11. Birthday attendance

Given the following two tables, write a query to return the fraction of students, rounded to two decimal places, who attended school (attendance = 1) on their birthday.

```python
import pandas as pd
import numpy  as np

data1 = {'student_id'    : [1,2,3,1,2,3,1,2,3,4],
         'school_date'    : ['03-APR-20',
                             '03-APR-20',
                             '03-APR-20',
                             '04-APR-20',
                             '04-APR-20',
                             '04-APR-20',
                             '05-APR-20',
                             '05-APR-20',
                             '05-APR-20',
                             '05-APR-20'],
          'attendance'    : [0,1,1,1,1,1,0,1,1,1]
        }

data2 = {'student_id'    : [1,2,3,4],
          'school_id'     : [2,1,1,2],
          'grade'         : [5,4,3,4],
          'date_birth'    : ['03-APR-12',
                             '04-APR-13',
                             '05-APR-14',
                             '03-APR-13']
        }
attendance  = pd.DataFrame(data1)
students    = pd.DataFrame(data2)


attendance['school_date']=pd.to_datetime(attendance['school_date'],format="%d-%b-%y")
print(f'attendance table:\n{attendance}')
```

```
attendance table:
     student_id school_date  attendance
0             1  2020-04-03           0
1             2  2020-04-03           1
2             3  2020-04-03           1
3             1  2020-04-04           1
4             2  2020-04-04           1
5             3  2020-04-04           1
6             1  2020-04-05           0
7             2  2020-04-05           1
8             3  2020-04-05           1
9             4  2020-04-05           1
```

```python
students['date_birth']=pd.to_datetime(students['date_birth'],format="%d-%b-%y")
print(f'students table:\n{students}')
```

```
students table:
    student_id  school_id  grade date_birth
0            1          2      5 2012-04-03
1            2          1      4 2013-04-04
2            3          1      3 2014-04-05
3            4          2      4 2013-04-03
```

```
df1=attendance[['student_id','attendance']].copy()
df1['month']=attendance['school_date'].dt.month
df1['day']=attendance['school_date'].dt.day
print(f'day and month attendance:\n{df1}')
```

```
day and month attendance:
    student_id  attendance  month  day
0            1           0      4    3
1            2           1      4    3
2            3           1      4    3
3            1           1      4    4
4            2           1      4    4
5            3           1      4    4
6            1           0      4    5
7            2           1      4    5
8            3           1      4    5
9            4           1      4    5
```

```
df2=students[['student_id']].copy()
df2['month']=students['date_birth'].dt.month
df2['day']=students['date_birth'].dt.day
print(f'day and month birthday:\n{df2}')
```

```
day and month birthday:
    student_id  month  day
0            1      4    3
1            2      4    4
2            3      4    5
3            4      4    3
```

```
df3=pd.merge(df2,df1,on=['student_id','month','day'],how='left')
print(f'attendance on birthday dates:\n{df3}')
```

```
attendance on birthday dates:
    student_id  month  day  attendance
0            1      4    3         0.0
1            2      4    4         1.0
2            3      4    5         1.0
3            4      4    3         NaN
```

```
attendance_copy=attendance[['student_id','attendance']].copy()
attendance_copy['month']=attendance['school_date'].dt.month
attendance_copy['day']=attendance['school_date'].dt.day
students_copy=students[['student_id']].copy()
students_copy['month']=students['date_birth'].dt.month
students_copy['day']=students['date_birth'].dt.day
```

```python
attendance_on_birthday=(pd.merge(students_copy,
                                 attendance_copy,
                                 on=['student_id','month','day'],
                                 how='left'
                        )
)
fraction=attendance_on_birthday['attendance'].mean()
print(f'fraction of birthday attendance:\n{round(fraction,2)}')
```

```
fraction of birthday attendance:
0.67
```