

12. час: Модификације табеле и рачунање у табели

На овом часу ћемо говорити о:

1. додавању нове врсте или колоне табели; и
2. транспонувању табеле.

12.1. Додавање нове врсте или колоне табели

Понекад је важно да вредности које смо израчунали на основу података у табели (на пример, просеке оцена по ученицима) и сачувамо у табели. Ако табела нема за то предвиђену колону или врсту лако је можемо додати!

Вратимо се поново на пример са оценама ученика кога смо видели на раније:

```
In [1]: # изврши ову ћелију
import pandas as pd
razred = [
    ["Ана", 5, 3, 5, 2, 4, 5],
    ["Бојан", 5, 5, 5, 5, 5, 5],
    ["Влада", 4, 5, 3, 4, 5, 4],
    ["Гордана", 5, 5, 5, 5, 5, 5],
    ["Дејан", 3, 4, 2, 3, 3, 4],
    ["Ђорђе", 4, 5, 3, 4, 5, 4],
    ["Елена", 3, 3, 3, 4, 2, 3],
    ["Жаклина", 5, 5, 4, 5, 4, 5],
    ["Зоран", 4, 5, 4, 4, 3, 5],
    ["Ивана", 2, 2, 2, 2, 2, 5],
    ["Јасна", 3, 4, 5, 4, 5, 5]]
ocene = pd.DataFrame(razred)
ocene.columns=["Име", "Српски", "Енглески", "Математика", "Физика", "Хемија", "Ликовно"]
ocene1 = ocene.set_index("Име")
ocene1
```

```
Out[1]:
```

	Српски	Енглески	Математика	Физика	Хемија	Ликовно
Име						
Ана	5	3	5	2	4	5
Бојан	5	5	5	5	5	5
Влада	4	5	3	4	5	4
Гордана	5	5	5	5	5	5
Дејан	3	4	2	3	3	4
Ђорђе	4	5	3	4	5	4
Елена	3	3	3	4	2	3
Жаклина	5	5	4	5	4	5
Зоран	4	5	4	4	3	5
Ивана	2	2	2	2	2	5
Јасна	3	4	5	4	5	5

	Српски	Енглески	Математика	Физика	Хемија	Ликовно
Име						
Ана	5	3	5	2	4	5
Бојан	5	5	5	5	5	5
Влада	4	5	3	4	5	4
Гордана	5	5	5	5	5	5
Дејан	3	4	2	3	3	4
Ђорђе	4	5	3	4	5	4
Елена	3	3	3	4	2	3
Жаклина	5	5	4	5	4	5
Зоран	4	5	4	4	3	5
Ивана	2	2	2	2	2	5
Јасна	3	4	5	4	5	5

У овом низу наредни смо пров увезли библиотеку *pandas* са скраћеним именом "pd", од листе података смо направили табелу, тали колонама имена и индексирали смо табелу по имену ученика.

Као што смо раније видели, лако можемо да израчунамо просек оцена сваког ученика и да те податке испишемо. Међутим, ако желимо да просеке запамтимо у табели, треба нам нова колона. Нова колона се табели додаје тако што се просто напише:

```
In [14]: # изврши ову ћелију
ocene1["ПросекУч"] = 0.0
```

Ако покушамо да непостојећој колони додамо неку вредност, систем ће сам на крај табеле додати нову колону и попунити је наведеним вредностима:

```
In [15]: # изврши ову ћелију
osene1
```

```
Out[15]:
```

	Српски	Енглески	Математика	Физика	Хемија	Ликовно	ПросекУч
Име							
Ана	5	3	5	2	4	5	0.0
Бојан	5	5	5	5	5	5	0.0
Влада	4	5	3	4	5	4	0.0
Гордана	5	5	5	5	5	5	0.0
Дејан	3	4	2	3	3	4	0.0
Ђорђе	4	5	3	4	5	4	0.0
Елена	3	3	3	4	2	3	0.0
Жаклина	5	5	4	5	4	5	0.0
Зоран	4	5	4	4	3	5	0.0
Ивана	2	2	2	2	2	5	0.0
Јасна	3	4	5	4	5	5	0.0

Сада ћемо у једном *for*-циклусу да прођемо кроз индексну колону табеле и да за сваки ред табеле израчунамо просек бројева уписаних у колоне "Српски"--"Ликовно". (Не смемо да рачунамо просек целог реда, јер редови сада садрже и колону "ПросекУч" која не сме да се укључи у рачун просека!)

```
In [16]: # изврши ову ћелију
for ucenik in osene1.index:
    osene1.loc[ucenik, "ПросекУч"] = osene1.loc[ucenik, "Српски":"Ликовно"].mean()
```

Ево како изгледа нова табела:

```
In [17]: # изврши ову ћелију
осене1
```

```
Out[17]:
```

	Српски	Енглески	Математика	Физика	Хемија	Ликовно	ПросекУч
Име							
Ана	5	3	5	2	4	5	4.000000
Бојан	5	5	5	5	5	5	5.000000
Влада	4	5	3	4	5	4	4.166667
Гордана	5	5	5	5	5	5	5.000000
Дејан	3	4	2	3	3	4	3.166667
Ђорђе	4	5	3	4	5	4	4.166667
Елена	3	3	3	4	2	3	3.000000
Жаклина	5	5	4	5	4	5	4.666667
Зоран	4	5	4	4	3	5	4.166667
Ивана	2	2	2	2	2	5	2.500000
Јасна	3	4	5	4	5	5	4.333333

Да бисмо израчунали просечну оцену за сваки предмет, додаћемо нову врсту:

```
In [18]: # изврши ову ћелију
осене1.loc["ПросекПр"] = 0.0
осене1
```

```
Out[18]:
```

	Српски	Енглески	Математика	Физика	Хемија	Ликовно	ПросекУч
Име							
Ана	5.0	3.0	5.0	2.0	4.0	5.0	4.000000
Бојан	5.0	5.0	5.0	5.0	5.0	5.0	5.000000
Влада	4.0	5.0	3.0	4.0	5.0	4.0	4.166667
Гордана	5.0	5.0	5.0	5.0	5.0	5.0	5.000000
Дејан	3.0	4.0	2.0	3.0	3.0	4.0	3.166667
Ђорђе	4.0	5.0	3.0	4.0	5.0	4.0	4.166667
Елена	3.0	3.0	3.0	4.0	2.0	3.0	3.000000
Жаклина	5.0	5.0	4.0	5.0	4.0	5.0	4.666667
Зоран	4.0	5.0	4.0	4.0	3.0	5.0	4.166667
Ивана	2.0	2.0	2.0	2.0	2.0	5.0	2.500000
Јасна	3.0	4.0	5.0	4.0	5.0	5.0	4.333333
ПросекПр	0.0	0.0	0.0	0.0	0.0	0.0	0.000000

Овде треба да застанемо за тренутак и да се подсетимо да се запис облика `осене1["ПросекУч"]` односи на колоне табеле тако да ће наредба

```
ocene1["ПросекУч"] = 0.0
```

додати нову колону попуњену нулама, док се запис `ocene1.loc["ПросекПр"]` односи на врсте табеле, па ће наредба

```
ocene1.loc["ПросекПр"] = 0.0
```

додати нову врсту попуњену нулама (што се и десило у примеру).

```
In [19]: # изврши ову ћелију
for predmet in ocene1.columns:
    ocene1.loc["ПросекПр", predmet] = ocene1.loc["Ана":"Јасна", predmet].mean()
```

```
In [20]: ocene1
```

```
Out[20]:
```

	Српски	Енглески	Математика	Физика	Хемија	Ликовно	ПросекУч
Име							
Ана	5.000000	3.000000	5.000000	2.000000	4.000000	5.000000	4.000000
Бојан	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000
Влада	4.000000	5.000000	3.000000	4.000000	5.000000	4.000000	4.166667
Гордана	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000
Дејан	3.000000	4.000000	2.000000	3.000000	3.000000	4.000000	3.166667
Ђорђе	4.000000	5.000000	3.000000	4.000000	5.000000	4.000000	4.166667
Елена	3.000000	3.000000	3.000000	4.000000	2.000000	3.000000	3.000000
Жаклина	5.000000	5.000000	4.000000	5.000000	4.000000	5.000000	4.666667
Зоран	4.000000	5.000000	4.000000	4.000000	3.000000	5.000000	4.166667
Ивана	2.000000	2.000000	2.000000	2.000000	2.000000	5.000000	2.500000
Јасна	3.000000	4.000000	5.000000	4.000000	5.000000	5.000000	4.333333
ПросекПр	3.909091	4.181818	3.727273	3.818182	3.909091	4.545455	4.015152

12.2. Транспонованье табеле

Замена врста и колона табеле се зове *транспонованье*. Приликом транспонованья имена колона полазне табеле постају индекси нове табеле, док индексна колона полазне табеле постаје одређује имена колона нове табеле:

	Пол	Старост	Тежина	Висина
Име				
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
Елена	ж	14	49	161
Жаклина	ж	15	52	164
Зоран	м	15	57	167
Ивана	ж	13	45	158
Јасна	ж	14	51	162

Име	Ана	Бојан	Влада	Гордана	Дејан	Ђорђе	Елена	Жаклина	Зоран	Ивана	Јасна
Пол	ж	м	м	ж	м	м	ж	ж	м	ж	ж
Старост	13	14	13	15	15	13	14	15	15	13	14
Тежина	46	52	47	54	56	45	49	52	57	45	51
Висина	160	165	157	165	163	159	161	164	167	158	162

Табела се транспонује тако што се на њу примени функција `T` која као резултат враћа нову, транспоновану табелу. Ево примера:

```
In [2]: # изврши ову ћелију
import pandas as pd
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],
    ["Елена", "ж", 14, 49, 161],
    ["Жаклина", "ж", 15, 52, 164],
    ["Зоран", "м", 15, 57, 167],
    ["Ивана", "ж", 13, 45, 158],
    ["Јасна", "ж", 14, 51, 162]]
tabela = pd.DataFrame(podaci)
tabela.columns=["Име", "Пол", "Старост", "Тежина", "Висина"]
tabela1=tabela.set_index("Име")

tabela2 = tabela1.T
```

Погледајмо сада како изгледа полазна табела и како изгледа транспонована табела, шта су им врсте, а шта колоне. Кренимо од полазне табеле:

```
In [22]: # изврши ову ћелију
tabela1
```

```
Out[22]:
```

	Пол	Старост	Тежина	Висина
Име				
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
Елена	ж	14	49	161
Жаклина	ж	15	52	164
Зоран	м	15	57	167
Ивана	ж	13	45	158
Јасна	ж	14	51	162

```
In [23]: # изврши ову ћелију
tabela1.index
```

```
Out[23]: Index(['Ана', 'Бојан', 'Влада', 'Гордана', 'Дејан', 'Ђорђе', 'Елена',
               'Жаклина', 'Зоран', 'Ивана', 'Јасна'],
              dtype='object', name='Име')
```

```
In [24]: # изврши ову ћелију
tabela1.columns
```

```
Out[24]: Index(['Пол', 'Старост', 'Тежина', 'Висина'], dtype='object')
```

А сада транспонована табела:

```
In [25]: # изврши ову ћелију
tabela2
```

```
Out[25]:
```

	Име	Ана	Бојан	Влада	Гордана	Дејан	Ђорђе	Елена	Жаклина	Зоран	Ивана	Јасна
Пол	ж	м	м	ж	м	м	ж	ж	ж	м	ж	ж
Старост	13	14	13	15	15	13	14	15	15	13	14	
Тежина	46	52	47	54	56	45	49	52	57	45	51	
Висина	160	165	157	165	163	159	161	164	167	158	162	

```
In [26]: # изврши ову ћелију
tabela2.index
```

```
Out[26]: Index(['Пол', 'Старост', 'Тежина', 'Висина'], dtype='object')
```

```
In [27]: # изврши ову ћелију
tabela2.columns
```

```
Out[27]: Index(['Ана', 'Бојан', 'Влада', 'Гордана', 'Дејан', 'Ђорђе', 'Елена',
               'Жаклина', 'Зоран', 'Ивана', 'Јасна'],
              dtype='object', name='Име')
```

Ево поново примера са оценама:

```
In [3]: # изврши ову ћелију
razred = [
    ["Ана", 5, 3, 5, 2, 4, 5],
    ["Бојан", 5, 5, 5, 5, 5, 5],
    ["Влада", 4, 5, 3, 4, 5, 4],
    ["Гордана", 5, 5, 5, 5, 5, 5],
    ["Дејан", 3, 4, 2, 3, 3, 4],
    ["Ђорђе", 4, 5, 3, 4, 5, 4],
    ["Елена", 3, 3, 3, 4, 2, 3],
    ["Жаклина", 5, 5, 4, 5, 4, 5],
    ["Зоран", 4, 5, 4, 4, 3, 5],
    ["Ивана", 2, 2, 2, 2, 2, 5],
    ["Јасна", 3, 4, 5, 4, 5, 5]]
ocene = pd.DataFrame(razred)
ocene.columns=["Име", "Српски", "Енглески", "Математика", "Физика", "Хемија", "Ликовно"]
ocene1 = ocene.set_index("Име")
ocene1
```

```
Out[3]:
```

	Српски	Енглески	Математика	Физика	Хемија	Ликовно
Име						
Ана	5	3	5	2	4	5
Бојан	5	5	5	5	5	5
Влада	4	5	3	4	5	4
Гордана	5	5	5	5	5	5
Дејан	3	4	2	3	3	4
Ђорђе	4	5	3	4	5	4
Елена	3	3	3	4	2	3
Жаклина	5	5	4	5	4	5
Зоран	4	5	4	4	3	5
Ивана	2	2	2	2	2	5
Јасна	3	4	5	4	5	5

Како смо раније већ видели, просек оцена по предметима добијамо лако:


```
In [25]: # изврши ову ћелију
for predmet in ocene1.columns:
    print(predmet, "->", ocene1[predmet].mean())
```

```
Српски -> 3.909090909090909
Енглески -> 4.181818181818182
Математика -> 3.727272727272727
Физика -> 3.8181818181818183
Хемија -> 3.909090909090909
Ликовно -> 4.545454545454546
```

Да бисмо добили просек оцена по ученицима, можемо да приступимо врстама табеле користећи функцију `loc` како смо то већ видели, али можемо и овако. Прво ћемо да транспонујемо табелу:

```
In [29]: # изврши ову ћелију
ocene2 = ocene1.T
ocene2
```

```
Out[29]:
```

	Име	Ана	Бојан	Влада	Гордана	Дејан	Ђорђе	Елена	Жаклина	Зоран	Ивана	Јасна
Српски	5	5	4	5	3	4	3	5	4	2	3	
Енглески	3	5	5	5	4	5	3	5	5	2	4	
Математика	5	5	3	5	2	3	3	4	4	2	3	
Физика	2	5	4	5	3	4	4	5	4	2	4	
Хемија	4	5	5	5	3	5	2	4	3	2	3	
Ликовно	5	5	4	5	4	4	3	5	5	5	3	

па ћемо на транспоновану табелу применити исту идеју (рачунање просека по колонама, јер су колоне транспоноване табеле заправо врсте полазне табеле.):

```
In [30]: # изврши ову ћелију
for ucenik in ocene2.columns:
    print(ucenik, "->", ocene2[ucenik].mean())
```

```
Ана -> 4.0
Бојан -> 5.0
Влада -> 4.166666666666667
Гордана -> 5.0
Дејан -> 3.1666666666666665
Ђорђе -> 4.166666666666667
Елена -> 3.0
Жаклина -> 4.666666666666667
Зоран -> 4.166666666666667
Ивана -> 2.5
Јасна -> 4.333333333333333
```

12.3. Задаци

Задатак 1. Биолози су до данас описали око два милиона врта живих бића. Сва она су подељена у пет царстава и приближан број врста живих бића по царствима је дат следећом табелом:

Царство	Број врста
Животиње	1.400.000
Билјке	290.000
Гљиве	100.000
Протисти	200.000
Монере	10.000

Додај нову врсту табели која се зове "Укупно" и у њу унеси укупан број врста. Немој рачунати ручно, већ примени функцију `sum` на одговарајућу колону табеле.

Задатак 2. У наредној ћелији дати су подаци о тежини (у килограмима) и дужини/висини (у центиметрима) једног дечака од његовог шестог месеца па до његових 6,5 година. Прикажи ове податке графиком.

In [31]:

```
# изврши ову ћелију
merenja = ["6 мес", "1,5 год", "2,5 год", "3,5 год", "4,5 год", "5,5 год", "6,5 год"]
masaKG = [5.9, 11.5, 14.8, 20.5, 22.0, 24.2, 29.0]
visinaCM = [62.0, 84.0, 97.0, 115.0, 122.5, 131.5, 135.0]
```

Транспонуј табелу, тако транспонованој табели додај нову колону, па за свако мерење (што је сада врста у новој, транспонованој табели) израчунај *BMI* (*body mass index*) по формули:

$$BMI = \frac{\text{masa u kilogramima}}{(\text{visina u metrima})^2}$$

Задатак 3. У следећој табели су приказане највише и најниже температуре (у Целзијусима) икада измерене на континентима:

Континент:	Европа	Азија	Африка	Северна Америка	Јужна Америка	Аустралија	Антарктик
Највиша забележена темп:	48	54	55	56.7	48.9	50.7	19.8
Најнижа забележена темп:	-58.1	-67.8	-23.9	-63	-32.8	-23	-89.2

(Подаци су преузети са странице <https://www.space.com/17816-earth-temperature.html>) (<https://www.space.com/17816-earth-temperature.html>). Додај нови ред овој табели па у њега упиши максимални температурни распон за сваки континент, који се добија када се од највише забележене температуре одузме најнижа забележена температура.

In []:

