

10. час: Утврђивање градива

Сада ћемо мало застати и осврнути се на материјал са последња два часа.

Задатак 1. Изврши ћелију испод, па одговори на следећа питања:

```
In [2]: # изврши ову ћелију
import pandas as pd
import matplotlib.pyplot as plt

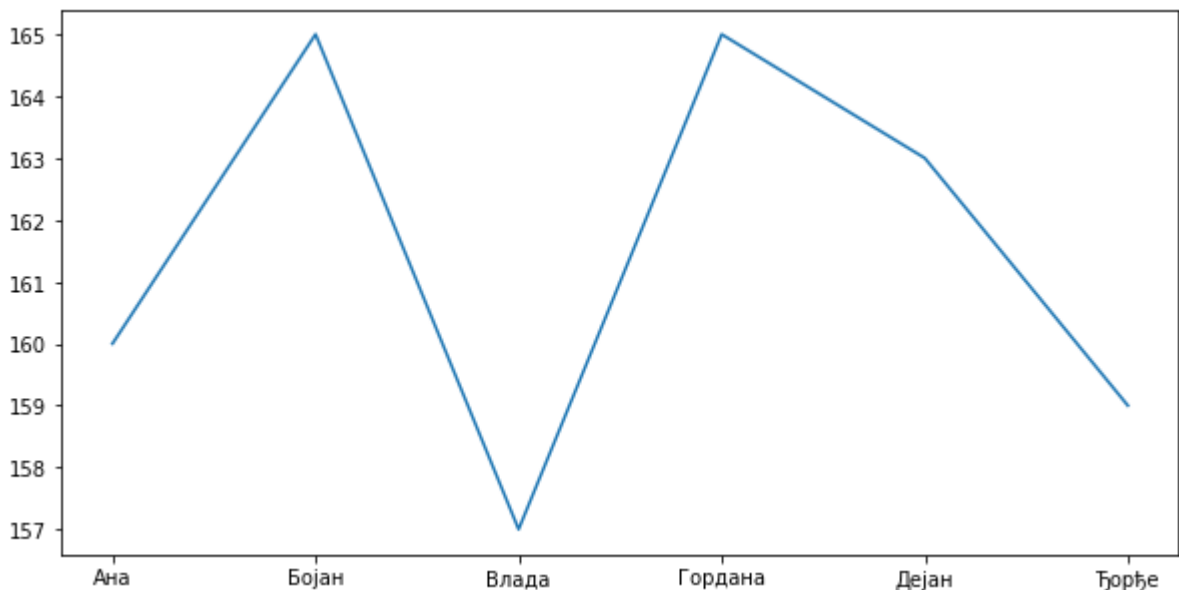
podaci = [
    ["Ана", "ж", 13, 46, 160],
    ["Бојан", "м", 14, 52, 165],
    ["Влада", "м", 13, 47, 157],
    ["Гордана", "ж", 15, 54, 165],
    ["Дејан", "м", 15, 56, 163],
    ["Ђорђе", "м", 13, 45, 159]]

tabela = pd.DataFrame(podaci)
tabela.columns=["Име", "Пол", "Старост", "Тежина", "Висина"]
tabela1=tabela.set_index("Име")

print(tabela1["Висина"].min(), tabela1["Висина"].mean(), tabela1["Висина"].max())

plt.figure(figsize=(10,5))
plt.plot(tabela1.index, tabela1["Висина"])
plt.show()
plt.close()
```

157 161.5 165



1. Зашто учитавамо две библиотеке?
2. Зашто не радимо са подацима у облику листе, већ компликујемо себи живот структуром података *DataFrame* из библиотеке *pandas*?

3. Како се зову колоне табеле `tabela` ?
4. У чему је разлика између табела `tabela` и `tabela1` ?
5. У чему је разлика између функција `min` и `mean` ?
6. Шта представља израз `tabela1.index` ?
7. Како треба изменити програм па да на дијаграму поред висине добијемо и графички приказ тежине деце?
8. Шта је вредност израза `tabela1.loc["Ђорђе"]` ?
9. Шта је вредност израза `tabela1.loc["Ђорђе", "Висина"]` ?
10. Шта је вредност израза `tabela.loc["Ђорђе", "Висина"]` ?

Задатак 2. Ево трошкова живота једне породице током једне године, по месецима (сви износи су представљени у динарима):

Ставка	Јан	Феб	Мар	Апр	Мај	Јун	Јул	Авг	Сеп	Окт	Нов	
Станарина	8.251	8.436	8.524	8.388	8.241	8.196	8.004	7.996	7.991	8.015	8.353	8
Струја	4.321	4.530	4.115	3.990	3.985	3.726	3.351	3.289	3.295	3.485	3.826	3
Телефон (фиксни)	1.425	1.538	1.623	1.489	1.521	1.485	1.491	1.399	1.467	1.531	1.410	1
Телефон (мобилни)	2.181	2.235	2.073	1.951	1.989	1.945	3.017	2.638	2.171	1.831	1.926	1
ТВ и интернет	2.399	2.399	2.399	2.399	2.399	2.399	2.399	2.399	2.399	2.399	2.399	2
Превоз	1.830	1.830	1.830	1.830	1.950	1.950	1.450	1.450	1.950	1.950	2.050	2
Храна	23.250	23.780	24.019	24.117	24.389	24.571	24.736	24.951	25.111	25.389	25.531	25
Остало	4.500	3.700	5.100	3.500	2.750	4.250	7.320	8.250	3.270	4.290	3.200	8

У ћелији испод су исти подаци представљени листом:



```
In [1]: # изврши ову ћелију
troskovi = [
    ["Станарина", 8251, 8436, 8524, 8388, 8241, 8196, 8004, 7996, 7991, 8015, 8353,
    ["Струја", 4321, 4530, 4115, 3990, 3985, 3726, 3351, 3289, 3295, 3485, 3826, 38
    ["Телефон (фиксни)", 1425, 1538, 1623, 1489, 1521, 1485, 1491, 1399, 1467, 1531
    ["Телефон (мобилни)", 2181, 2235, 2073, 1951, 1989, 1945, 3017, 2638, 2171, 183
    ["ТВ и интернет", 2399, 2399, 2399, 2399, 2399, 2399, 2399, 2399, 2399, 2399, 2
    ["Превоз", 1830, 1830, 1830, 1830, 1950, 1950, 1450, 1450, 1950, 1950, 2050, 20
    ["Храна", 23250, 23780, 24019, 24117, 24389, 24571, 24736, 24951, 25111, 25389,
    ["Остало", 4500, 3700, 5100, 3500, 2750, 4250, 7320, 8250, 3270, 4290, 3200, 83
]
```

(а) Додај нову ћелију испод ове па представи табелу структуром *DataFrame*. Индексирај табелу.

(б) Израчунај и испиши *укупне* трошкове ове породице по месецима (колико је породица укупно потрошила у јануарну, колико у фебруару итд).

(в) Израчунај и испиши *просечну* потрошњу ове породице по ставкама (колико је породица просечно потрошила на станарину, колико на струју, итд).

Задатак 3. У табели испод су дати подаци о продаји неких производа у пет пословних јединица једне компаније (бројеви представљају број продатих комада у једном месецу):

Производ	ПЈ1	ПЈ2	ПЈ3	ПЈ4	ПЈ5
Ципеле	5	17	3	11	9
Кошуља	8	6	7	4	0
Каиш	4	1	3	5	1
Панталоне	4	2	6	4	5
Чарапе (пар)	8	9	7	4	9
Кравата	1	0	3	2	4

Следећа табела садржи цене ових производа у динарима:

Производ	Цена (дин)
Ципеле	11.250
Кошуља	6.500
Каиш	4.750
Панталоне	2.500
Чарапе (пар)	750
Кравата	3.500

Ћелија испод садржи податке из ове две табеле предстаљене у облику листе:

```
In [1]: # изврши ову ћелију
proizvodi = [
    ["Ципеле",      5, 17, 3, 11, 9],
    ["Кошуља",     8, 6, 7, 4, 0],
    ["Каиш",        4, 1, 3, 5, 1],
    ["Панталоне",   4, 2, 6, 4, 5],
    ["Чарапе (пар)", 8, 9, 7, 4, 9],
    ["Кравата",     1, 0, 3, 2, 4]]
cene = [
    ["Ципеле",      11250],
    ["Кошуља",     6500],
    ["Каиш",        4750],
    ["Панталоне",   2500],
    ["Чарапе (пар)", 750],
    ["Кравата",     3500]]
```

(а) Додај нову ћелију испод ове па представи обе табеле структуром *DataFrame*. Индексирај обе табеле.

(б) Додај нову ћелију испод ове па израчунај колико је укупно у том месецу продато ципела, кошуља, каишева, панталона, чарапа и кравата.

(е) Додај нову ћелију испод ове па израчунај колико је у том месецу компанија зарадила на продаји ципела, колико на продаји кошуља, колико на продаји каишева, итд.

(е*) Додај нову ћелију испод ове па израчунај и испиши зараду сваке пословне јединице у том месецу.