Series ma'lumotlar tuzilmasi

461.22-guruh talabasi Eshmatov B

```
import pandas as pd
import numpy as np
from pandas import Series
```

1-mashq O'zbekiston Respublikasi viloyatlaridan iborat bir o'lchamli Series obyektini yarating

```
viloyat=Series(['Samarqand','Toshkent','Navoiy','Xorazm','Sirdaryo','Andijon','Fargona','Surxondaryo','Qashqadaryo','Buxoro','Jizzax','Xo
```

2-mashq. Yuqorida yaratilgan Series obyektining barcha qiymaltarini chiqaring

```
viloyat
```

```
0       Samarqand
1        Toshkent
2          Navoiy
3          Xorazm
4        Sirdaryo
5         Andijon
6         Fargona
7     Surxondaryo
8     Qashqadaryo
9          Buxoro
10         Jizzax
11         Xorazm
dtype: object
```

3-mashq. Yuqoridagi Series obyektining eng birinchi qiymatini chiqaring

```
print(viloyat.head(1));
```

```
0    Samarqand
dtype: object
```

4-mashq. Yuqoridagi Series obyektining eng oxirgi qiymatini chiqaring

```
print(viloyat.tail(1));
```

```
11    Xorazm
dtype: object
```

5-mashq. Yuqoridagi Series obyektining o'rtasidan 5 ta qiymat chiqaring

```
print(viloyat[5:8]);
```

```
5        Andijon
6        Fargona
7    Surxondaryo
dtype: object
```

6-mashq. Quyidagi Series obyektini yarating. Ahamiyat bering. Series obyekti indeksi sifatida viloyatlar nomi ishlatilgan

```
viloyatlar=Series([1899000,1384700,2597000,910500,1200000,1862000,767500,2029000,1200000,2322000,648100,1676000,4450000],
                  index=['Andijon','Buxoro','Fargona','Jizzax','Xorazm','Namangan','Navoiy','Qashqadaryo','Qoraqalpogiston Respublikasi',
print(viloyatlar)
```

```
Andijon                         1899000
Buxoro                          1384700
Fargona                         2597000
Jizzax                           910500
Xorazm                          1200000
Namangan                        1862000
Navoiy                           767500
Qashqadaryo                     2029000
Qoraqalpogiston Respublikasi    1200000
Samarqand                       2322000
Sirdaryo                         648100
Surxondaryo                     1676000
```

```
Toshkent                        4450000
dtype: int64
```

7-mashq. Yuqoridagi Series obyektidan aholi soni 1 milliondan kam va ko'p bo'lgan viloyatlarni alohida chiqaring

```
print('1000000 dan kop\n')
print(viloyatlar[viloyatlar >= 1000000])
print('\n\n')
print('1000000 dan kam\n')
print(viloyatlar[viloyatlar < 1000000])
```

```
1000000 dan kop

Andijon                         1899000
Buxoro                          1384700
Fargona                         2597000
Xorazm                          1200000
Namangan                        1862000
Qashqadaryo                     2029000
Qoraqalpogiston Respublikasi    1200000
Samarqand                       2322000
Surxondaryo                     1676000
Toshkent                        4450000
dtype: int64



1000000 dan kam

Jizzax      910500
Navoiy      767500
Sirdaryo    648100
dtype: int64
```

8-mashq. Quyidagi lug'atda viloyatlar va ularning maydonlari berilgan. Lug'atdan yangi Series obyektini yarating va unga "Hududlar hajmi" deb nom bering

```
viloyatlar_nomi={
    'Andijon':[4200],
    'Buxoro':[39400],
    'Fargona':[6800],
    'Xorazm':[6300],
    'Namangan':[7900],
    'Navoiy':[110800],
    'Qashqadaryo':[28400],
    'Qoraqalpogiston Respublikasi':[160000],
    'Samarqand':[16400],
    'Sirdaryo':[5100],
    'Surxondaryo':[20800],
    'Jizzax':[20500],
    'Toshkent':[15300],
}
chiqar=Series(viloyatlar_nomi)
chiqar.name='Hududlar hajmi'
print(chiqar.name)
chiqar
```

```
Hududlar hajmi
Andijon                          [4200]
Buxoro                          [39400]
Fargona                          [6800]
Xorazm                           [6300]
Namangan                         [7900]
Navoiy                         [110800]
Qashqadaryo                     [28400]
Qoraqalpogiston Respublikasi    [160000]
Samarqand                       [16400]
Sirdaryo                         [5100]
Surxondaryo                     [20800]
Jizzax                          [20500]
Toshkent                        [15300]
Name: Hududlar hajmi, dtype: object
```

9-mashq. Yuqoridagi Series obyektidan hududi eng katta va eng kichik elementlarni chiqaring

```
np.max(chiqar.values)
```

```
[160000]
```

10-mashq. Yuqoridagi hudulardan O'zbekiston Respublikasi hududini hisoblang

```
print(np.sum(chiqar.sum()))
```

```
441900
```

11-mashq. Yuqoridagi hudulardan O'zbekiston Respublikasi hududini hisoblang va eng katta qiymati, eng kichik qiymatiga nisbatini oluvchi dastur kodini yozing

```
print("Hududi \n",np.sum(chiqar.sum()))
print('\n\n katta kichik nisbati\n')
print((np.max(chiqar.sum())) / (np.min(chiqar.sum())))
```

```
Hududi
 441900


 katta kichik nisbati

38.095238095238095
```

12-mashq. Yuqoridagi hudulardan O'zbekiston Respublikasi hududini hisoblang va eng katta qiymati, eng kichik qiymati yig'indisini oluvchi dastur kodini yozing

```
print("Hududi \n",np.sum(chiqar.sum()))
print('\n\n katta kichik yigindisi')
print((np.max(chiqar.sum())) + (np.min(chiqar.sum())))
```

```
Hududi
 441900


 katta kichik yigindisi
164200
```

13-mashq. Yuqoridagi hudulardan O'zbekiston Respublikasi hududini eng katta qiymati, eng kichik qiymati ko'paytmasi va eng kichik qiymatni chiqaruvchi dastur kodini yozing

```
print("Eng katta qiymat \n", np.max(chiqar.sum()))
print("Eng kichik qiymat \n", np.min(chiqar.sum()))
print("Kopaytmasi \n", (np.max(chiqar.sum()))*(np.min(chiqar.sum())))
```

```
Eng katta qiymat
 160000
Eng kichik qiymat
 4200
Kopaytmasi
 672000000
```

14-mashq. Series obyektini yarating va isnull(), notnull metodlarini ishlating va vazifasini tushuntiring.

```
chiqarish = pd.Series(['Olma', 'Anor', 'Nok', 'Behi', 'Gilos', 'Uzum', 'Banan', 'Shaftoli'], index = ['a', 's', 'd', 'f', 'g', 'h', 'j',
print(chiqarish.isnull())
print('\n')
print(chiqarish.notnull())
```

```
a    False
s    False
d    False
f    False
g    False
h    False
j    False
k    False
dtype: bool


a    True
s    True
d    True
f    True
g    True
h    True
j    True
k    True
dtype: bool
```

15-mashq. Series obyektini yarating va iloc, loc, max, min, mean metodlarini ishlating va vazifasini tushuntiring.

```
print(chiqarish.iloc[2]) # iloc() metodi elementga faqat index orqali murojaatni ta'minlaydi
print(chiqarish.iloc[4:6])

print('\n')
print(chiqarish.loc['s']) #loc() metodi elementga index nomi yoki bog'lanish uslubi orqali murojaat qilishni ta'minlaydi
print(chiqarish.loc['g':'k'])
```

```
Nok
g    Gilos
h     Uzum
dtype: object


Anor
g       Gilos
h        Uzum
j       Banan
k    Shaftoli
dtype: object
```

DataFrame ma'lumot tuzilmasi bilan ishlash

1-mashq. Quyida ko'rsatilgan dataframeni yarating va df deb nomlang

```
vil1 = viloyatlar.index
vil2 = []
for ab in vil1:
  if(len(ab) < 20):
    vil2.append(ab + ' viloyati')
  else :
    vil2.append(ab)
viloyatlar
```

```
Andijon                         1899000
Buxoro                          1384700
Fargona                         2597000
Jizzax                           910500
Xorazm                          1200000
Namangan                        1862000
Navoiy                           767500
Qashqadaryo                     2029000
Qoraqalpogiston Respublikasi    1200000
Samarqand                       2322000
Sirdaryo                         648100
Surxondaryo                     1676000
Toshkent                        4450000
dtype: int64
```

```
data = {
    'hudud': vil2,
    'markazi':['Andijon shahri', 'Buxoro shahri', 'Fargona shahri', 'Jizzax shahri', 'Urganch shahri', 'Namangan shahri', 'Navoiy shahri
    'maydoni': chiqar.sum(),
    'aholisi': viloyatlar.values
}
df = pd.DataFrame(data)
df
```

	hudud	markazi	maydoni	aholisi
0	Andijon viloyati	Andijon shahri	4200	1899000
1	Buxoro viloyati	Buxoro shahri	39400	1384700
2	Fargona viloyati	Fargona shahri	6800	2597000
3	Jizzax viloyati	Jizzax shahri	6300	910500
4	Xorazm viloyati	Urganch shahri	7900	1200000
5	Namangan viloyati	Namangan shahri	110800	1862000
6	Navoiy viloyati	Navoiy shahri	28400	767500
7	Qashqadaryo viloyati	Qarshi shahri	160000	2029000
8	Qoraqalpogiston Respublikasi	Nukus shahri	16400	1200000
9	Samarqand viloyati	Samarqand shahri	5100	2322000
10	Sirdaryo viloyati	Guliston shahri	20800	648100
11	Surxondaryo viloyati	Termiz shahri	20500	1676000
12	Toshkent viloyati	Nurafshon shahri	15300	4450000

Next steps: Generate code with df View recommended plots

2-mashq. Yuqoridagi df ning yuqorisidan 7 ta qator chiqaring

```
df.head(7)
```

	hudud	markazi	maydoni	aholisi
0	Andijon viloyati	Andijon shahri	4200	1899000
1	Buxoro viloyati	Buxoro shahri	39400	1384700
2	Fargona viloyati	Fargona shahri	6800	2597000
3	Jizzax viloyati	Jizzax shahri	6300	910500
4	Xorazm viloyati	Urganch shahri	7900	1200000
5	Namangan viloyati	Namangan shahri	110800	1862000
6	Navoiy viloyati	Navoiy shahri	28400	767500

Next steps: Generate code with df View recommended plots

3-mashq. Yuqoridagi df dan tasodifiy 5 ta qator chiqaring

```
df.sample(n=5)
```

	hudud	markazi	maydoni	aholisi
9	Samarqand viloyati	Samarqand shahri	5100	2322000
11	Surxondaryo viloyati	Termiz shahri	20500	1676000
3	Jizzax viloyati	Jizzax shahri	6300	910500
7	Qashqadaryo viloyati	Qarshi shahri	160000	2029000
5	Namangan viloyati	Namangan shahri	110800	1862000

4-mashq. Yuqoridagi df ning 'Hudud' va 'Aholisi' ustinlarini alohida df ko'rinishida chiqaring

```
hudud = df['hudud']
aholi = df['aholisi']
print(hudud, "\n")
print(aholi)
```

```
0                Andijon viloyati
1                 Buxoro viloyati
2                Fargona viloyati
3                 Jizzax viloyati
4                 Xorazm viloyati
5               Namangan viloyati
6                 Navoiy viloyati
7            Qashqadaryo viloyati
8    Qoraqalpogiston Respublikasi
```

```
9              Samarqand viloyati
10              Sirdaryo viloyati
11           Surxondaryo viloyati
12              Toshkent viloyati
Name: hudud, dtype: object

0     1899000
1     1384700
2     2597000
3      910500
4     1200000
5     1862000
6      767500
7     2029000
8     1200000
9     2322000
10     648100
11    1676000
12    4450000
Name: aholisi, dtype: int64
```

5-mashq. Yuqoridagi df ning Markaz ustunini Series obyekti ko'rinishida chiqaring

```
chiqarish = Series(df['markazi'])
print(chiqarish)
```

```
0       Andijon shahri
1        Buxoro shahri
2       Fargona shahri
3        Jizzax shahri
4       Urganch shahri
5      Namangan shahri
6        Navoiy shahri
7        Qarshi shahri
8         Nukus shahri
9     Samarqand shahri
10     Guliston shahri
11       Termiz shahri
12    Nurafshon shahri
Name: markazi, dtype: object
```

6-mashq. Yuqoridagi df dan aholi soni 1.5 milliondan oshiq bo'lgan hududlarni chiqaring

```
print(df[df.aholisi > 1500000])
```

```
                     hudud           markazi  maydoni  aholisi
0       Andijon viloyati    Andijon shahri     4200  1899000
2       Fargona viloyati    Fargona shahri     6800  2597000
5      Namangan viloyati   Namangan shahri   110800  1862000
7   Qashqadaryo viloyati     Qarshi shahri   160000  2029000
9     Samarqand viloyati  Samarqand shahri     5100  2322000
11  Surxondaryo viloyati     Termiz shahri    20500  1676000
12     Toshkent viloyati  Nurafshon shahri    15300  4450000
```

7-mashq. Yuqoridagi df dan maydoni 20000kv.km dan kichik bo'lgan hududlarni chiqaring

```
print(df[df.maydoni < 20000])
```

```
                            hudud           markazi  maydoni  aholisi
0              Andijon viloyati    Andijon shahri     4200  1899000
2              Fargona viloyati    Fargona shahri     6800  2597000
3               Jizzax viloyati     Jizzax shahri     6300   910500
4               Xorazm viloyati    Urganch shahri     7900  1200000
8   Qoraqalpogiston Respublikasi      Nukus shahri    16400  1200000
9            Samarqand viloyati  Samarqand shahri     5100  2322000
12            Toshkent viloyati  Nurafshon shahri    15300  4450000
```

8-mashq. Yuqoridagi df dan, Aholisi 2 milliondan oshiq hududlar orasida Samarqand viloyati bor yoki yo'qligini tekshiring

```
samvil = df.hudud[df.aholisi > 2000000]
if 'Samarqand viloyati' in samvil.values :
  print('Mavjud')
else :
  print('Mavjud emas')
```

```
Mavjud
```

9-mashq. Yuqoridagi df dan faqatgina Andijon, Farg'ona va Namangan viloyatlarini ajratib ko'rsating

```
print(df[df.hudud == 'Andijon viloyati'])
print(df[df.hudud=='Fargona viloyati'])
print(df[df.hudud=='Namangan viloyati'])
```

```
              hudud         markazi  maydoni  aholisi
0  Andijon viloyati  Andijon shahri     4200  1899000
              hudud         markazi  maydoni  aholisi
2  Fargona viloyati  Fargona shahri     6800  2597000
              hudud          markazi  maydoni  aholisi
5  Namangan viloyati  Namangan shahri   110800  1862000
```

10-mashq. Yuqoridagi df dan Samarqanda va Toshkent viloyatlarining aholisi va maydonini alohida df ko'rinishida chiqaring loc orqali

```
samtosh = df.loc[[9,12]]
samtosh
```

	hudud	markazi	maydoni	aholisi
9	Samarqand viloyati	Samarqand shahri	5100	2322000
12	Toshkent viloyati	Nurafshon shahri	15300	4450000

Next steps: Generate code with samtosh View recommended plots

11-mashq. Yuqoridagi df dan yangi df_new DataFrame yarating. Yangi dataframega faqatgina aholisi 1.5mln dan kam bo'lgan hududlarni ko'chiring.

```
df_new = df[df.aholisi < 1500000]
df_new
```

	hudud	markazi	maydoni	aholisi
1	Buxoro viloyati	Buxoro shahri	39400	1384700
3	Jizzax viloyati	Jizzax shahri	6300	910500
4	Xorazm viloyati	Urganch shahri	7900	1200000
6	Navoiy viloyati	Navoiy shahri	28400	767500
8	Qoraqalpogiston Respublikasi	Nukus shahri	16400	1200000
10	Sirdaryo viloyati	Guliston shahri	20800	648100

Next steps: Generate code with df_new View recommended plots

12-mashq. Aholisi 1milliondan ko'p, lekin maydoni 10000kv.km dan kam hududlarni chiqaring

```
print(df[(df[df.aholisi > 100000]).maydoni <  10000])
```

```
                hudud           markazi  maydoni  aholisi
0    Andijon viloyati    Andijon shahri     4200  1899000
2    Fargona viloyati    Fargona shahri     6800  2597000
3     Jizzax viloyati     Jizzax shahri     6300   910500
4     Xorazm viloyati    Urganch shahri     7900  1200000
9  Samarqand viloyati  Samarqand shahri     5100  2322000
```

13-mashq. Aholisi 1milliondan kam, lekin maydoni 10000kv.km dan ko'p hududlarni chiqaring. IKKITA SHARTNI BIRLASHTIRISH UCHUN MANTIQIY VA AMALIDAN FOYDALANASIZ "&" SHU BELGI

```
for i in range(13):
  ahomay = df.loc[i]
  if ahomay.maydoni >10000 & ahomay.aholisi < 1000000 :
    print(ahomay.hudud)
```

```
 Buxoro viloyati
 Xorazm viloyati
 Namangan viloyati
 Navoiy viloyati
 Qashqadaryo viloyati
 Qoraqalpogiston Respublikasi
 Sirdaryo viloyati
 Surxondaryo viloyati
 Toshkent viloyati
```

14-mashq. Yuqoridagi df dan Markaz ustunini tashlab yuboring

```
markaz = df.drop('markazi', axis=1)
markaz
```

	hudud	maydoni	aholisi
0	Andijon viloyati	4200	1899000
1	Buxoro viloyati	39400	1384700
2	Fargona viloyati	6800	2597000
3	Jizzax viloyati	6300	910500
4	Xorazm viloyati	7900	1200000
5	Namangan viloyati	110800	1862000
6	Navoiy viloyati	28400	767500
7	Qashqadaryo viloyati	160000	2029000
8	Qoraqalpogiston Respublikasi	16400	1200000
9	Samarqand viloyati	5100	2322000
10	Sirdaryo viloyati	20800	648100
11	Surxondaryo viloyati	20500	1676000
12	Toshkent viloyati	15300	4450000

Next steps: Generate code with `markaz` View recommended plots

15-mashq. Yuqoridagi df dan Sirdaryo viloyati aholisini chiqaring. at orqali

```
print(df.at[10, 'aholisi'])
```

```
648100
```

16-mashq. "cars.xlsx" formatdagi dataset berilgan. Uning ustunlari yig'indisini chiqaruvchi dastur kodini yozing va natijasini oling.

```
dataset = pd.read_excel('cars.xlsx')
dataset.sum()
```

```
Make           AcuraAcuraAcuraAcuraAcuraAcuraAcuraAudiAudiAud...
Model           MDX RSX Type S 2dr TSX 4dr TL 4dr 3.5 RL 4dr ...
Type           SUVSedanSedanSedanSedanSedanSportsSedanSedanSe...
Origin         AsiaAsiaAsiaAsiaAsiaAsiaAsiaEuropeEuropeEurope...
DriveTrain     AllFrontFrontFrontFrontFrontRearFrontFrontFron...
MSRP                                                    14027638
Invoice                                                 12846292
EngineSize                                                1368.2
Cylinders                                                 2474.0
Horsepower                                                 92399
MPG_City                                                   8586
MPG_Highway                                               11489
Weight                                                  1531364
Wheelbase                                                 46290
Length                                                    79763
dtype: object
```

17-mashq. "cars.xlsx" formatdagi dataset berilgan. Uning "price" ustunining eng kichik elemnti idisini va eng katta elementi idisini chiqaruvchi dastur kodini yozing va natijasini oling.

```
print(dataset[['Invoice']].idxmin())
print("\n")
print(dataset[['Invoice']].idxmax())
```

```
Invoice    206
dtype: int64


Invoice    334
dtype: int64
```

18-mashq. "cars.xlsx" formatdagi dataset berilgan. Datasetda describe() va info() metodlarini ishlating va natijasini oling. Bu metodlarni vazifasini tushuntiring.

```
dataset.describe()
```

	MSRP	Invoice	EngineSize	Cylinders	Horsepower	MPG_City	MP
count	428.000000	428.000000	428.000000	426.000000	428.000000	428.000000	
mean	32774.855140	30014.700935	3.196729	5.807512	215.885514	20.060748	
std	19431.716674	17642.117750	1.108595	1.558443	71.836032	5.238218	
min	10280.000000	9875.000000	1.300000	3.000000	73.000000	10.000000	
25%	20334.250000	18866.000000	2.375000	4.000000	165.000000	17.000000	
50%	27635.000000	25294.500000	3.000000	6.000000	210.000000	19.000000	
75%	39205.000000	35710.250000	3.900000	6.000000	255.000000	21.250000	
max	192465.000000	173560.000000	8.300000	12.000000	500.000000	60.000000	

```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 428 entries, 0 to 427
Data columns (total 15 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   Make         428 non-null    object 
 1   Model        428 non-null    object 
 2   Type         428 non-null    object 
 3   Origin       428 non-null    object 
 4   DriveTrain   428 non-null    object 
 5   MSRP         428 non-null    int64  
 6   Invoice      428 non-null    int64  
 7   EngineSize   428 non-null    float64
 8   Cylinders    426 non-null    float64
 9   Horsepower   428 non-null    int64  
 10  MPG_City     428 non-null    int64  
 11  MPG_Highway  428 non-null    int64  
 12  Weight       428 non-null    int64  
 13  Wheelbase    428 non-null    int64  
 14  Length       428 non-null    int64  
dtypes: float64(2), int64(8), object(5)
memory usage: 50.3+ KB
```

19-mashq. "cars.xlsx" formatdagi dataset berilgan. Datasetda corr() va corwith() metodlarini ishlating va natijasini oling. Bu metodlarni vazifasini tushuntiring.

```
print(dataset['Weight'].corr(dataset['Length']))
```

```
0.690020710909717
```

```
dataset1 = dataset[['Invoice', 'Length', 'Weight']]
dataset2 = dataset[['Weight']]
dataset1.corrwith(dataset2['Weight'])
```

```
Invoice    0.442332
Length     0.690021
Weight     1.000000
dtype: float64
```

20-mashq. "cars.xlsx" formatdagi dataset berilgan. Datasetda 3 ta ustunning bir-biriga korelyatsiyasi natijasini oling.

```
print(dataset[['Invoice', 'Weight', 'Length']].corr())
```

```
             Invoice    Weight    Length
Invoice  1.000000  0.442332  0.166586
```