Weighted window functions: part 2

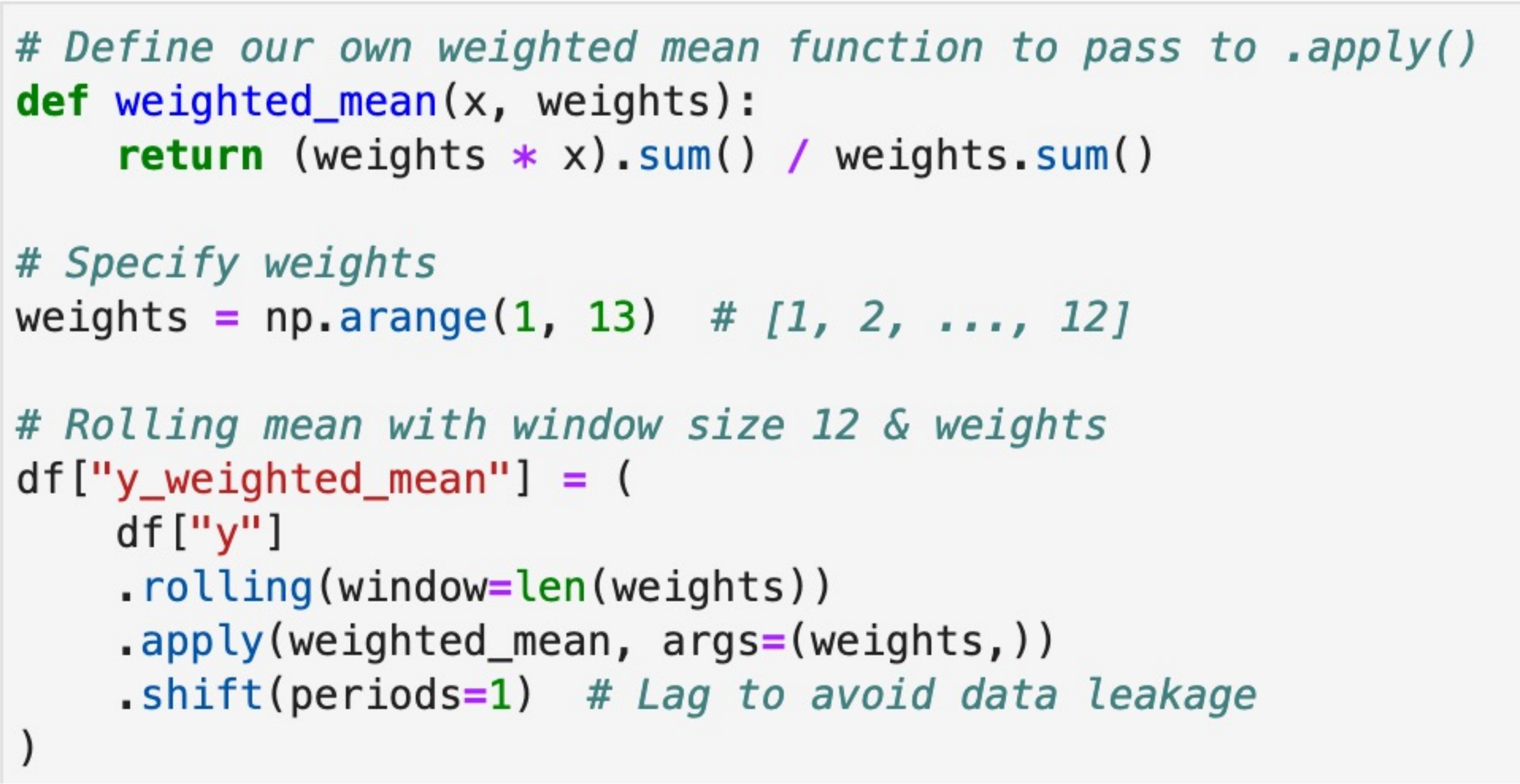
Window features

Weighted statistics

- Each statistic (e.g., the mean) has a formula for a weighted version (e.g., weighted mean).

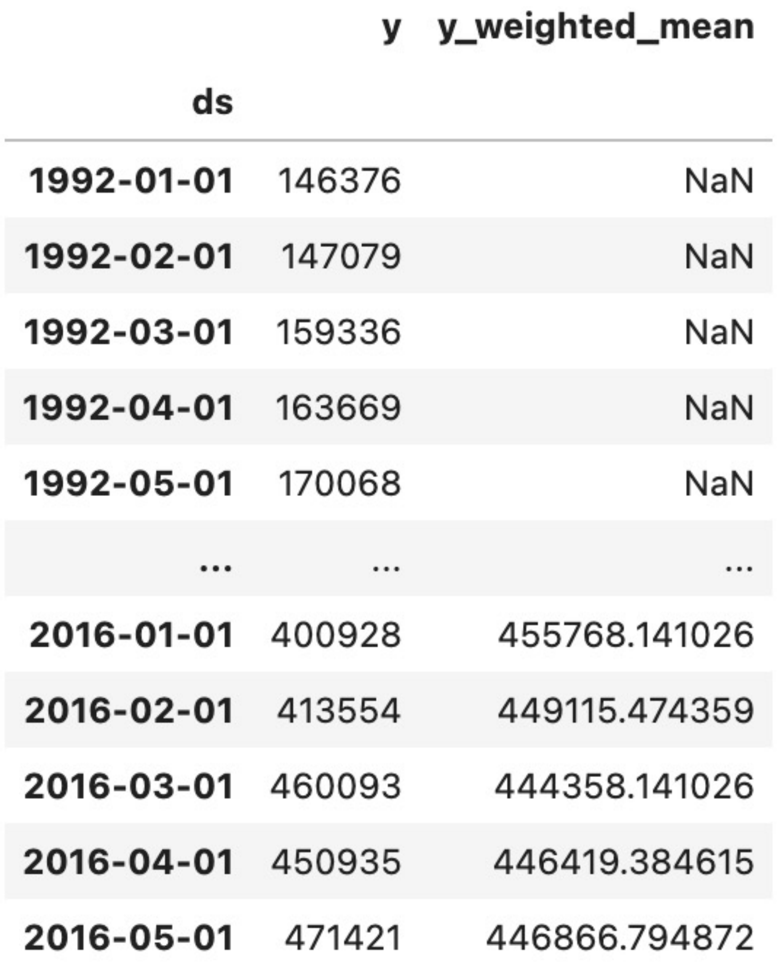
	Unweighted	**Weighted**
Mean	$\hat{\mu} = \frac{\sum_{i=1}^{N} x_i}{N}$	$\hat{\mu}_w = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i}$
Variance	$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N} (x_i - \hat{\mu})^2}{N}$	$\hat{\sigma}_w{}^2 = \frac{\sum_{i=1}^{N} w_i (x_i - \hat{\mu}_w)^2}{\sum_{i=1}^{N} w_i}$

Weighted window in Pandas



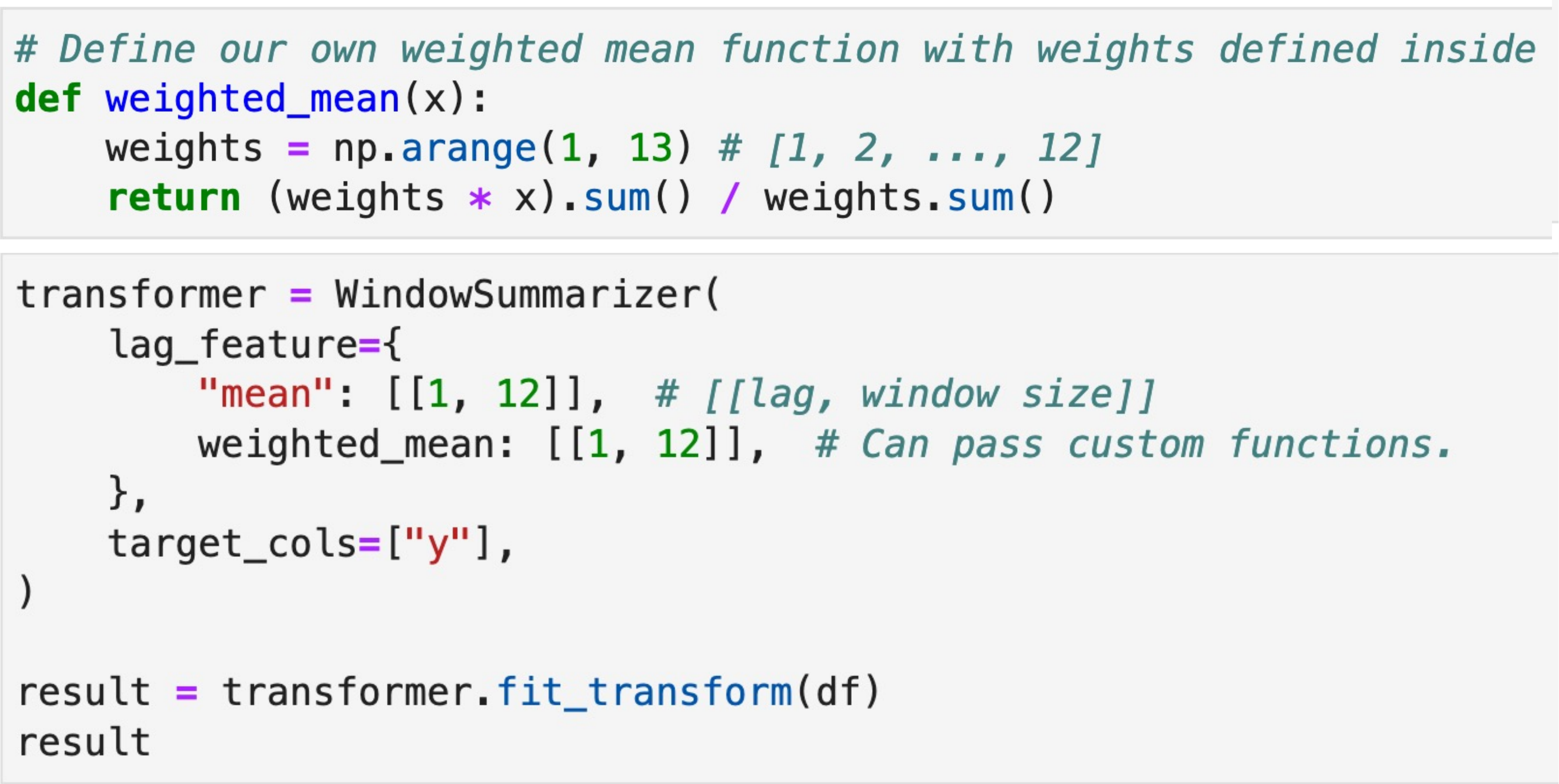
```
# Define our own weighted mean function to pass to .apply()
def weighted_mean(x, weights):
    return (weights * x).sum() / weights.sum()

# Specify weights
weights = np.arange(1, 13)  # [1, 2, ..., 12]

# Rolling mean with window size 12 & weights
df["y_weighted_mean"] = (
    df["y"]
    .rolling(window=len(weights))
    .apply(weighted_mean, args=(weights,))
    .shift(periods=1)  # Lag to avoid data leakage
)
```

Weighted window in Pandas



	y	y_weighted_mean
ds		
1992-01-01	146376	NaN
1992-02-01	147079	NaN
1992-03-01	159336	NaN
1992-04-01	163669	NaN
1992-05-01	170068	NaN
...	...	...
2016-01-01	400928	455768.141026
2016-02-01	413554	449115.474359
2016-03-01	460093	444358.141026
2016-04-01	450935	446419.384615
2016-05-01	471421	446866.794872

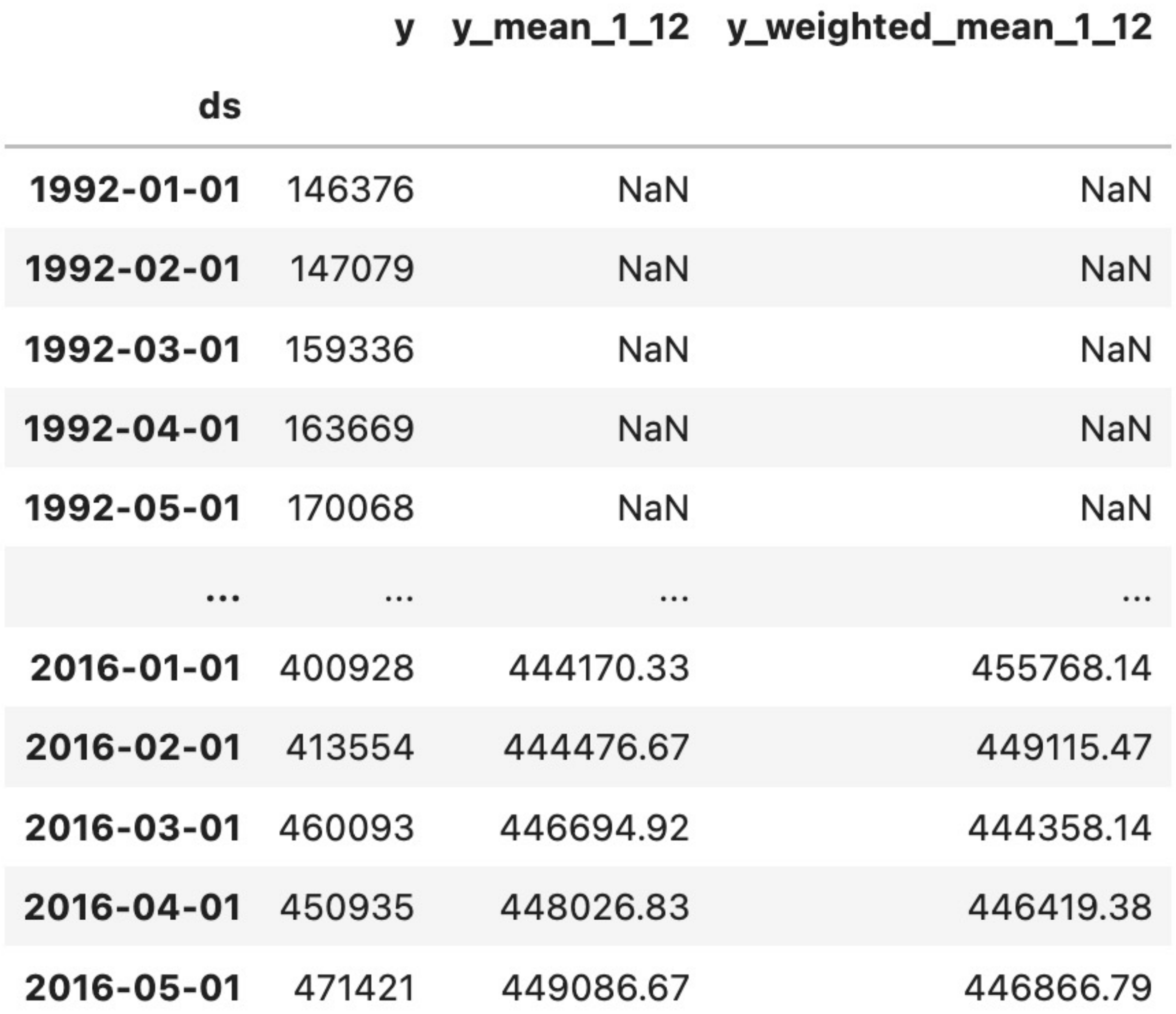
Weighted window in sktime



```
# Define our own weighted mean function with weights defined inside
def weighted_mean(x):
    weights = np.arange(1, 13) # [1, 2, ..., 12]
    return (weights * x).sum() / weights.sum()
```

```
transformer = WindowSummarizer(
    lag_feature={
        "mean": [[1, 12]],  # [[lag, window size]]
        weighted_mean: [[1, 12]],  # Can pass custom functions.
    },
    target_cols=["y"],
)

result = transformer.fit_transform(df)
result
```

Weighted window in sktime



	y	y_mean_1_12	y_weighted_mean_1_12
ds			
1992-01-01	146376	NaN	NaN
1992-02-01	147079	NaN	NaN
1992-03-01	159336	NaN	NaN
1992-04-01	163669	NaN	NaN
1992-05-01	170068	NaN	NaN
...	...	...	...
2016-01-01	400928	444170.33	455768.14
2016-02-01	413554	444476.67	449115.47
2016-03-01	460093	446694.92	444358.14
2016-04-01	450935	448026.83	446419.38
2016-05-01	471421	449086.67	446866.79