

Час 6. Проценти и секторски дијаграми ¶

На овом часу ћеш научити:

1. како да извршиш *фреквенцијску анализу* низа података (што је једно компликовано име за једну веома једноставну и корисну ствар);
2. како да израчунаш заступљеност неке величине у процентима; и
3. како да прикажеш заступљеност неких величина изражену у процентима на секторском дијаграму.

6.1. Фреквенцијска анализа низа података

Фреквенцијска анализа низа података се своди на то да се преброје подаци у низу и да се на основу тако добијених *фреквенција података* покуша доћи до неког закључка (реч *фреквенција* значи "учесталост"). На пример, у једном разреду има 30 ученика и њихове оцене из информатике су дате у следећем низу:

```
In [1]: # изврши ову ћелију!  
oceneInf = [3, 4, 5, 4, 5, 3, 4, 5, 2, 4, 5, 4, 5, 4, 2, 3, 1, 4, 5, 4, 3, 2, 3, 4]
```

Фреквенцијска анализа овог низа се своди на то да се утврди колико има петица, четворки, тројки, двојки и јединица у том разреду. Уместо да то урадимо ручно (рецимо, да направимо табелицу па да "лупамо рецке"), пустићемо Пајтон да то уради за нас. Пајтон има уграђену функцију `count` која може да преброји колико пута се у некој листи јавља неки елемент:

```
In [2]: # изврши ову ћелију!  
ocena5 = oceneInf.count(5)  
ocena4 = oceneInf.count(4)  
ocena3 = oceneInf.count(3)  
ocena2 = oceneInf.count(2)  
ocena1 = oceneInf.count(1)  
print("Ocene u razredu su raspodeljene kako sledi:")  
print("Ocenu 5 ima", ocena5, "ucenika")  
print("Ocenu 4 ima", ocena4, "ucenika")  
print("Ocenu 3 ima", ocena3, "ucenika")  
print("Ocenu 2 ima", ocena2, "ucenika")  
print("Ocenu 1 ima", ocena1, "ucenika")
```

```
Ocene u razredu su raspodeljene kako sledi:  
Ocenu 5 ima 9 ucenika  
Ocenu 4 ima 11 ucenika  
Ocenu 3 ima 6 ucenika  
Ocenu 2 ima 3 ucenika  
Ocenu 1 ima 1 ucenika
```

6.2. Проценти

Реч *проценат* потиче од латинског *pro centum* што значи "у стотини". Зато је

$$47\% = \frac{47}{100} = \text{четрдесет седам од сто.}$$

На пример, у једној школи која има 856 ученика њих 25% су одлични. Колико има одличних ученика у тој школи?

Одговор. У тој школи има 214 одличних ученика јер је

$$856 \cdot 25\% = 856 \cdot \frac{25}{100} = 214.$$

У истој школи има 326 ученика са добрим просеком. Који је то процент укупног броја ученика?

Одговор. Нека је то $x\%$ ученика. Онда је

$$856 \cdot x\% = 856 \cdot \frac{x}{100} = 326.$$

Одатле се лако добија да је

$$x = \frac{326 \cdot 100}{856} \approx 38,08.$$

Дакле, приближно 38,08% ученика те школе има добар успех.

Вратићемо се сада на пример са оценама из информатике и приказаћемо податке овај пут у процентима. Ако си извршио претходна два мала Пајтон програма систем у меморији већ има све податке које смо до сада израчунали, па можемо да наставимо овако:

```
In [3]: # изврши ову ћелију!
ukupno = len(oceneInf)
procenat5 = 100.0 * ocena5 / ukupno
procenat4 = 100.0 * ocena4 / ukupno
procenat3 = 100.0 * ocena3 / ukupno
procenat2 = 100.0 * ocena2 / ukupno
procenat1 = 100.0 * ocena1 / ukupno
print("Ocene u razredu su raspodeljene kako sledi:")
print("Ocenu 5 ima ", procenat5, "% ucenika", sep="")
print("Ocenu 4 ima ", procenat4, "% ucenika", sep="")
print("Ocenu 3 ima ", procenat3, "% ucenika", sep="")
print("Ocenu 2 ima ", procenat2, "% ucenika", sep="")
print("Ocenu 1 ima ", procenat1, "% ucenika", sep="")
```

Ocene u razredu su raspodeljene kako sledi:

Ocenu 5 ima 30.0% ucenika
Ocenu 4 ima 36.666666666666664% ucenika
Ocenu 3 ima 20.0% ucenika
Ocenu 2 ima 10.0% ucenika
Ocenu 1 ima 3.3333333333333335% ucenika

Дакле, фреквенцијска анализа нам даје вредности у *конкретном износу* (рецимо, у разреду има 9 петица, 11 четворки и тако даље). Ако ове податке представимо процентима добијамо *дистрибуцију* оцена која нам говори о *заступљености* оцена у *релативним* односима. Тако, горњим секторским дијаграмом је заправо показано да у овом разреду има 30% петица ($9 / 30 = 30\%$), 36,67% четворки ($11 / 30 = 36,67\%$), 20% тројки ($6 / 30 = 20\%$), 10% двојки ($3 / 30 = 10\%$) и 3,33% јединица ($1 / 30 = 3,33\%$).

6.3. Секторски дијаграми

У ситуацијама када се приказује колико процената које компоненте учествује у саставу неке целине погодно је податке приказати *секторским дијаграмом* који представља круг исечен на исечке попут пице. Круг тада представља целину (100%), док исечци представљају компоненте које учествују у целини исказане у процентима.

На пример, оцене из информатике су у једном разреду са 30 ученика расподељене овако:

Оцена	Заступљеност оцене
5	30,00%
4	36,67%
3	20,00%
2	10,00%
1	3,33%

Податке из ове табеле ћемо илустровати секторским дијаграмом. Прво ћемо увести библиотеку:

```
In [4]: # изврши ову ћелију!  
import matplotlib.pyplot as plt
```

Онда ћемо податке записати помоћу два низа овако:

```
In [5]: # изврши ову ћелију!  
procenti = [30.00, 36.67, 20.00, 10.00, 3.33]  
ocene    = ["5", "4", "3", "2", "1"]
```

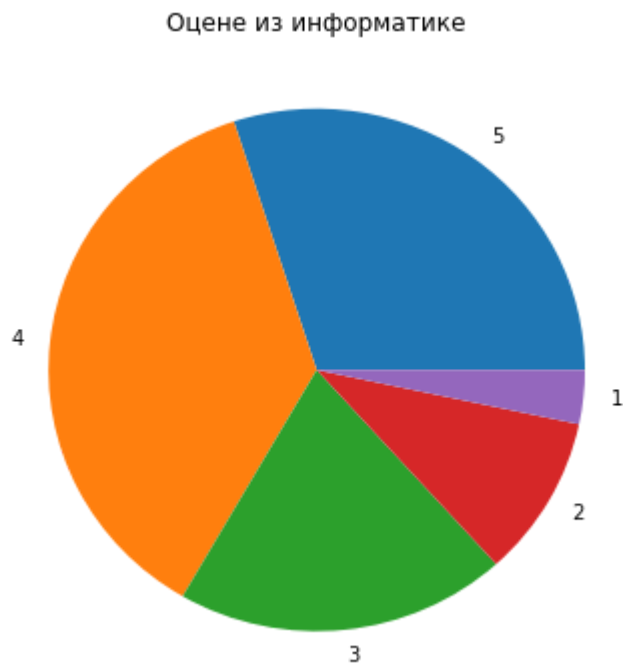
И сада можемо да пређемо на представљање података дијаграмом. Функција која податке представља секторским дијаграмом се зове `pie` зато што се у америчком сленгу секторски дијаграми зову *pie charts* = "тортасти дијаграми". Први аргумент представља низ процената, док се другим задају ознаке (енгл. *label* = ознака):

```
In [6]: # изврши ову ћелију!  
plt.pie(procenti, labels=ocene)  
plt.title("Оцене из информатике")  
plt.show()  
plt.close()
```



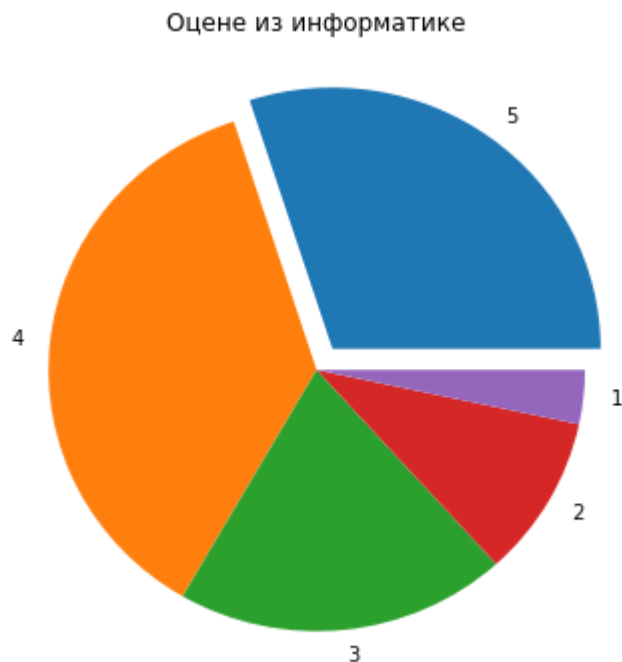
Није лоше, али није ни сјајно јер је дијаграм мало јајаст. Овај проблем можемо решити тако што ћемо рећи систему за цртање да дијаграм нацрта унутар квадрата тако што ћемо поставити димензије слике које одређују квадрат:

```
In [7]: # изврши ову ћелију!  
plt.figure(figsize=(6,6))  
plt.pie(procenti, labels=ocene)  
plt.title("Оцене из информатике")  
plt.show()  
plt.close()
```



Ако желимо да нагласимо број петица у разреду можемо одговарајући сектор да "измакнемо" мало из средишта. То се постиже аргументом `explode` функције `pie` који за сваки податак у низу каже колико треба да га измакнемо из средишта (0 = не треба изместити сектор из средишта дијаграма; што је већи број, то је и измештање веће).

```
In [8]: # изврши ову ћелију!
procenti = [30.00, 36.67, 20.00, 10.00, 3.33]
ocene    = ["5", "4", "3", "2", "1"]
izmestanje = [0.1, 0, 0, 0, 0]
plt.figure(figsize=(6,6))
plt.pie(procenti, labels=ocene, explode=izmestanje)
plt.title("Оцене из информатике")
plt.show()
plt.close()
```



Као други пример посматрајмо структуру наше атмосфере. Наша атмосфера је смеша разних гасова, а сув ваздух има следећи састав:

Гас	Заступљеност (%)
Азот	78,08
Кисеоник	20,94
Аргон	0,93
Угљен диоксид	0,05

Важно је напоменути да *ово није тачан састав наше атмосфере*: у њој има и других гасова, али у траговима (како то хемичари воле да кажу) па они нису приказани у табели.

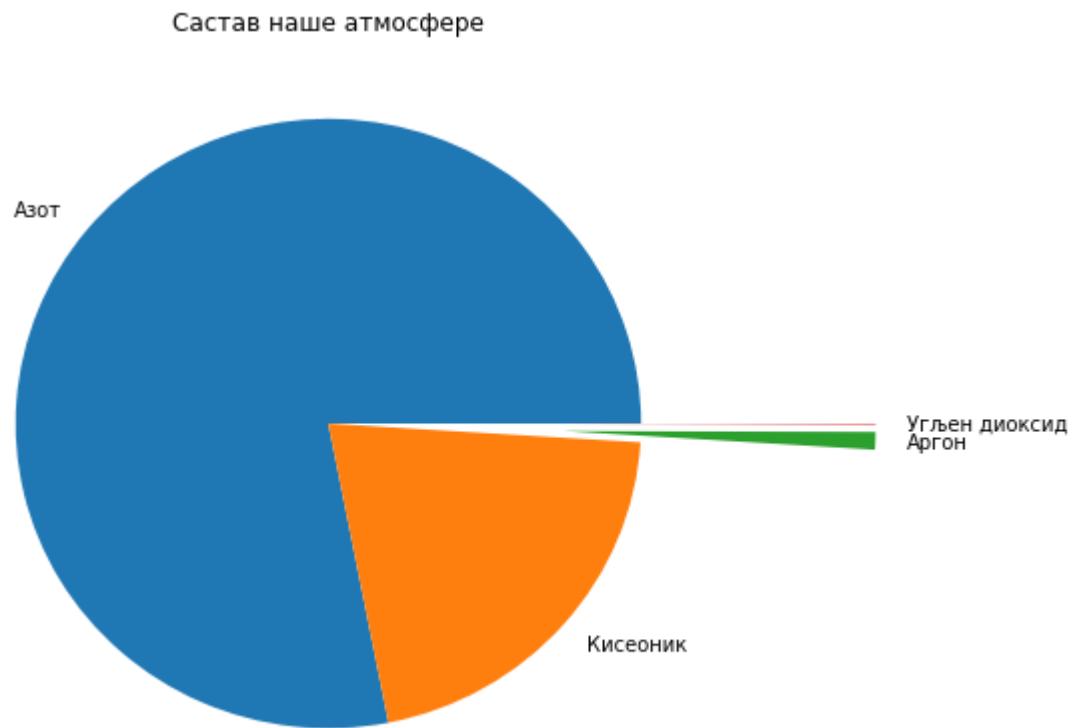
Структуру атмосфере ћемо представити секторским дијаграмом овако:

```
In [9]: # изврши ову ћелију!
procenti = [78.08, 20.94, 0.93, 0.05]
oznake = ["Азот", "Кисеоник", "Аргон", "Угљен диоксид"]
plt.figure(figsize=(7,7))
plt.pie(procenti, labels=oznake)
plt.title("Састав наше атмосфере")
plt.show()
plt.close()
```



Опет наилазимо на проблем: ознаке за последња два податка су се преклопиле јер се ради о веома уском секторима. Да бисмо решили проблем можемо још више да повећамо дијаграм, а можемо пробати и да последња два "мала" сектора "измакнемо" из средишта, овако:

```
In [10]: # изврши ову ћелију!
procenti = [78.08, 20.94, 0.93, 0.05]
oznake = ["Азот", "Кисеоник", "Аргон", "Угљен диоксид"]
izmestanje = [0, 0, 0.75, 0.75]
plt.figure(figsize=(7,7))
plt.pie(procenti, labels=oznake, explode=izmestanje)
plt.title("Састав наше атмосфере")
plt.show()
plt.close()
```

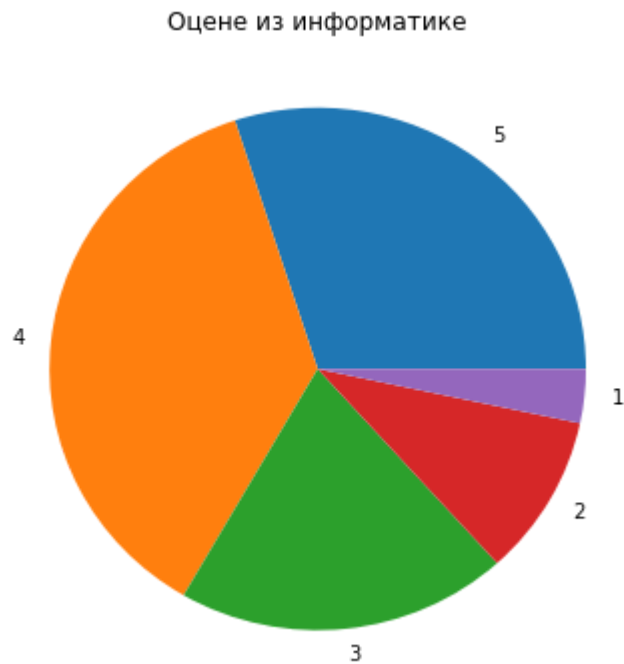


Низ вредности који се прослеђује функцији `pie` не морају бити проценти (односно, то не мора бити низ бројева чији збир је 100). Низ вредности може бити низ произвољних вредности, а функција `pie` ће их аутоматски прерачунати у проценте које ће приказати на дијаграму. Рецимо, у примеру са самог почетка имамо овакву расподелу оцена из информатике:

Оцена	Број ученика
5	9
4	11
3	6
2	3
1	1
Укупно	30

Ево Пајтон програма који приказује податке из табеле:

```
In [11]: # изврши ову ћелију!  
podaci = [9, 11, 6, 3, 1]  
oznake = ["5", "4", "3", "2", "1"]  
plt.figure(figsize=(6,6))  
plt.pie(podaci, labels=oznake)  
plt.title("Оцене из информатике")  
plt.show()  
plt.close()
```



Ако желимо да нагласимо двојке и четворке, то можемо учинити на следећи начин:


```
In [12]: # изврши ову ћелију!
podaci = [9, 11, 6, 3, 1]
oznake = ["5", "4", "3", "2", "1"]
izdvoj = [0, 0.2, 0, 0.2, 0]
plt.figure(figsize=(6,6))
plt.pie(podaci, labels=oznake, explode=izdvoj)
plt.title("Оцене из информатике")
plt.show()
plt.close()
```



6.4. Задаци

Задатак 1. (а) Једна књига је коштала 2460 динара, а онда је књижара снизила њену цену за 20%. Колико сада кошта књига?

(б) Друга књига у истој књижари након снижења од 20% кошта 1480 динара. Колико је коштала пре снижења?

Задатак 2. Према подацима Републичког завода за статистику старосна структура становника Републике Србије на дан 31.12.2017. изгледа овако:

Старосна група	Број становника
Предшколски узраст (0-6)	457.414
Основношколски узраст (7-14)	547.577
Средњошколски узраст (15-18)	280.080
Радно способни (19-65)	4.435.126
Пензионери (65 и више)	1.281.247
Укупно	7.001.444

Изразити старосну структуру становника Републике Србије на дан 31.12.2017. у процентима и илустровати податке секторским дијаграмом

Задатак 3. Изабел је, гледајући како пада киша у Маконду, одлучила да целе једне године води метеоролошки дневник. Кадогд би киша пала у понедељак у дневник би уписала 1; ако би киша пала у уторак у дневник би уписала 2; за среду би уписала 3 и тако редом до недеље (за кишовите недеље у дневник би уписивала 7). Тако је добила следећи низ бројева:

```
In [13]: IzabelinDnevnik = [1,2,4,7,2,4,7,6,7,5,6,7,3,5,7,1,3,6,2,3,4,2,3,1,4,7,7,  
                           6,5,6,4,5,6,2,3,4,5,1,3,4,2,5,7,2,3,5,3,5,7,6,7,2,3,7,  
                           1,2,3,4,5,6,7,2,7,3,4,1,5,6,1,2,4,5,6,7,1,3,4,1,2,3,4,  
                           2,5,7,6,4,5,6,1,3,7,5,7,1,2,3,7,7,3,4,7,1,2,4,7,4,7,2,  
                           3,4,4,6,8,1,7,7,7,3,4,5,6,7,1,2,4,7,1,2,3,1,7,2,7]
```

(а) Колико кишних дана је Изабел уписала у свој дневник? (Додај нову ћелију испод ове и израчунај користећи функцију `len`.)

(б) Исажи број кишних дана у процентима, ако знаш да та година није била преступна. (Додај нову ћелију испод ове и израчунај.)

(в) Утврди колико је Изабел регистровала кишних понедељака, уторака, среда, четвртака, петак и субота у тој години. (Додај нову ћелију испод ове и израчунај користећи функцију `count`.)

(г) Исажи у процентима број кишних понедељака у односу на број свих понедељака у тој години, којих је било 53. (Додај нову ћелију испод ове и израчунај.)

(д) Утврди најкишнији дан у недељи у Изабелином дневнику и прикажи податке добијене под
(в) секторским дијаграмом. Издвој у дијаграму најкишнији дан. (Додај нову ћелију испод ове.)