|  |  |
| --- | --- |
| **Gerb-BMSTU_01** | **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**

**«Библотека Pandas»**

**ДИСЦИПЛИНА: «Высокоуровневое программирование»**

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

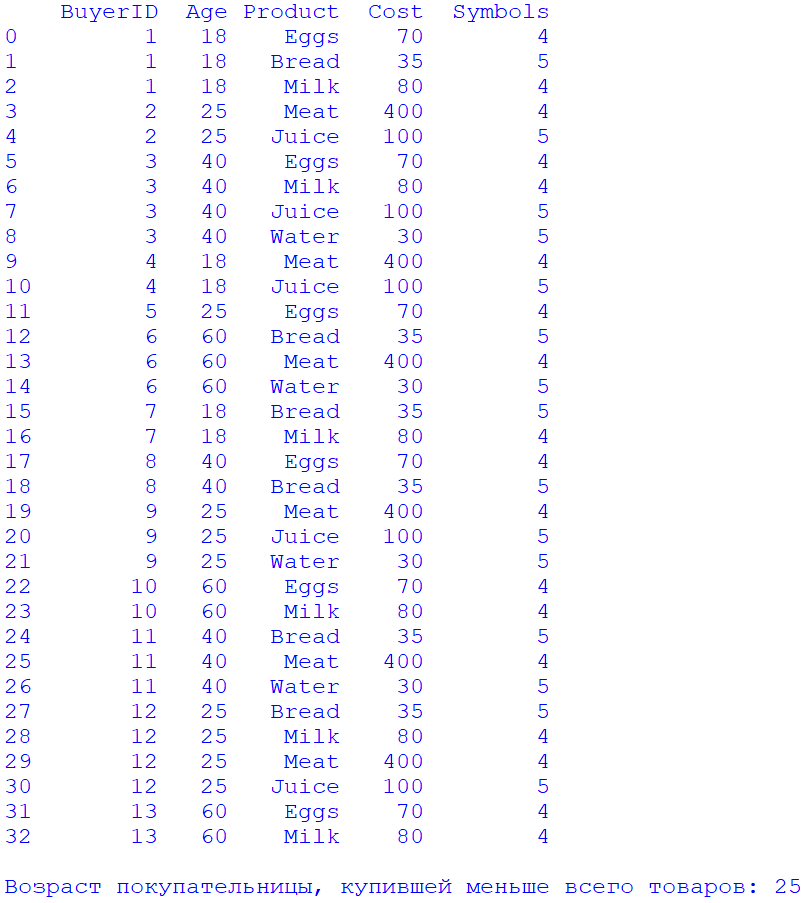
[o[2], o[2], o[2] + h, o[2] + h, o[2]],

[o[2], o[2], o[2] + h, o[2] + h, o[2]]])

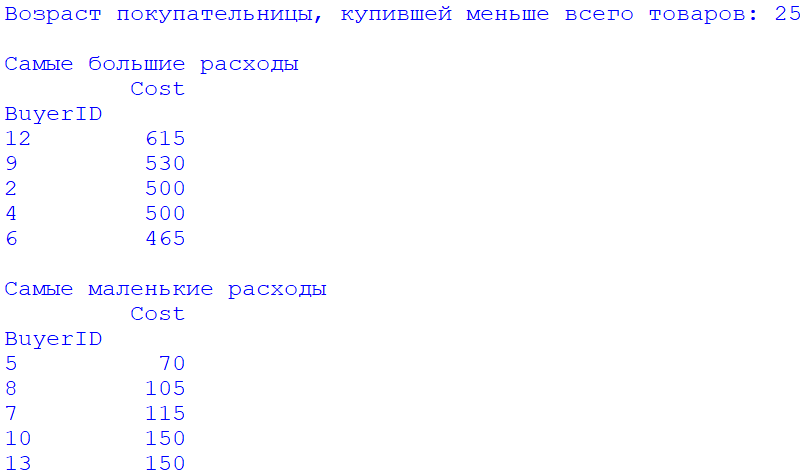
ax.plot\_surface(X, Y, Z, rstride=1, cstride=1)

plt.show()

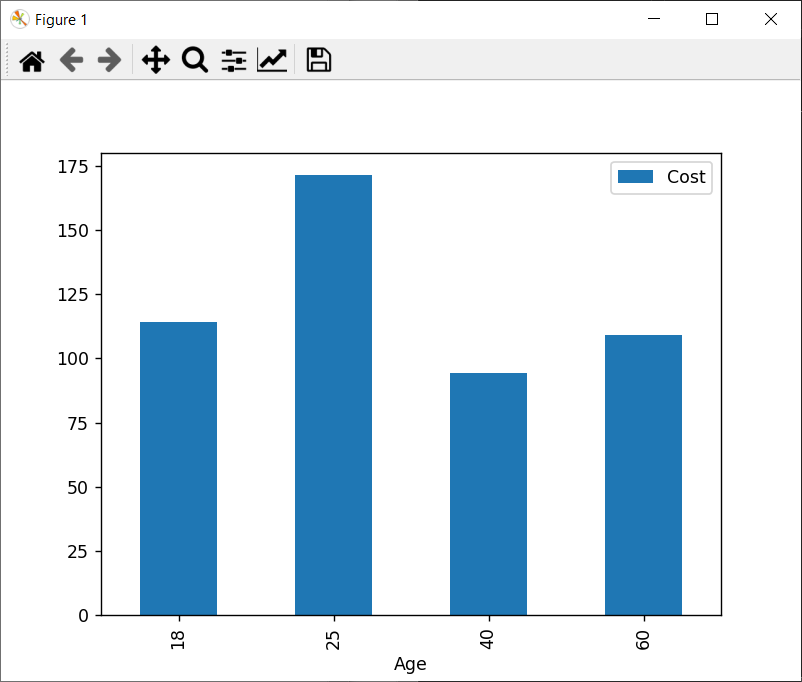
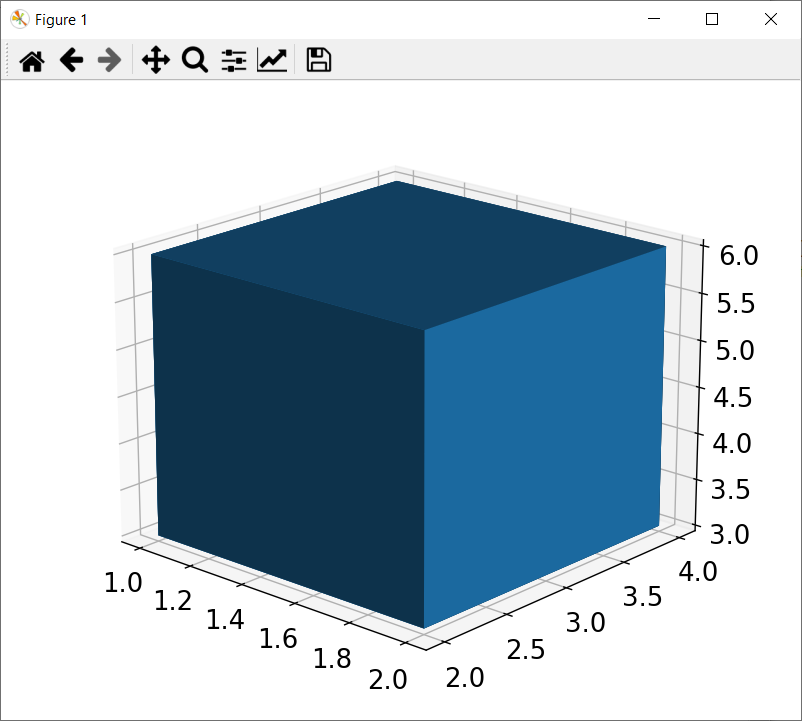
**Результаты:**



**Рисунок 1.1.** Результат



**Рисунок 1.2.** Результат

**Рисунок 1.3.** Результат

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