```
Лабораторна робота №5. Варіант 15
ІП-13 Ал Хадам Мурат Резгович
Робота з часовими рядами в Pandas
Мета роботи: Ознайомитись з основними представленнями часових
рядів в Pandas та операціями з ними.
Завдання до лабораторної роботи:
1.Побудувати графік зміни об'ємів продаж: а) загальний; б) за 2006 рік; в) за лютий 2020 року; г) за лютий 2008 – травень 2010;
д) за 2016 та 2018 на одному графіку (паралельно). 2.3 найти середні значення ціни на час відкриття біржі: а) за 2012 рік; б) за
кожний рік; в) за кожний тиждень четвертого кварталу 2016 року. г) Розрахувати і зобразити зміни ціни на час відкриття біржі у
відсотках за кожний день впродовж осені 2011 року. д) Знайти та зобразити графічно ковзне середнє ціни на час відкриття
біржі за 2019 рік з вікном в місяць.
In [1]:
import pandas as pd
import matplotlib.pyplot as plt
plt.rcParams['figure.figsize'] = (10, 6)
df = pd.read_csv('Amazon.csv', parse_dates=['Date'], index_col='Date')
df
Out[1]:
                                                               Adj Close
                                                                          Volume
                  Open
                              High
                                           Low
                                                      Close
       Date
 1997-05-15
               2.437500
                           2.500000
                                       1.927083
                                                               1.958333 72156000
                                                   1.958333
1997-05-16
               1.968750
                           1.979167
                                       1.708333
                                                   1.729167
                                                               1.729167 14700000
1997-05-19
               1.760417
                           1.770833
                                       1.625000
                                                   1.708333
                                                               1.708333
                                                                          6106800
1997-05-20
               1.729167
                           1.750000
                                       1.635417
                                                   1.635417
                                                               1.635417
                                                                          5467200
1997-05-21
                           1.645833
                                       1.375000
                                                   1.427083
                                                               1.427083 18853200
               1.635417
2020-07-27 3062.000000 3098.000000
                                    3015.770020 3055.209961
                                                            3055.209961
                                                                         4170500
2020-07-28 3054.270020 3077.090088
                                    2995.760010 3000.330078
                                                            3000.330078
                                                                         3126700
2020-07-29 3030.989990 3039.159912 2996.770020 3033.530029
                                                            3033.530029
                                                                          2974100
2020-07-30 3014.000000
                        3092.000000
                                    3005.000000 3051.879883
                                                                          6128300
2020-07-31 3244.00000 3246.820068 3151.00000 3164.679932 3164.679932
                                                                          8085500
5842 rows × 6 columns
Завдання 1
Побудувати графік зміни об'ємів продаж: а) загальний;
In [2]:
plt.plot(df['Volume'])
plt.title("Загальна зміна об'ємів продажів з роками")
plt.xlabel('Дата')
plt.ylabel("Об'єм продажів")
plt.show()
                                     Загальна зміна об'ємів продажів з роками
        1e8
    1.0
    0.8
 Об'єм продажів
    0.6
    0.4
    0.2
    0.0
                                                        2008
                      2000
                                       2004
                                                                         2012
                                                                                          2016
                                                                                                           2020
                                                            Дата
б) за 2006 рік;
In [3]:
year_2006 = df.loc['2006-01-01':'2006-12-31']['Volume']
year_2006
Out[3]:
Date
              7582200
2006-01-03
2006-01-04
              7440900
2006-01-05
              5417200
              6152900
2006-01-06
2006-01-09
              8943100
2006-12-22
              5585100
2006-12-26
              4473600
              3536500
2006-12-27
              4545500
2006-12-28
              4197000
2006-12-29
Name: Volume, Length: 251, dtype: int64
In [4]:
plt.plot(year_2006, label='3a 2006')
plt.title("Зміна об'ємів продажів за 2006")
plt.xlabel('Дата')
plt.ylabel("Об'єм продажів")
plt.legend(loc='best')
plt.show()
                                          Зміна об'ємів продажів за 2006
                                                                                                         3a 2006
    7
    6
    5
 Об'єм продажів
    3
    2
    1
    0
                                                                                          2006-11
       2006-01
                       2006-03
                                        2006-05
                                                        2006-07
                                                                         2006-09
                                                                                                           2007-01
                                                          Дата
в) за лютий 2020 року;
In [5]:
february_2020 = df.loc['2020-02-01':'2020-02-28']['Volume']
february_2020
Out[5]:
Date
2020-02-03
              5899100
2020-02-04
              5289300
2020-02-05
              4376200
2020-02-06
              3183000
2020-02-07
              5095300
2020-02-10
              5056200
2020-02-11
              5746000
2020-02-12
              3334300
2020-02-13
              3031800
2020-02-14
              2606200
2020-02-18
              2945600
2020-02-19
              2561200
2020-02-20
              3131300
2020-02-21
              4646300
2020-02-24
              6547000
2020-02-25
              6219100
2020-02-26
              5224600
2020-02-27
              8112100
2020-02-28
              9493800
Name: Volume, dtype: int64
plt.plot(february_2020, label='3a лютий 2020')
plt.title("Зміна об'ємів продажів за лютий 2020")
plt.xlabel('Дата')
plt.ylabel("Об'єм продажів")
plt.legend(loc='best')
plt.show()
                                     Зміна об'ємів продажів за лютий 2020
       1e6
              За лютий 2020
    9
    8
 Об'єм продажів
    5
    4
    3
              2020-02-05
                             2020-02-09
                                             2020-02-13
                                                            2020-02-17
                                                                            2020-02-21
                                                                                            2020-02-25
                                                                                                           2020-02-29
                                                         Дата
г) за лютий 2008 – травень 2010;
In [7]:
feb2008_may2010 = df.loc['2008-02-01':'2010-05-31']['Volume']
feb2008_may2010
Out[7]:
Date
2008-02-01
              16361000
              9155200
2008-02-04
2008-02-05
              9633900
              12399500
2008-02-06
2008-02-07
              14501700
2010-05-24
               4535900
2010-05-25
               6913400
2010-05-26
               6964600
2010-05-27
               4747700
2010-05-28
               3954800
Name: Volume, Length: 586, dtype: int64
plt.plot(feb2008_may2010, label='За лютий 2008 - травень 2010')
plt.title("Зміна об'ємів продажів за лютий 2008 - травень 2010")
plt.xlabel('Дата')
plt.ylabel("Об'єм продажів")
plt.legend(loc='best')
plt.xticks(rotation=45)
plt.show()
                            Зміна об'ємів продажів за лютий 2008 - травень 2010
       1e7
              За лютий 2008 - травень 2010
    5
    4
 Об'єм прода
    3
    2
    1
                                                          Дата
д) за 2016 та 2018 на одному графіку (паралельно).
In [9]:
df_2016 = df.loc['2016-01-01':'2016-12-31']['Volume']
df_2018 = df.loc['2018-01-01':'2018-12-31']['Volume']
df_2018.index = df_2018.index.map(lambda x: x.replace(year=2016))
In [10]:
plt.plot(df_2016, label='2016')
plt.plot(df_2018, label='2018')
plt.title("Зміна об'ємів продажів у 2016 та 2018 роках")
plt.xlabel('Дата')
plt.ylabel("Об'єм продажів")
plt.legend()
plt.gca().xaxis.set_major_formatter(plt.matplotlib.dates.DateFormatter('%m-%d'))
plt.show()
                                   Зміна об'ємів продажів у 2016 та 2018 роках
        1e7
                                                                                                              2016
                                                                                                              2018
    1.4
    1.2
    1.0
 Об'єм продажів
    0.8
    0.6
    0.4
    0.2
          01-01
                                                            07-01
                                                                                             11-01
                           03-01
                                           05-01
                                                                             09-01
                                                                                                              01-01
                                                            Дата
Завдання 2
Знайти середні значення ціни на час відкриття біржі: а) за 2012 рік;
In [11]:
open_2012 = df.loc['2012-01-01':'2012-12-31']['Open']
Out[11]:
Date
2012-01-03 175.889999
2012-01-04 179.210007
2012-01-05 175.940002
2012-01-06 178.070007
2012-01-09 182.759995
               . . .
2012-12-24 257.299988
2012-12-26 257.149994
2012-12-27 248.279999
2012-12-28 245.850006
2012-12-31 243.750000
Name: Open, Length: 250, dtype: float64
In [12]:
print('Середня вартість ціни при відкритті біржі станом на 2012 рік:', round(open_2012.mean(), 2))
Середня вартість ціни при відкритті біржі станом на 2012 рік: 220.01
б) за кожний рік;
In [13]:
avg_per_year = df['Open'].resample('Y').mean()
avg_per_year
Out[13]:
Date
1997-12-31
               3.114486
1998-12-31 15.550967
1999-12-31 69.408720
2000-12-31
             46.602307
2001-12-31
             12.264262
              16.406349
2002-12-31
             37.606905
2003-12-31
2004-12-31 43.557500
2005-12-31 39.820397
2006-12-31 35.872231
2007-12-31 67.114821
2008-12-31 69.839051
2009-12-31
              87.071270
2010-12-31 138.958373
2011-12-31 196.669682
2012-12-31 220.011040
2013-12-31 297.875119
2014-12-31 332.798452
2015-12-31 478.124961
2016-12-31 699.756587
2017-12-31
             968.275619
2018-12-31 1644.072709
2019-12-31 1788.746190
2020-12-31
             2304.611433
Freq: A-DEC, Name: Open, dtype: float64
в) за кожний тиждень четвертого кварталу 2016 року.
In [14]:
avg_week_q4_2016 = df.loc['2016-10-01':'2016-12-31']['Open'].resample('W').mean()
avg_week_q4_2016
Out[14]:
Date
2016-10-09
              840.879993
2016-10-16
              836.470007
2016-10-23
             817.471997
             822.050000
2016-10-30
2016-11-06 778.359998
2016-11-13 767.029993
2016-11-20
             745.142004
             780.599991
2016-11-27
             760.559998
2016-12-04
2016-12-11
             763.081995
             768.178015
2016-12-18
2016-12-25
             766.042004
2017-01-01
             769.630005
Freq: W-SUN, Name: Open, dtype: float64
г) Розрахувати і зобразити зміни ціни на час відкриття біржі у відсотках за кожний день впродовж осені 2011 року.
In [15]:
avg_autumn_2011 = df.loc['2011-09-01':'2011-11-30']['Open']
percentage_change = (avg_autumn_2011 / avg_autumn_2011.shift(1) - 1) * 100
percentage_change
Out[15]:
Date
2011-09-01
2011-09-02
           -2.945000
2011-09-06
           -1.995787
2011-09-07
             6.851589
2011-09-08
            -0.228519
               . . .
2011-11-23
             3.268254
2011-11-25
           -1.372627
2011-11-28 0.651221
2011-11-29
             1.633188
2011-11-30 -0.010270
Name: Open, Length: 63, dtype: float64
In [16]:
plt.plot(percentage_change.index, percentage_change, linestyle='--', color='b')
plt.title('Зміна курсу у відсотковому відношенні за кожний день впродовж осені 2011 року.')
plt.xlabel('Дата')
plt.ylabel('Зміни середніх значення ціни на час відкриття біржі у відсотках')
plt.grid(True)
plt.show()
            Зміна курсу у відсотковому відношенні за кожний день впродовж осені 2011 року.
 середніх значення ціни на час відкриття біржі у відсотках
       5
     -5
    -10
```

-15

In [17]:

Out[17]: Date

2019-01-02 2019-01-03

2019-01-04

2019-01-07 2019-01-08

2019-12-24

2019-12-26

2019-12-27

2019-12-30

2019-12-31

plt.xlabel('Дата') plt.ylabel('Ціна') plt.xticks(rotation=45)

plt.legend() plt.show()

In [18]:

2011-09-01

avg_open_prices_2019

1465.199951

1520.010010

1530.000000 1602.310059

1664.689941 . . .

1793.810059

1801.010010

1882.920044

1874.000000

1842.000000 Name: Open, Length: 252, dtype: float64

2011-09-15

avg_open_prices_2019 = df.loc['2019-01-01':'2019-12-31']['Open']

month_rolling_mean_2019 = avg_open_prices_2019.rolling(window=30).mean()

plt.plot(avg_open_prices_2019, label='Середня ціна на час відкриття за день') plt.plot(month_rolling_mean_2019, label='Ковзне середнє (місячне вікно)')

plt.title('Ковзне середнє за середньою ціною на час відкриття біржі в 2019 рік з вікном в місяць')

2011-10-01

д) Знайти та зобразити графічно ковзне середнє ціни на час відкриття біржі за 2019 рік з вікном в місяць.

2011-10-15

Дата

2011-11-01

2011-11-15

2011-12-01

