08. Upgrade rate by product action

Given the following two tables, return the fraction of users, rounded to two decimal places, who accessed feature two (type: F2 in events table) and upgraded to premium within the first 30 days of signing up.

```
import pandas as pd
import numpy  as np

data1 = {'user_id'   : [1,2,3,4,5,6,7],
         'name'      : ['John','Jane','Jill','Josh','Jean','Justin','Jeremy'],
         'join_date' : ['14-feb-20',
                        '14-feb-20',
                        '15-feb-20',
                        '15-feb-20',
                        '16-feb-20',
                        '17-feb-20',
                        '18-feb-20']
        }

data2 = {'user_id'     : [1,2,2,3,4,1,3],
         'Type'        : ['F1','F2','P','F2','F2','P','P'],
         'access_date' : ['1-mar-20',
                          '2-mar-20',
                          '12-mar-20',
                          '15-mar-20',
                          '15-mar-20',
                          '16-mar-20',
                          '22-mar-20']
        }

users  = pd.DataFrame(data1)
events = pd.DataFrame(data2)


users['join_date'] = pd.to_datetime(users['join_date'])
print(f'users:\n{users}')
```

```
users:
   user_id    name  join_date
0        1    John 2020-02-14
1        2    Jane 2020-02-14
2        3    Jill 2020-02-15
3        4    Josh 2020-02-15
4        5    Jean 2020-02-16
5        6  Justin 2020-02-17
6        7  Jeremy 2020-02-18
<ipython-input-49-7f5c3f43b052>:1: UserWarning: Could not infer format, so each
  users['join_date'] = pd.to_datetime(users['join_date'])
```

```
events['access_date'] = pd.to_datetime(events['access_date'])
print(f'events:\n {events}')
```

```
events:
    user_id Type access_date
0         1   F1  2020-03-01
1         2   F2  2020-03-02
2         2    P  2020-03-12
3         3   F2  2020-03-15
4         4   F2  2020-03-15
5         1    P  2020-03-16
6         3    P  2020-03-22
<ipython-input-50-b31e0c40a0cf>:1: UserWarning: Could not infer format, so each
  events['access_date'] = pd.to_datetime(events['access_date'])
```

```
df1=(users.drop(columns=['name'])
          .merge(events.query("Type=='F2'")[['user_id']]
                 ,on='user_id'
                 ,how='inner'
           )
)
print(f'Who are F2 Users:\n{df1}')
```

```
Who are F2 Users:
   user_id  join_date
0        2 2020-02-14
1        3 2020-02-15
2        4 2020-02-15
```

```
df2=events.query("Type=='P'")[['user_id','access_date']]
print(f'Who are Premium Users:\n{df2}')
```

```
Who are Premium Users:
   user_id access_date
2        2  2020-03-12
5        1  2020-03-16
6        3  2020-03-22
```

```
rop(columns=['name'])
erge(events.query("Type=='F2'")[['user_id']]
     ,on='user_id'
     ,how='inner'

erge(events.query("Type=='P'")[['user_id','access_date']]
     ,on='user_id'
     ,how='left'


irst30']=premium_upgrade['access_date']-premium_upgrade['join_date']<=pd.Timedelta(days=30)
hin first 30 days:\n{premium_upgrade}\n')
um_upgrade['WithinFirst30'].mean()


pgrade_rate}')
```

```
Was upgraded within first 30 days:
   user_id  join_date access_date  WithinFirst30
0        2 2020-02-14  2020-03-12           True
```

```
1         3 2020-02-15  2020-03-22          False
2         4 2020-02-15         NaT          False

upgrade_rate: 0.33
```