Projekt analityczny: linie lotnicze

Jędrzej Konopko

22 04 2020

Zadanie 1

```
zadanie1 <-DBI::dbGetQuery(con, "SELECT
                           AVG(arr_delay_new) AS [avg_delay]
                           FROM Flight_delays;")
zadanie1
```

```
##   avg_delay
## 1  15.91152
```

Zadanie 2

```
zadanie2 <-DBI::dbGetQuery(con, "SELECT
                           MAX(arr_delay_new) AS [max_delay]
                           FROM Flight_delays;")
zadanie2
```

```
##   max_delay
## 1      1895
```

Zadanie 3

```
zadanie3 <-DBI::dbGetQuery(con, "SELECT
                           TOP 1 arr_delay_new,fl_date,
                           carrier,origin_city_name,dest_city_name
                           FROM Flight_delays
                           WHERE arr_delay_new IS NOT NULL
                           ORDER BY arr_delay_new DESC;")
zadanie3
```

```
##   arr_delay_new    fl_date carrier origin_city_name  dest_city_name
## 1          1895 2017-07-26      AA         Kona, HI Los Angeles, CA
```

Zadanie 4

```
zadanie4 <-DBI::dbGetQuery(con, "SELECT
    AVG(F.arr_delay_new) AS [avg_delay],
    W.weekday_name
    FROM Flight_delays F
        LEFT OUTER JOIN Weekdays W
```

```
        ON F.day_of_week=W.weekday_id
    GROUP BY W.weekday_name
    ORDER BY avg_delay DESC;")
zadanie4
```

```
##   avg_delay weekday_name
## 1  20.80747       Friday
## 2  18.04801       Monday
## 3  16.10514    Wednesday
## 4  15.64696     Thursday
## 5  15.21876     Saturday
## 6  12.88056      Tuesday
## 7  12.77606       Sunday
```

Zadanie 5

```
zadanie5 <-DBI::dbGetQuery(con, "SELECT AVG(F.arr_delay_new) AS [avg_delay],
       A.airline_name
FROM Flight_delays F
    LEFT OUTER JOIN Airlines A
        ON F.airline_id=A.airline_id
GROUP BY A.airline_name
HAVING SUM(CASE WHEN F.origin='SFO' THEN 1 ELSE 0 END) > 0
ORDER BY avg_delay DESC;")
zadanie5
```

```
##    avg_delay               airline_name
## 1  28.841148        JetBlue Airways: B6
## 2  18.980300 Frontier Airlines Inc.: F9
## 3  18.375314 American Airlines Inc.: AA
## 4  16.950403  United Air Lines Inc.: UA
## 5  16.808273  SkyWest Airlines Inc.: OO
## 6  13.964467         Virgin America: VX
## 7  13.823983 Southwest Airlines Co.: WN
## 8  12.258788   Delta Air Lines Inc.: DL
## 9   7.453928   Alaska Airlines Inc.: AS
## 10  4.202719 Hawaiian Airlines Inc.: HA
```

Zadanie 6

```
zadanie6 <-DBI::dbGetQuery(con, "SELECT
((SELECT COUNT(*)
FROM(
SELECT AVG(F.arr_delay_new) AS [avg_del]
FROM Flight_delays F
LEFT OUTER JOIN Airlines A
ON F.airline_id=A.airline_id
GROUP BY A.airline_name
HAVING AVG(arr_delay_new)>10) AS [del_airlines])*1.0)/
```

```
((SELECT COUNT(*)
FROM(
SELECT AVG(F.arr_delay_new) AS [avg_del]
FROM Flight_delays F
LEFT OUTER JOIN Airlines A
ON F.airline_id=A.airline_id
GROUP BY A.airline_name) AS [all_airlines])*1.0) AS [late_proportion]")
zadanie6
```

```
##   late_proportion
## 1       0.8333333
```

Zadanie 7

```
zadanie7 <-DBI::dbGetQuery(con, "SELECT (Avg((dep_delay_new)*(arr_delay_new))
-(Avg(dep_delay_new)*Avg(arr_delay_new)))
/(StDevP(dep_delay_new)*StDevP(arr_delay_new)) AS [Pearsons r]
FROM Flight_delays;")
zadanie7
```

```
##   Pearsons r
## 1    0.97371
```

Zadanie 8

```
zadanie8 <-DBI::dbGetQuery(con, "SELECT TOP 1 -(aa.avg_delay1- bb.avg_delay2)
AS [delay_increase],aa.airline_name
FROM
(SELECT AVG(F.arr_delay_new) AS [avg_delay1],
       A.airline_name
FROM Flight_delays F
    LEFT OUTER JOIN Airlines A
        ON F.airline_id=A.airline_id
WHERE fl_date BETWEEN '2017-07-01' AND '2017-07-23'
GROUP BY A.airline_name) AS aa,

(SELECT AVG(F.arr_delay_new) AS [avg_delay2],
       A.airline_name
FROM Flight_delays F
    LEFT OUTER JOIN Airlines A
        ON F.airline_id=A.airline_id
WHERE fl_date BETWEEN '2017-07-24' AND '2017-07-31'
GROUP BY A.airline_name) as bb
where aa.airline_name =bb.airline_name
ORDER BY delay_increase DESC")
zadanie8
```

```
##   delay_increase               airline_name
## 1       0.584763 Southwest Airlines Co.: WN
```

Zadanie 9

```
zadanie9 <-DBI::dbGetQuery(con, "SELECT
       A.airline_name
FROM Flight_delays F
    LEFT OUTER JOIN Airlines A
        ON F.airline_id=A.airline_id
WHERE origin='SFO' AND dest='EUG'
AND A.airline_name IN (
SELECT
       A.airline_name
FROM Flight_delays F
    LEFT OUTER JOIN Airlines A
        ON F.airline_id=A.airline_id
WHERE origin='SFO' AND dest='PDX'

GROUP BY A.airline_name
)
GROUP BY A.airline_name
ORDER BY A.airline_name")
zadanie9
```

```
##                airline_name
## 1 SkyWest Airlines Inc.: OO
## 2 United Air Lines Inc.: UA
```
