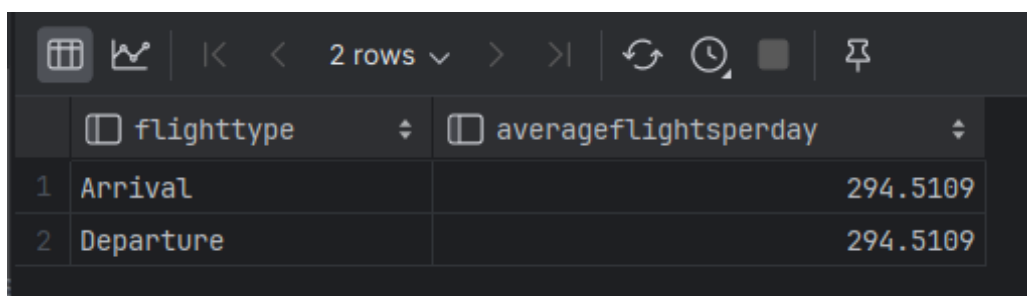
weekdayorweekend	carrier	averagedeparturedelay
Weekday	OO	19.2615
Weekend	OO	10.7917
Weekend	B6	10.5331
Weekend	9E	8.3674
Weekday	B6	7.1307
Weekend	DL	7.0178
Weekend	AS	6.8571
Weekend	MQ	6.6996
Weekday	AA	5.6163
Weekend	AA	5.3277

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;
```

	flighttype	averageflightsperday
1	Arrival	294.5109
2	Departure	294.5109

2 rows

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

Query:

```sql
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;
```

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;
```

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

Query:
```sql
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;
```

	carrier	destination	temperature	windspeed	condition	aircrafttailnumber	ontimedeparturerate
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

1-500 ∨ of 501+   CSV ∨

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

	weathercondition	windcondition	averagedeparturedelay	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

191 rows

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:
```sql
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;
```

	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;
```

	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
12	11	B6	3.46	3101	796	25.67
13	11	DL	2.65	2396	594	24.79
14	11	HA	-5.39	28	1	3.57
15	11	MQ	-1.12	339	52	15.34
16	11	OO	9.02	246	56	22.76
17	11	YX	1.26	162	20	12.35
18	12	9E	9.65	1344	388	28.87
19	12	AA	9.04	1101	335	30.43
20	12	AS	9.69	347	132	38.04

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;
```

	winddirection	windspeed	averagedeparturedelay	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:
```sql
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

View:

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.

```sql
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,
    ROUND(AVG(dd.DEP_DELAY),4) 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;
```

24 rows

	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	35.6111
5	Fog / Windy	28	9	25.8929
6	Cloudy / Windy	341	73	18.695
7	Fair / Windy	479	91	15.3695
8	Light Snow	76	17	15.1447
9	Partly Cloudy / Windy	558	110	14.7867
10	Rain	351	97	14.4957
11	Wintry Mix	83	21	12.3373
12	Light Rain	1865	324	9.5957
13	Light Rain / Windy	288	43	6.2361
14	Mostly Cloudy / Windy	1442	211	5.8558
15	Fair	4408	572	5.8044
16	Partly Cloudy	2974	393	5.2616
17	Cloudy	4902	574	5.2336
18	Mostly Cloudy	8820	1037	4.8478
19	Light Drizzle	194	19	3.9588
20	Rain / Windy	25	4	3.8