

# How can I calculate the time delta in months using Pandas?

Authored by  
**stats writer**

July 1, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can I calculate the time delta in months using Pandas?*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=164868>

Pandas is a popular data analysis library in Python that offers a variety of functions for manipulating and analyzing data. One of its useful features is the ability to calculate the time delta in months between two dates. This can be achieved by using the built-in function "pd.to\_timedelta()" which converts a series of dates into a time delta object. Then, by utilizing the "months" attribute, the time delta in months can be calculated. This functionality is particularly helpful in analyzing time series data and making comparisons between different time periods.

## Pandas: Calculate Timedelta in Months

You can use the following function to calculate a timedelta in months between two columns of a pandas DataFrame:

```
def month_diff(x, y):  
    end = x.dt.to_period('M').view(dtype='int64')  
    start = y.dt.to_period('M').view(dtype='int64')  
    return end-start
```

The following example shows how to use this function in practice.

Example: Calculate Timedelta in Months in Pandas

Suppose we have the following pandas DataFrame:

```
import pandas as pd
```

```
#create DataFrame
```

```
df = pd.DataFrame({'event': ,  
'start_date': ,  
'end_date': })
```

**#convert start date and end date columns to datetime**

```
df = pd.to_datetime(df)
```

```
df = pd.to_datetime(df)
```

**#view DataFrame**

```
print(df)
```

```
event start_date end_date
```

```
0 A 2021-01-01 2021-06-08
```

```
1 B 2021-02-01 2021-02-09
```

```
2 C 2021-04-01 2021-08-01
```

Now suppose we'd like to calculate the `timedelta` (in months) between the `start_date` and `end_date` columns.

To do so, we'll first define the following function:

**#define function to calculate timedelta in months between two columns**

```
def month_diff(x, y):
```

```
end = x.dt.to_period('M').view(dtype='int64')
```

```
start = y.dt.to_period('M').view(dtype='int64')  
return end-start
```

Next, we'll use this function to calculate the `timedelta` in months between the `start_date` and `end_date` columns:

```
#calculate month difference between start date and end  
date columns
```

```
df = month_diff(df.end_date, df.start_date)
```

```
#view updated DataFrame
```

```
df
```

```
event start_date end_date month_difference
```

```
0 A 2021-01-01 2021-06-08 5
```

```
1 B 2021-02-01 2021-02-09 0
```

```
2 C 2021-04-01 2021-08-01 4
```

The `month_difference` column displays the `timedelta` (in months) between the `start_date` and `end_date` columns.

Additional Resources

The following tutorials explain how to perform other common operations in pandas: