In [53]:
```
%%html

<h1>List of chapters in this book:<h1>
<h2>Chapter 11 - Time series</h2>
```

List of chapters in this book:

Chapter 11 - Time series

In [1]:
```
%%html
<h2>Chapter 11 - Time series</h2>
Code examples are taken from <a href="https://github.com/wesm/pydata-book/blob/3rd-edition/ch11.ipynb">https://gith
```

Chapter 11 - Time series

Code examples are taken from https://github.com/wesm/pydata-book/blob/3rd-edition/ch11.ipynb

In [3]:
```
import numpy as np
import pandas as pd
np.random.seed(12345)
import matplotlib.pyplot as plt
plt.rc("figure", figsize=(10, 6))
PREVIOUS_MAX_ROWS = pd.options.display.max_rows
pd.options.display.max_columns = 20
pd.options.display.max_rows = 20
pd.options.display.max_colwidth = 80
np.set_printoptions(precision=4, suppress=True)
```

In [4]:
```
datestrs = ["2011-07-06 12:00:00", "2011-08-06 00:00:00"]
pd.to_datetime(datestrs)
```

Out[4]:
```
DatetimeIndex(['2011-07-06 12:00:00', '2011-08-06 00:00:00'], dtype='datetime64[ns]', freq=None)
```

In [6]:
```
from datetime import datetime
dates = [datetime(2011, 1, 2), datetime(2011, 1, 5),
         datetime(2011, 1, 7), datetime(2011, 1, 8),
         datetime(2011, 1, 10), datetime(2011, 1, 12)]
ts = pd.Series(np.random.standard_normal(6), index=dates)
ts
```

Out[6]:
```
2011-01-02   -0.204708
2011-01-05    0.478943
2011-01-07   -0.519439
2011-01-08   -0.555730
2011-01-10    1.965781
2011-01-12    1.393406
dtype: float64
```

In [9]:
```
ts.index
```

Out[9]:
```
DatetimeIndex(['2011-01-02', '2011-01-05', '2011-01-07', '2011-01-08',
               '2011-01-10', '2011-01-12'],
              dtype='datetime64[ns]', freq=None)
```

In [10]:
```
longer_ts = pd.Series(np.random.standard_normal(1000),
                      index=pd.date_range("2000-01-01", periods=1000))
longer_ts
```

Out[10]:
```
2000-01-01    0.092908
2000-01-02    0.281746
2000-01-03    0.769023
2000-01-04    1.246435
2000-01-05    1.007189
                ...
2002-09-22    0.930944
2002-09-23   -0.811676
2002-09-24   -1.830156
2002-09-25   -0.138730
2002-09-26    0.334088
Freq: D, Length: 1000, dtype: float64
```

In [11]:
```
longer_ts["2002-09"]
```

Out[11]:
```
2002-09-01    0.380339
2002-09-02   -1.067035
2002-09-03    0.255452
2002-09-04    2.111287
2002-09-05   -0.634190
                ...
2002-09-22    0.930944
2002-09-23   -0.811676
2002-09-24   -1.830156
2002-09-25   -0.138730
2002-09-26    0.334088
Freq: D, Length: 26, dtype: float64
```

In [13]:
```
dates = pd.date_range("2000-01-01", periods=100, freq="W-WED")
dates
long_df = pd.DataFrame(np.random.standard_normal((100, 4)),
                       index=dates,
                       columns=["Colorado", "Texas",
                                "New York", "Ohio"])
long_df.loc["2001-05"]
```

Out[13]:

	Colorado	Texas	New York	Ohio
2001-05-02	-0.006045	0.490094	-0.277186	-0.707213
2001-05-09	-0.560107	2.735527	0.927335	1.513906
2001-05-16	0.538600	1.273768	0.667876	-0.969206
2001-05-23	1.676091	-0.817649	0.050188	1.951312
2001-05-30	3.260383	0.963301	1.201206	-1.852001

In [14]:
```
%%html
<h3>Date ranges, Frequencies, and Shifting</h3>
```

Date ranges, Frequencies, and Shifting

In [16]:
```
index = pd.date_range("2012-04-01","2012-06-01")
index
```

Out[16]:
```
DatetimeIndex(['2012-04-01', '2012-04-02', '2012-04-03', '2012-04-04',
               '2012-04-05', '2012-04-06', '2012-04-07', '2012-04-08',
               '2012-04-09', '2012-04-10', '2012-04-11', '2012-04-12',
               '2012-04-13', '2012-04-14', '2012-04-15', '2012-04-16',
               '2012-04-17', '2012-04-18', '2012-04-19', '2012-04-20',
               '2012-04-21', '2012-04-22', '2012-04-23', '2012-04-24',
               '2012-04-25', '2012-04-26', '2012-04-27', '2012-04-28',
               '2012-04-29', '2012-04-30', '2012-05-01', '2012-05-02',
               '2012-05-03', '2012-05-04', '2012-05-05', '2012-05-06',
               '2012-05-07', '2012-05-08', '2012-05-09', '2012-05-10',
               '2012-05-11', '2012-05-12', '2012-05-13', '2012-05-14',
               '2012-05-15', '2012-05-16', '2012-05-17', '2012-05-18',
               '2012-05-19', '2012-05-20', '2012-05-21', '2012-05-22',
               '2012-05-23', '2012-05-24', '2012-05-25', '2012-05-26',
               '2012-05-27', '2012-05-28', '2012-05-29', '2012-05-30',
               '2012-05-31', '2012-06-01'],
              dtype='datetime64[ns]', freq='D')
```

In [17]:
```
pd.date_range(start="2012-04-01", periods=20)
```

Out[17]:
```
DatetimeIndex(['2012-04-01', '2012-04-02', '2012-04-03', '2012-04-04',
               '2012-04-05', '2012-04-06', '2012-04-07', '2012-04-08',
               '2012-04-09', '2012-04-10', '2012-04-11', '2012-04-12',
               '2012-04-13', '2012-04-14', '2012-04-15', '2012-04-16',
               '2012-04-17', '2012-04-18', '2012-04-19', '2012-04-20'],
              dtype='datetime64[ns]', freq='D')
```

In [18]:
```
pd.date_range(end="2012-06-01", periods=20)
```

Out[18]:
```
DatetimeIndex(['2012-05-13', '2012-05-14', '2012-05-15', '2012-05-16',
               '2012-05-17', '2012-05-18', '2012-05-19', '2012-05-20',
               '2012-05-21', '2012-05-22', '2012-05-23', '2012-05-24',
               '2012-05-25', '2012-05-26', '2012-05-27', '2012-05-28',
               '2012-05-29', '2012-05-30', '2012-05-31', '2012-06-01'],
              dtype='datetime64[ns]', freq='D')
```

In [20]:
```
pd.date_range("2000-01-01", "2000-12-01", freq="BM")
```

```
/tmp/ipykernel_76239/3742551278.py:1: FutureWarning: 'BM' is deprecated and will be removed in a future version, ple
ase use 'BME' instead.
  pd.date_range("2000-01-01", "2000-12-01", freq="BM")
```

Out[20]:
```
DatetimeIndex(['2000-01-31', '2000-02-29', '2000-03-31', '2000-04-28',
               '2000-05-31', '2000-06-30', '2000-07-31', '2000-08-31',
               '2000-09-29', '2000-10-31', '2000-11-30'],
              dtype='datetime64[ns]', freq='BME')
```

In [22]:
```
%%html
Base time series frequencies
<img src="images/time_series_freq1.png" width=400>
<img src="images/time_series_freq2.png" width=400>
```

Base time series frequencies

 No description has been provided for this image

 No description has been provided for this image

In [23]:
```
%%html

<h3>Frequencies and Date offsets</h3>
```

Frequencies and Date offsets

In [28]:
```
from pandas.tseries.offsets import Hour, Minute
hour = Hour()
hour
```

Out[28]:
```
<Hour>
```

In [30]:
```
Hour(2) + Minute(30)
```

Out[30]:
```
<150 * Minutes>
```

In [29]:
```
pd.date_range("2000-01-01", "2000-01-03 23:59", freq="4H")
```

```
/tmp/ipykernel_76239/3449897904.py:1: FutureWarning: 'H' is deprecated and will be removed in a future version, plea
se use 'h' instead.
  pd.date_range("2000-01-01", "2000-01-03 23:59", freq="4H")
```

Out[29]:
```
DatetimeIndex(['2000-01-01 00:00:00', '2000-01-01 04:00:00',
               '2000-01-01 08:00:00', '2000-01-01 12:00:00',
               '2000-01-01 16:00:00', '2000-01-01 20:00:00',
               '2000-01-02 00:00:00', '2000-01-02 04:00:00',
               '2000-01-02 08:00:00', '2000-01-02 12:00:00',
               '2000-01-02 16:00:00', '2000-01-02 20:00:00',
               '2000-01-03 00:00:00', '2000-01-03 04:00:00',
               '2000-01-03 08:00:00', '2000-01-03 12:00:00',
               '2000-01-03 16:00:00', '2000-01-03 20:00:00'],
              dtype='datetime64[ns]', freq='4h')
```

In [32]:
```
pd.date_range("2000-01-01", periods=10, freq="1h30min")
```

Out[32]:
```
DatetimeIndex(['2000-01-01 00:00:00', '2000-01-01 01:30:00',
               '2000-01-01 03:00:00', '2000-01-01 04:30:00',
               '2000-01-01 06:00:00', '2000-01-01 07:30:00',
               '2000-01-01 09:00:00', '2000-01-01 10:30:00',
               '2000-01-01 12:00:00', '2000-01-01 13:30:00'],
              dtype='datetime64[ns]', freq='90min')
```

In [33]:
```
monthly_dates = pd.date_range("2012-01-01", "2012-09-01", freq="WOM-3FRI")
list(monthly_dates)
```

Out[33]:
```
[Timestamp('2012-01-20 00:00:00'),
 Timestamp('2012-02-17 00:00:00'),
 Timestamp('2012-03-16 00:00:00'),
 Timestamp('2012-04-20 00:00:00'),
 Timestamp('2012-05-18 00:00:00'),
 Timestamp('2012-06-15 00:00:00'),
 Timestamp('2012-07-20 00:00:00'),
 Timestamp('2012-08-17 00:00:00')]
```

In [35]:
```
%%html

<h3>Shifting (leading and Lagging) Data</h3>
```

Shifting (leading and Lagging) Data

In [36]:
```
ts = pd.Series(np.random.standard_normal(4),
               index=pd.date_range("2000-01-01", periods=4, freq="M"))
ts
```

```
/tmp/ipykernel_76239/3562605038.py:2: FutureWarning: 'M' is deprecated and will be removed in a future version, plea
se use 'ME' instead.
  index=pd.date_range("2000-01-01", periods=4, freq="M"))
```

Out[36]:
```
2000-01-31   -0.066748
2000-02-29    0.838639
2000-03-31   -0.117388
2000-04-30   -0.517795
Freq: ME, dtype: float64
```

In [37]:
```
ts.shift(2)
```

Out[37]:
```
2000-01-31         NaN
2000-02-29         NaN
2000-03-31   -0.066748
2000-04-30    0.838639
Freq: ME, dtype: float64
```

In [38]:
```
ts.shift(-2)
```

Out[38]:
```
2000-01-31   -0.117388
2000-02-29   -0.517795
2000-03-31         NaN
2000-04-30         NaN
Freq: ME, dtype: float64
```

In [39]:
```
ts.shift(2, freq="M")
```

```
/tmp/ipykernel_76239/903147437.py:1: FutureWarning: 'M' is deprecated and will be removed in a future version, pleas
e use 'ME' instead.
  ts.shift(2, freq="M")
```

Out[39]:
```
2000-03-31   -0.066748
2000-04-30    0.838639
2000-05-31   -0.117388
2000-06-30   -0.517795
Freq: ME, dtype: float64
```

In [41]:
```
%%html

<h3>Time Zone Handling</h3>
```

Time Zone Handling

In [42]:
```
dates = pd.date_range("2012-03-09 09:30", periods=6)
ts = pd.Series(np.random.standard_normal(len(dates)), index=dates)
ts
```

Out[42]:
```
2012-03-09 09:30:00   -0.116696
2012-03-10 09:30:00    2.389645
2012-03-11 09:30:00   -0.932454
2012-03-12 09:30:00   -0.229331
2012-03-13 09:30:00   -1.140330
2012-03-14 09:30:00    0.439920
Freq: D, dtype: float64
```

In [44]:
```
print(ts.index.tz)
```

```
None
```

In [47]:
```
pd.date_range("2012-03-09 09:30", periods=10, tz="UTC")
```

Out[47]:
```
DatetimeIndex(['2012-03-09 09:30:00+00:00', '2012-03-10 09:30:00+00:00',
               '2012-03-11 09:30:00+00:00', '2012-03-12 09:30:00+00:00',
               '2012-03-13 09:30:00+00:00', '2012-03-14 09:30:00+00:00',
               '2012-03-15 09:30:00+00:00', '2012-03-16 09:30:00+00:00',
               '2012-03-17 09:30:00+00:00', '2012-03-18 09:30:00+00:00'],
              dtype='datetime64[ns, UTC]', freq='D')
```

In [51]:
```
# Conversion from naive to localized
ts_utc = ts.tz_localize("UTC")
```

```
ts_utc
```

Out[51]:
```
2012-03-09 09:30:00+00:00   -0.116696
2012-03-10 09:30:00+00:00    2.389645
2012-03-11 09:30:00+00:00   -0.932454
2012-03-12 09:30:00+00:00   -0.229331
2012-03-13 09:30:00+00:00   -1.140330
2012-03-14 09:30:00+00:00    0.439920
Freq: D, dtype: float64
```

In [52]:
```
# Convert to another timezone
ts_utc.tz_convert("America/New_York")
```

Out[52]:
```
2012-03-09 04:30:00-05:00   -0.116696
2012-03-10 04:30:00-05:00    2.389645
2012-03-11 05:30:00-04:00   -0.932454
2012-03-12 05:30:00-04:00   -0.229331
2012-03-13 05:30:00-04:00   -1.140330
2012-03-14 05:30:00-04:00    0.439920
Freq: D, dtype: float64
```

In [ ]:
```
```

In [ ]:
```
```