

Projekt analityczny: linie lotnicze

26.04.2021

```
library(odbc)

## Warning: package 'odbc' was built under R version 4.0.5

library(DBI)
con <- DBI::dbConnect(odbc::odbc(),
  Driver = "SQL Server",
  Server = "mssql-2017.labs.wmi.amu.edu.pl",
  Database = "dbad_flights",
  Trusted_Connection = "True")
```

Pytania i odpowiedzi

a. Jakie było średnie opóźnienie przylotu?

```
SELECT AVG(arr_delay_new) [avg_delay]
FROM Flight_delays
```

Table 1: 1 records

avg_delay
15.91152

b. Jakie było maksymalne opóźnienie przylotu?

```
SELECT MAX(arr_delay_new) AS max_delay FROM Flight_delays
```

Table 2: 1 records

max_delay
1895

c. Który lot miał największe opóźnienie przylotu?

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

Table 3: 1 records

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

d.Które dni tygodnia są najgorsze do podróżowania?

```
SELECT AVG(arr_delay_new) AS avg_delay , w.weekday_name
FROM Flight_delays f JOIN Weekdays w ON f.day_of_week = w.weekday_id
GROUP BY w.weekday_name
ORDER BY AVG(arr_delay_new) 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

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

```
SELECT AVG(arr_delay_new) AS avg_delay, airline_name
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

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

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

```
SELECT (SELECT SUM(ile)
FROM (SELECT COUNT(DISTINCT a.airline_name) AS ile
FROM Flight_delays f JOIN Airlines a ON f.airline_id = a.airline_id
GROUP BY a.airline_name
HAVING AVG(arr_delay_new) > 10) AS razem) * 1.0 /
(SELECT COUNT(DISTINCT airline_id) FROM Flight_delays) AS late_proportion
```

Table 6: 1 records

late_proportion
0.8333333

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

```
SELECT DISTINCT a.airline_name
FROM Flight_delays f JOIN Airlines a ON f.airline_id = a.airline_id
WHERE a.airline_id IN
(SELECT DISTINCT airline_id FROM Flight_delays
WHERE origin = 'SFO' AND dest = 'PDX' AND airline_id IN
(SELECT DISTINCT airline_id FROM Flight_delays
WHERE origin = 'SFO' AND dest = 'EUG'))
```

Table 7: 2 records

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

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

```
SELECT AVG(arr_delay_new) AS 'avg_delay', origin, dest
FROM Flight_delays
WHERE (origin = 'MDW' OR origin = 'ORD') AND (dest IN ('SFO', 'SJC', 'OAK'))
AND crs_dep_time > 1400
GROUP BY origin, dest
ORDER BY AVG(arr_delay_new) DESC
```

Table 8: 5 records

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