

# Linie lotnicze - projekt 2023

Barbara Ratajczak

2023-04-14

**Zadanie 1** Jakie było średnie opóźnienie przylotu?

```
SELECT AVG(arr_delay_new) 'avg_delay'
FROM   Flight_delays;
```

Table 1: Wyniki

avg_delay
15.91152

**Zadanie 2** Jakie było maksymalne opóźnienie przylotu?

```
SELECT MAX(arr_delay_new) 'max_delay'
FROM   Flight_delays;
```

Table 2: Wyniki

max_delay
1895

**Zadanie 3** Który lot miał największe opóźnienie przylotu?

```
SELECT carrier, origin_city_name, dest_city_name, fl_date, arr_delay_new
FROM   Flight_delays
WHERE  arr_delay_new = (SELECT MAX(arr_delay_new) 'max_delay'
                        FROM   Flight_delays);
```

Table 3: Wyniki

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

**Zadanie 4** Które dni tygodnia są najgorsze do podróżowania?

```
SELECT  weekday_name, AVG(arr_delay_new) 'avg_delay'
FROM    Flight_delays
JOIN    Weekdays
        ON weekday_id = day_of_week
GROUP BY weekday_name
ORDER BY avg_delay DESC;
```

Table 4: Wyniki

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

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

```
SELECT  airline_name, AVG(F2.arr_delay_new) 'avg delay'
FROM    Flight_delays F
JOIN    Airlines A
        ON A.airline_id = F.airline_id
JOIN    Flight_delays F2
        ON F2.airline_id = F.airline_id
WHERE   F.origin = 'SFO'
GROUP BY airline_name
ORDER BY "avg delay" DESC;
```

Table 5: Wyniki

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

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

```

SELECT CAST(COUNT(*) AS REAL)/(SELECT COUNT(DISTINCT airline_id) FROM Flight_delays) AS 'late_proportion'
FROM   (SELECT   airline_name, AVG(arr_delay_new) 'avg delay'
        FROM     Flight_delays F
        JOIN     Airlines A
              ON  A.airline_id = F.airline_id
        GROUP BY airline_name) AS t
WHERE  "avg delay" > 10;

```

Table 6: Wyniki

late_proportion
0.8333333

**Zadanie 7** Jak opóźnienia wylotów wpływają na opóźnienia przylotów?

```

WITH T1(avg_arr_delay, avg_dep_delay, std_arr, std_dep)
AS (
    SELECT AVG(arr_delay_new),
           AVG(dep_delay_new),
           STDEVP(arr_delay_new),
           STDEVP(dep_delay_new)
    FROM   Flight_delays
    WHERE  arr_delay_new > 0),
T2(cov)
AS (
    SELECT SUM((arr_delay_new - avg_arr_delay)*(dep_delay_new - avg_dep_delay))/COUNT(arr_delay_new)
    FROM   T1
    JOIN   Flight_delays
          ON arr_delay_new > 0),
T3(Pearson_r)
AS (
    SELECT cov/(std_arr*std_dep)
    FROM   T2
    CROSS JOIN T1)
SELECT Pearson_r AS [Pearson r]
FROM   T3;

```

Table 7: Wyniki

Pearson r
0.9735109

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

```

WITH T1(airline, avg_del_before)
AS (
    SELECT airline_name, AVG(arr_delay_new)
    FROM   Flight_delays F

```

```

        JOIN      Airlines A
              ON A.airline_id = F.airline_id
        WHERE     day_of_month BETWEEN 1 AND 23
        GROUP BY  airline_name),
T2(airline, avg_del_after)
AS (
    SELECT      airline_name, AVG(arr_delay_new)
    FROM        Flight_delays F
    JOIN        Airlines A
              ON A.airline_id = F.airline_id
    WHERE       day_of_month BETWEEN 24 AND 31
    GROUP BY    airline_name),
T3(airline, avg_before, avg_after)
AS (
    SELECT T1.airline, avg_del_before, avg_del_after
    FROM    T1
    JOIN    T2
          ON T1.airline = T2.airline)
SELECT     airline,
          avg_after-avg_before AS 'delay_increase'
FROM       T3
WHERE      avg_after-avg_before = (SELECT MAX(avg_after-avg_before)
                                FROM T3);

```

Table 8: Wyniki

airline	delay_increase
Southwest Airlines Co.: WN	0.584763

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

```

SELECT DISTINCT airline_name
FROM    Flight_delays F
JOIN    Airlines A
      ON A.airline_id = F.airline_id
WHERE   origin = 'SFO' AND dest = 'PDX'
      AND airline_name IN (SELECT DISTINCT airline_name
                          FROM    Flight_delays F
                          JOIN    Airlines A
                                ON A.airline_id = F.airline_id
                          WHERE   origin = 'SFO' AND dest = 'EUG');

```

Table 9: Wyniki

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

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

```
SELECT  origin, dest, AVG(arr_delay_new) 'avg delay'
FROM    Flight_delays
WHERE   (origin = 'MDW' OR origin = 'ORD')
        AND (dest = 'OAK' OR dest = 'SJC' OR dest = 'SFO')
        AND (crs_dep_time > 1400)
GROUP BY origin, dest
ORDER BY [avg delay] DESC;
```

Table 10: Wyniki

origin	dest	avg delay
ORD	SFO	22.19253
MDW	SFO	19.85714
MDW	SJC	17.20000
ORD	SJC	14.81111
MDW	OAK	12.12903