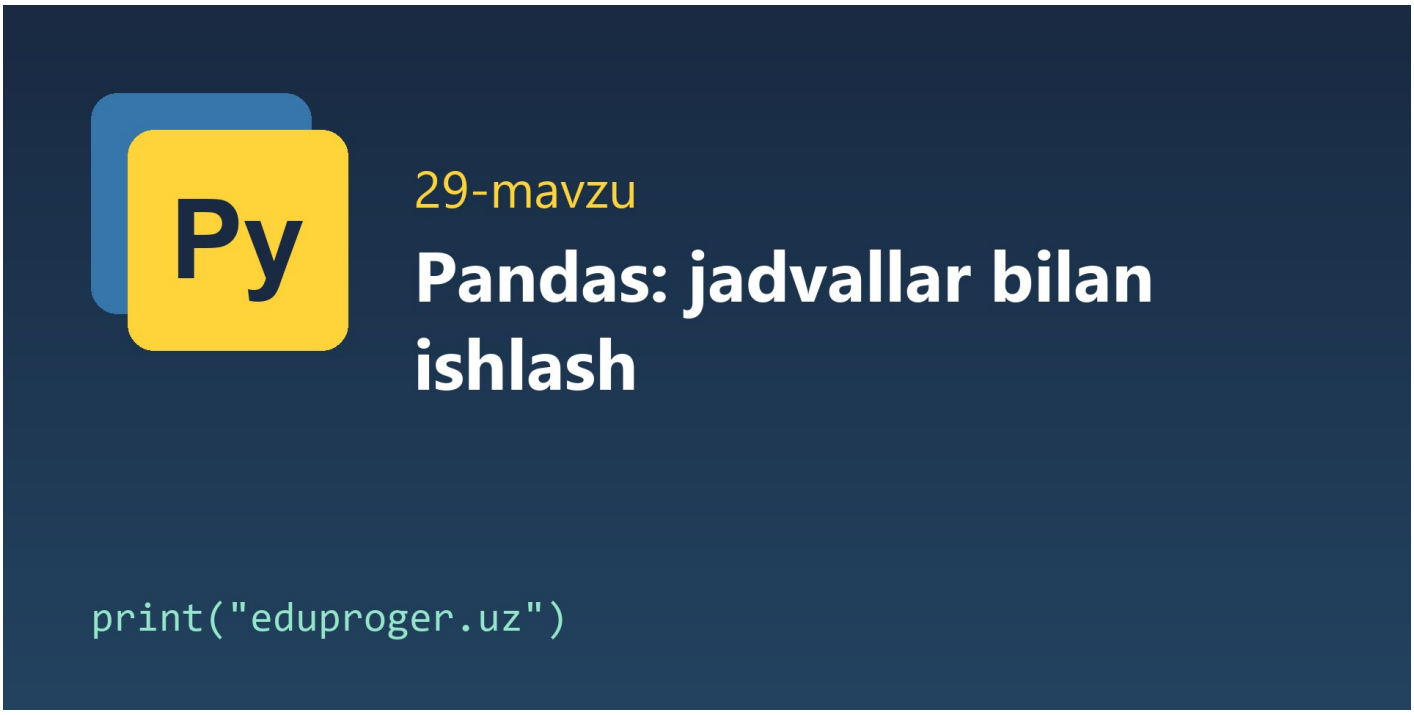
29. Pandas: jadvallar bilan ishlash



Pandas — Excel jadvallarini Python ichida boshqarishga o'xshaydi, faqat ancha kuchliroq: millionlab qator, avtomatlashtirish, murakkab hisobotlar. Ma'lumotlar tahlilchisi ishining asosiy qismi aynan Pandas bilan bajariladi. U NumPy ustiga qurilgan.

> **O'rnatish**
>
> O'rnatish: `pip install pandas`. Saytdagi muharrir uni birinchi `import pandas` da yuklaydi (10–20 soniya, keyin tez).

1. Series va DataFrame

Series — nomlangan bitta ustun, **DataFrame** — ustunlari nomlangan jadval:

```
import pandas as pd
ballar = pd.Series([78, 95, 88], index=["Ali", "Malika", "Sardor"])
print(ballar)
print(ballar["Malika"], ballar.mean())
```

Natija:

```
Ali       78
Malika    95
Sardor    88
dtype: int64
95 87.0
```

```
import pandas as pd
df = pd.DataFrame({
    "ism": ["Ali", "Malika", "Sardor", "Zarina", "Bekzod", "Laylo"],
    "sinf": [9, 9, 10, 10, 9, 10],
    "fan": ["mat", "inf", "mat", "inf", "inf", "mat"],
    "ball": [78, 95, 88, 64, 52, 91],
})
print(df)
print(df.shape)
```

Natija:

```
      ism  sinf  fan  ball
0     Ali     9  mat    78
1  Malika     9  inf    95
2  Sardor    10  mat    88
3  Zarina    10  inf    64
4  Bekzod     9  inf    52
5   Laylo    10  mat    91
(6, 4)
```

2. CSV dan o'qish va birinchi qarash

```
import pandas as pd
with open("talabalar.csv", "w", encoding="utf-8") as f:
    f.write("ism,sinf,fan,ball\nAli,9,mat,78\nMalika,9,inf,95\nSardor,10,mat,8

df = pd.read_csv("talabalar.csv")
print(df.head(3))
print(df.dtypes)
print(df["ball"].describe().round(1))
```

Natija:

```
      ism  sinf  fan  ball
0     Ali     9  mat    78
1  Malika     9  inf    95
2  Sardor    10  mat    88
ism       str
sinf    int64
fan       str
ball    int64
dtype: object
count     6.0
mean     78.0
std      16.9
min      52.0
25%      67.5
50%      83.0
75%      90.2
max      95.0
Name: ball, dtype: float64
```

Metod	Nima beradi
`df.head(n)` / `df.tail(n)`	Birinchi / oxirgi n qator
`df.shape`	(qatorlar, ustunlar)

Metod	Nima beradi
`df.info()`	Ustunlar, turlar, bo'sh qiymatlar
`df.describe()`	Sonli ustunlar statistikasi
`df.columns`	Ustun nomlari

3. Ustun tanlash va filtrlash

```
import pandas as pd
df = pd.DataFrame({"ism": ["Ali", "Malika", "Sardor", "Zarina", "Bekzod", "Lay
                   "fan": ["mat", "inf", "mat", "inf", "inf", "mat"], "ball":
print(df["ism"].tolist())
print(df[["ism", "ball"]].head(2))
print(df[df["ball"] >= 80])
print(df[(df["sinf"] == 10) & (df["fan"] == "mat")]["ism"].tolist())
```

Natija:

```
['Ali', 'Malika', 'Sardor', 'Zarina', 'Bekzod', 'Laylo']
      ism  ball
0     Ali    78
1  Malika    95
      ism  sinf  fan  ball
1  Malika     9  inf    95
2  Sardor    10  mat    88
5   Laylo    10  mat    91
['Sardor', 'Laylo']
```

4. Saralash va yangi ustun

```
import pandas as pd
df = pd.DataFrame({"ism": ["Ali", "Malika", "Sardor", "Zarina"], "ball": [78,
df["baho"] = df["ball"].apply(lambda b: 5 if b >= 86 else 4 if b >= 71 else 3
df["foiz"] = (df["ball"] / 100 * 100).round(1)
print(df.sort_values("ball", ascending=False))
```

Natija:

```
      ism  ball  baho  foiz
1  Malika    95     5  95.0
2  Sardor    88     5  88.0
0     Ali    78     4  78.0
3  Zarina    64     3  64.0
```

5. groupby — guruhlab hisoblash

Excel'dagi «svodnaya jadval» (pivot) ning Pandas'dagi ko'rinishi — `groupby`:

```
import pandas as pd
df = pd.DataFrame({"sinf": [9, 9, 10, 10, 9, 10], "fan": ["mat", "inf", "mat",
print(df.groupby("sinf")["ball"].mean().round(1))
print(df.groupby("fan")["ball"].agg(["count", "mean", "max"]).round(1))
print(df.pivot_table(values="ball", index="sinf", columns="fan", aggfunc="mean
```

Natija:

```
sinf
9     75.0
10    81.0
Name: ball, dtype: float64
     count  mean  max
fan
inf      3  70.3   95
mat      3  85.7   91
fan    inf   mat
sinf
9     73.5  78.0
10    64.0  89.5
```

6. Bo'sh qiymatlar

```
import pandas as pd
import numpy as np
df = pd.DataFrame({"ism": ["Ali", "Vali", "Gani"], "ball": [78, np.nan, 90]})
print(df["ball"].isna().sum(), "ta bo'sh")
print(df.dropna())
print(df.fillna(0))
print(df["ball"].mean())
```

Natija:

```
1 ta bo'sh
    ism  ball
0   Ali  78.0
2  Gani  90.0
    ism  ball
0   Ali  78.0
1  Vali   0.0
2  Gani  90.0
84.0
```

Pandas `NaN` (bo'sh) qiymatlarni o'rtacha va yig'indida avtomatik e'tiborsiz qoldiradi.

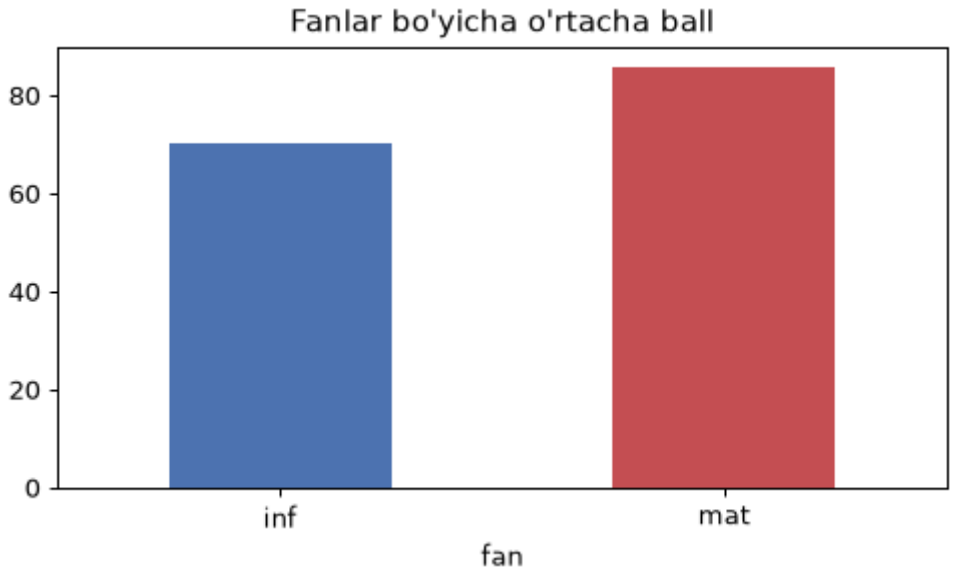
7. Natijani saqlash va grafik

```
import matplotlib
matplotlib.use("Agg")
import pandas as pd

df = pd.DataFrame({"sinf": [9, 9, 10, 10, 9, 10], "fan": ["mat", "inf", "mat",
ortacha = df.groupby("fan")["ball"].mean().round(1)
print(ortacha)
ax = ortacha.plot(kind="bar", color=["#4C72B0", "#C44E52"], figsize=(5.5, 3.4)
ax.figure.tight_layout()
ax.figure.savefig("pandas.png", dpi=90)
df.to_csv("hisobot.csv", index=False)
```

Natija:

```
fan
inf    70.3
mat    85.7
Name: ball, dtype: float64
```



8. Xulosa

- `import pandas as pd`; `pd.read_csv`, `df.head`, `df.describe`, `df.shape`.
- Ustun — `df["ball"]`, filtr — `df[df["ball"] > 80]`, bir nechta shart — `&`, `|` va qavslar.
- `sort_values`, `apply`, yangi ustun `df["x"] = ...`.
- `groupby(...).agg(...)`, `pivot_table`; bo'sh qiymatlar — `isna`, `dropna`, `fillna`; saqlash — `to_csv`.