Bazy Danych, Proj. analityczny: linie lotnicze

Jędrzej Rybczyński

17.04.2021

1. Jakie było średnie opóźnienie przylotu?

```
SELECT AVG(arr_delay_new) AS [avg_delay]
FROM [dbad_flights].[dbo].[Flight_delays];
```

Table 1: 1 records

avg_delay
15.91152

2. Jakie było maksymalne opóźnienie przylotu?

```
SELECT MAX(arr_delay_new) AS [max_delay]
FROM [dbad_flights].[dbo].[Flight_delays];
```

Table 2: 1 records

max_delay
1895

3. Który lot miał największe opóźnienie przylotu?

```
SELECT arr_delay_new, fl_date, carrier, origin_city_name, dest_city_name
FROM [dbad_flights].[dbo].[Flight_delays]
WHERE arr_delay_new = (SELECT MAX(arr_delay_new)
                       FROM [dbad_flights].[dbo].[Flight_delays])
```

Table 3: 1 records

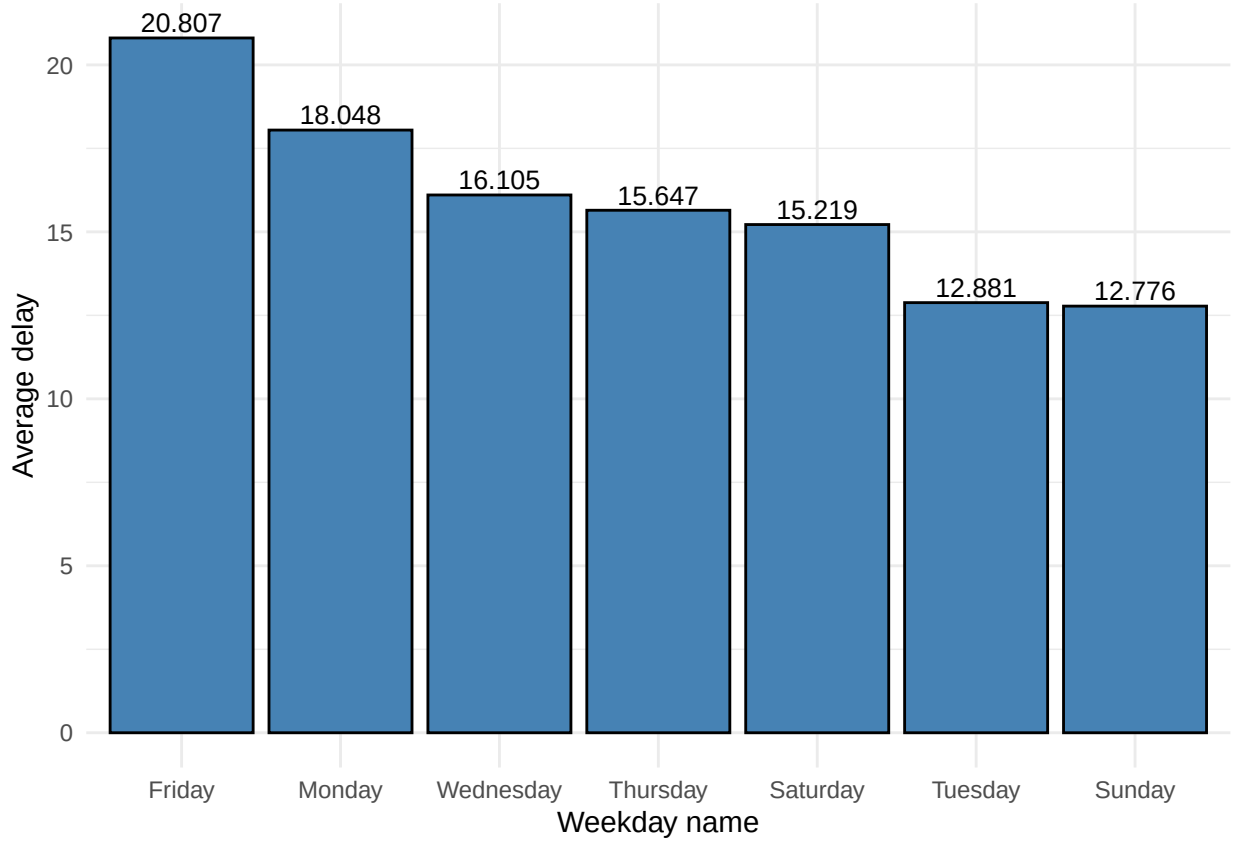
arr_delay_new	fl_date	carrier	origin_city_name	dest_city_name
1895	2017-07-26	AA	Kona, HI	Los Angeles, CA

4. Które dni tygodnia są najgorsze do podróżowania?

```
SELECT AVG(arr_delay_new) AS [avg_delay], W.weekday_name
FROM [dbad_flights].[dbo].[Flight_delays] D
INNER JOIN [dbad_flights].[dbo].[Weekdays] W
        ON D.day_of_week = W.weekday_id
GROUP BY W.weekday_name
ORDER BY avg_delay DESC;
```

Table 4: 7 records

avg_delay	weekday_name
20.80747	Friday
18.04801	Monday
16.10514	Wednesday
15.64696	Thursday
15.21876	Saturday
12.88056	Tuesday
12.77606	Sunday

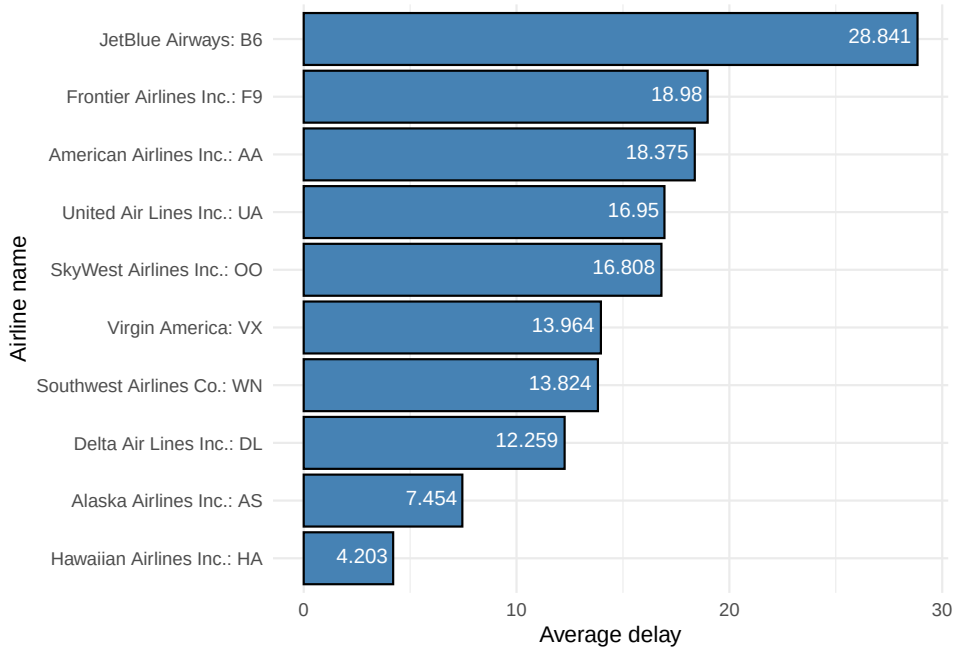


5. Które linie lotnicze latające z San Francisco (SFO) mają najmniejsze opóźnienia przylotu?

```
SELECT *
FROM (SELECT AVG(D.arr_delay_new) as [avg_delay], A.airline_name
        FROM [dbad_flights].[dbo].[Flight_delays] D
        INNER JOIN [dbad_flights].[dbo].[Airlines] A
                ON A.airline_id = D.airline_id
        GROUP BY A.airline_name) AS T
WHERE T.airline_name IN (SELECT A.airline_name
                         FROM   [dbad_flights].[dbo].[Airlines] A
                         WHERE  A.airline_name = T.airline_name
                                AND A.airline_id IN (SELECT D.airline_id
                                                     FROM   [dbad_flights].[dbo].[Flight_delays] D
                                                     WHERE  D.origin = 'SFO'
                                                            AND A.airline_id = D.airline_id))
ORDER BY T.avg_delay DESC;
```

Table 5: Displaying records 1 - 10

avg_delay	airline_name
28.841148	JetBlue Airways: B6
18.980300	Frontier Airlines Inc.: F9
18.375314	American Airlines Inc.: AA
16.950403	United Air Lines Inc.: UA
16.808273	SkyWest Airlines Inc.: OO
13.964467	Virgin America: VX
13.823983	Southwest Airlines Co.: WN
12.258788	Delta Air Lines Inc.: DL
7.453927	Alaska Airlines Inc.: AS
4.202719	Hawaiian Airlines Inc.: HA



6. Jaka część linii lotniczych ma regularne opóźnienia, tj. jej lot ma średnio co najmniej 10 min. opóźnienia?

```
SELECT  CAST(COUNT(TEMP.avg_delay1) AS DECIMAL) / CAST(COUNT(TEMP1.avg_delay2) AS DECIMAL)
        AS [late_propotion]
FROM   (SELECT AVG(D.arr_delay_new) as [avg_delay1]
        FROM   [dbad_flights].[dbo].[Flight_delays] D
        GROUP BY airline_id
        HAVING AVG(D.arr_delay_new) >= 10) AS TEMP
RIGHT OUTER JOIN (SELECT AVG(D.arr_delay_new) as [avg_delay2]
                  FROM   [dbad_flights].[dbo].[Flight_delays] D
                  GROUP BY airline_id) TEMP1
             ON  TEMP.avg_delay1 = TEMP1.avg_delay2;
```

Table 6: 1 records

late_propotion
0.8333333

7. Jak opóźnienia wylotów wpływają na opóźnienia przylotów?

```
SELECT (AVG(arr_delay_new * dep_delay_new) - (AVG(arr_delay_new) * AVG(dep_delay_new)))
       / (STDEVP(arr_delay_new) * STDEVP(dep_delay_new)) AS [Pearsons r]
FROM   [dbad_flights].[dbo].[Flight_delays]
```

Table 7: 1 records

Pearsons r
0.97371

8. Która linia lotnicza miała największy wzrost (różnica) średniego opóźnienia przylotów w ostatnim tygodniu miesiąca, tj. między 1-23 a 24-31 lipca?

```
WITH CTE_Delay_Increase
AS
(
    SELECT T1.airline_name, MAX(T2.delay - T1.delay1) AS [difference]
       FROM  (SELECT AVG(D.arr_delay_new) as [delay1], A1.airline_name
              FROM   [dbad_flights].[dbo].[Flight_delays] D
              INNER JOIN [dbad_flights].[dbo].[Airlines] A1
                      ON  D.airline_id = A1.airline_id AND D.day_of_month >= 1
                      AND D.day_of_month <= 23
              GROUP BY A1.airline_name) T1
       INNER JOIN (SELECT AVG(D.arr_delay_new) as [delay], A.airline_name
                   FROM   [dbad_flights].[dbo].[Flight_delays] D
                   INNER JOIN [dbad_flights].[dbo].[Airlines] A
                           ON  D.airline_id = A.airline_id AND D.day_of_month >= 24
                           AND D.day_of_month <= 31
                   GROUP BY A.airline_name) T2
                ON  T1.airline_name = T2.airline_name
       GROUP BY T1.airline_name
)
SELECT C.difference AS [delay_increase], C.airline_name
FROM   CTE_Delay_Increase C
WHERE  C.difference = (SELECT MAX(C.difference)
                       FROM CTE_Delay_Increase C);
```

Table 8: 1 records

delay_increase	airline_name
0.584763	Southwest Airlines Co.: WN

9. Które linie lotnicze latają zarówno na trasie SFO → PDX (Portland), jak i SFO → EUG (Eugene)?

```
SELECT DISTINCT A1.airline_name
FROM    [dbad_flights].[dbo].[Flight_delays] D
INNER JOIN [dbad_flights].[dbo].[Airlines] A1
        ON D.airline_id = A1.airline_id AND D.origin = 'SFO' AND D.dest = 'PDX'
INNER JOIN [dbad_flights].[dbo].[Flight_delays] D1
        ON D1.airline_id = A1.airline_id AND D1.origin = 'SFO' AND D1.dest = 'EUG'
```

Table 9: 2 records

airline_name
SkyWest Airlines Inc.: OO
United Air Lines Inc.: UA

10. Jak najszybciej dostać się z Chicago do Stanfordu, zakładając wylot po 14:00 czasu lokalnego?

```
SELECT  AVG(D.arr_delay_new) AS [avg_delay], D.origin, D.dest
FROM    [dbad_flights].[dbo].[Flight_delays] D
WHERE   D.origin IN ('MDW', 'ORD') AND D.dest IN ('SFO', 'SJC', 'OAK') AND D.crs_dep_time >= 1400
GROUP   BY D.origin, D.dest
ORDER BY AVG(D.arr_delay_new) DESC;
```

Table 10: 6 records

avg_delay	origin	dest
22.19253	ORD	SFO
19.85714	MDW	SFO
17.20000	MDW	SJC
15.51613	ORD	OAK
14.81111	ORD	SJC
12.12903	MDW	OAK