

Linie lotnicze Projekt 2022

Paweł Bielko

25 04 2022

Projekt nr 2 - Bazy danych - Linie lotnicze

Projekt przedstawiający polecenia wraz z wygenerowanymi tabelami z analizy bazy danych dotyczących opóźnień połączeń lotniczych w USA w lipcu 2017 r.

Zadanie 1

Jakie było średnie opóźnienie przylotu?

```
SELECT AVG(arr_delay_new) AS 'avg_delay'
FROM dbo.Flight_delays
WHERE arr_delay_new IS NOT NULL;
```

Table 1: 1 records

avg_delay
15.91152

Zadanie 2

Jakie było maksymalne opóźnienie przylotu?

```
SELECT MAX(arr_delay_new) AS 'max_delay'
FROM dbo.Flight_delays
WHERE arr_delay_new IS NOT NULL
```

Table 2: 1 records

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 dbo.Flight_delays
WHERE arr_delay_new >= ALL(SELECT arr_delay_new
                           FROM dbo.Flight_delays
                           WHERE arr_delay_new IS NOT NULL)
```

Table 3: 1 records

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 CASE WHEN day_of_week = 1 THEN 'Monday'
            WHEN day_of_week = 2 THEN 'Tuesday'
            WHEN day_of_week = 3 THEN 'Wednesday'
            WHEN day_of_week = 4 THEN 'Thursday'
            WHEN day_of_week = 5 THEN 'Friday'
            WHEN day_of_week = 6 THEN 'Saturday'
            WHEN day_of_week = 7 THEN 'Sunday'
            END AS 'weekday_name',
        AVG(arr_delay_new) AS 'avg_delay'
FROM dbo.Flight_delays
GROUP BY day_of_week
ORDER BY AVG(arr_delay_new) DESC;
```

Table 4: 7 records

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(arr_delay_new) AS avg_delay
FROM Flight_delays f JOIN Airlines a ON f.airline_id = a.airline_id
WHERE airline_name IN
(SELECT a.airline_name
 FROM Flight_delays f JOIN Airlines a ON f.airline_id = a.airline_id
 WHERE f.origin = 'SFO')
GROUP BY airline_name
ORDER BY AVG(arr_delay_new) DESC
```

Table 5: Displaying records 1 - 10

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(SELECT count(*)
from(select count(*) AS 'delay'
FROM dbo.Flight_delays
group by airline_id
having avg(arr_delay_new)>10) as A)*1.0/(SELECT count(*)
from(select count(*) AS 'delay'
FROM dbo.Flight_delays
group by airline_id) as B) as late_proportion
```

Table 6: 1 records

late_proportion
0.8333333

Zadanie 7

Jak opóźnienia wylotów wpływają na opóźnienia przylotów?

```
Select(((select count(*)
from Flight_delays )*
(select sum(isnull(arr_delay_new,0)*
isnull(dep_delay_new,0))
from Flight_delays))-
((select sum(isnull(dep_delay_new,0))
from Flight_delays)*
(select sum(isnull(arr_delay_new,0))
from Flight_delays)))/
((sqrt((select count(*)
from Flight_delays )*
(select sum(isnull(dep_delay_new,0)*
isnull(dep_delay_new,0))
from Flight_delays))-
(select sum(isnull(dep_delay_new,0))
```

```

from Flight_delays)*
(select sum(isnull(dep_delay_new,0))
from Flight_delays))*
(sqrt(((select count(*)
from Flight_delays )*
(select sum(isnull(arr_delay_new,0)*
isnull(arr_delay_new,0))
from Flight_delays))-
((select sum(isnull(arr_delay_new,0))
from Flight_delays)*
(select sum(isnull(arr_delay_new,0))
from Flight_delays)))) as 'Pearsons r'

```

Table 7: 1 records

Pearsons r
0.9717058

Zadanie 8

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

```

SELECT TOP 1 A.airline_name,(avg_delay_1-avg_delay_2) AS [delay_increase]
FROM Airlines A
JOIN (SELECT airline_id, AVG(arr_delay_new) AS [avg_delay_2]
FROM Flight_delays
WHERE month=7 AND (day_of_month BETWEEN 1 AND 23)
GROUP BY airline_id) F
ON F.airline_id=A.airline_id
JOIN (SELECT airline_id, AVG(arr_delay_new) AS [avg_delay_1]
FROM Flight_delays
WHERE month=7 AND (day_of_month BETWEEN 24 AND 31)
GROUP BY airline_id) P
ON P.airline_id=A.airline_id
ORDER BY delay_increase DESC

```

Table 8: 1 records

airline_name	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 P.airline_name from Flight_delays F
join Airlines P on P.airline_id=F.airline_id
where origin='SFO' and dest='PDX'
INTERSECT
select P.airline_name from Flight_delays F

```

```
join Airlines P on P.airline_id=F.airline_id
where origin='SFO' and dest='EUG'
```

Table 9: 2 records

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*1.000000000)as avg_delay
from Flight_delays
where crs_dep_time>1400
and (origin='ORD' or origin='MDW')
and (dest='SFO' or dest='OAK' or dest='SJC')
group by origin, dest
order by avg_delay DESC
```

Table 10: 5 records

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