



Министерство науки и высшего образования Российской Федерации  
Калужский филиал  
федерального государственного бюджетного  
образовательного учреждения высшего образования  
«Московский государственный технический университет имени Н.Э. Баумана  
(национальный исследовательский университет)»  
(КФ МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИУК «Информатика и управление»

КАФЕДРА ИУК4 «Программное обеспечение ЭВМ, информационные технологии»

## ДОМАШНЯЯ РАБОТА №2

### «Библиотека Pandas»

ДИСЦИПЛИНА: «Высокоуровневое программирование»

Выполнил: студент гр. ИУК4-32Б \_\_\_\_\_ (Подпись) ( Карельский М.К. )  
(Ф.И.О.)

Проверил: \_\_\_\_\_ (Подпись) ( Ерохин И.И. )  
(Ф.И.О.)

Дата сдачи (защиты):

Результаты сдачи (защиты):

- Балльная оценка:

- Оценка:

Калуга , 2021

**Цель:** формирование практических навыков работы с библиотеками numpy и pandas, а также применения средств визуализации данных языка Python.

**Задачи:**

1. Ознакомиться с функциональными возможностями библиотек numpy и pandas.
2. Изучить средства визуализации языка Python.

**Вариант 10**

**Задание:**

Считать данные из CSV файла в структуру DataFrame. Добавить в DataFrame еще один столбец, в котором содержится количество символов в строке идентификатора продукта. Определить возрастную группу покупательницы, которая приобрела меньше всего товаров (по стоимости). Построить график зависимости потраченных средств покупателей (среднее значение) от возрастной группы. Вывести результаты для 5 покупателей с самыми большими расходами и 5 с самыми маленькими. Построить трехмерную поверхность – прямоугольный параллелепипед (параметры параллелепипеда задать произвольно).

**Листинг:**

```
import pandas as pd, matplotlib.pyplot as plt, numpy as np
from mpl_toolkits.mplot3d import axes3d, Axes3D

df = pd.read_csv('data.csv')
symbols = df['Product'].str.len()
df['Symbols'] = symbols
print(df)

minCostAge = df.groupby(['BuyerID',
'Age']).sum()['Cost'].idxmin()[1]
print('\nВозраст покупательницы, купившей меньше всего товаров:',
minCostAge)

largest = df[['BuyerID',
'Cost']].groupby(['BuyerID']).sum().nlargest(5, 'Cost')
smallest = df[['BuyerID',
'Cost']].groupby(['BuyerID']).sum().nsmallest(5, 'Cost')
print('\nСамые большие расходы')
print(largest)
print('\nСамые маленькие расходы')
print(smallest)

df[['Age', 'Cost']].groupby(['Age']).mean().plot.bar()
plt.show()

plt.rc('font', size=15)
fig = plt.figure()
ax = Axes3D(fig, auto_add_to_figure=False)
```

```

fig.add_axes(ax)

o = (1, 2, 3)
l, w, h = 1, 2, 3
X = np.array([[o[0], o[0] + l, o[0] + l, o[0], o[0]],
[o[0], o[0] + l, o[0] + l, o[0], o[0]],
[o[0], o[0] + l, o[0] + l, o[0], o[0]],
[o[0], o[0] + l, o[0] + l, o[0], o[0]]])
Y = np.array([[o[1], o[1], o[1] + w, o[1] + w, o[1]],
[o[1], o[1], o[1] + w, o[1] + w, o[1]],
[o[1], o[1], o[1], o[1], o[1]],
[o[1] + w, o[1] + w, o[1] + w, o[1] + w, o[1] + w]])
Z = np.array([[o[2], o[2], o[2], o[2], o[2]],
[o[2] + h, o[2] + h, o[2] + h, o[2] + h, o[2] + h],
[o[2], o[2], o[2] + h, o[2] + h, o[2]],
[o[2], o[2], o[2] + h, o[2] + h, o[2]]])

ax.plot_surface(X, Y, Z, rstride=1, cstride=1)

plt.show()

```

### Результаты:

	BuyerID	Age	Product	Cost	Symbols
0	1	18	Eggs	70	4
1	1	18	Bread	35	5
2	1	18	Milk	80	4
3	2	25	Meat	400	4
4	2	25	Juice	100	5
5	3	40	Eggs	70	4
6	3	40	Milk	80	4
7	3	40	Juice	100	5
8	3	40	Water	30	5
9	4	18	Meat	400	4
10	4	18	Juice	100	5
11	5	25	Eggs	70	4
12	6	60	Bread	35	5
13	6	60	Meat	400	4
14	6	60	Water	30	5
15	7	18	Bread	35	5
16	7	18	Milk	80	4
17	8	40	Eggs	70	4
18	8	40	Bread	35	5
19	9	25	Meat	400	4
20	9	25	Juice	100	5
21	9	25	Water	30	5
22	10	60	Eggs	70	4
23	10	60	Milk	80	4
24	11	40	Bread	35	5
25	11	40	Meat	400	4
26	11	40	Water	30	5
27	12	25	Bread	35	5
28	12	25	Milk	80	4
29	12	25	Meat	400	4
30	12	25	Juice	100	5
31	13	60	Eggs	70	4
32	13	60	Milk	80	4

Возраст покупательницы, купившей меньше всего товаров: 25

Рисунок 1.1. Результат

```

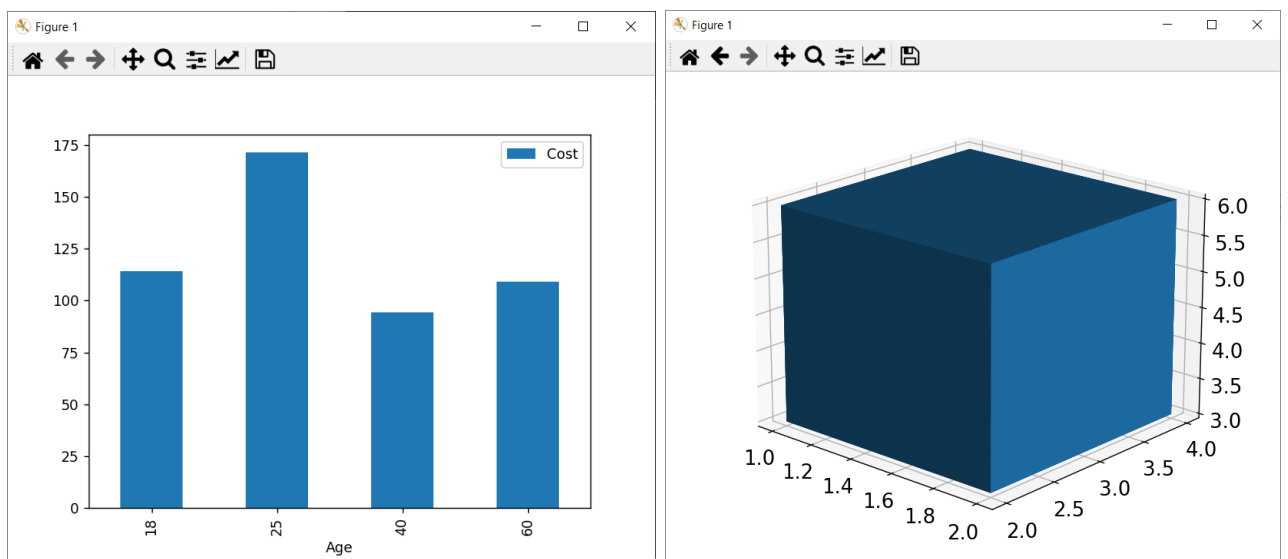
Возраст покупательницы, купившей меньше всего товаров: 25

Самые большие расходы
Cost
BuyerID
12      615
9       530
2       500
4       500
6       465

Самые маленькие расходы
Cost
BuyerID
5        70
8       105
7       115
10      150
13      150

```

**Рисунок 1.2. Результат**



**Рисунок 1.3. Результат**

**Вывод:** в ходе выполнения домашней работы были получены навыки работы с библиотеками Pandas, Numpy, Matplotlib.