

Задача 1 Даны значения зарплат из выборки выпускников: 100, 80, 75, 77, 89, 33, 45, 25, 65, 17, 30, 230, 24, 57, 55, 70, 75, 65, 84, 90, 150 Используя только встроенные питоновские функции и структуры данных (т.е. без библиотек numpy, pandas и др.) посчитать среднее арифметическое, смещённую и несмещённую оценки дисперсии, среднее квадратичное отклонение для данной выборки. Можно затем посчитать те же значения с использованием библиотек, чтобы проверить себя.

Рассчитаем среднее арифметическое ( $M_s$ )

In [3]:

```
a = [100, 80, 75, 77, 89, 33, 45, 25, 65, 17, 30, 230, 24, 57, 55]
n = 0
S = 0
for i in a:
    n += 1
    S += i
Ms = S/n
Ms
```

Out [3]:

73.14285714285714

Смещенная дисперсия ( $D$ )

In [4]:

```
MX = 0
for t in a:
    MX += (t - Ms)**2
D = MX/n
D
```

Out [4]:

2135.074829931973

Несмещенная дисперсия ( $D_2$ )

In [5]:

```
D2 = MX/(n-1)
D2
```

Out [5]:

2241.8285714285716

Среднее квадратичное отклонение (Sigma)

In [19]:

```
Sigma = D2**(0.5)
Sigma
```

Out [19]:

47.34795213553139

In [8]:

```
import numpy as np
```

In [9]:

```
np.mean(a)
```

Out [9]:

73.14285714285714

In [10]:

```
np.var(a)
```

Out [10]:

2135.0748299319725

In [11]:

```
np.std(a)
```

Out [11]:

46.206869943028735

Задача 2 Для выборки из задачи 1 найти (также без использования библиотек):

1. медиану, первый и третий квартили, интерквартильное расстояние,
2. выборсы в выборке (используя для этого метод как при построении "усов" из boxplot).

Возможные неоднозначности в вычислении квантилей можно разрешать любым способом.

In [12]:

```
a.sort()  
a[n//2], a[n//4], a[n//4*3]
```

Out[12]:

```
(70, 45, 84)
```

In [13]:

```
a[n//4*3] - a[n//4]
```

Out[13]:

```
39
```

In [14]:

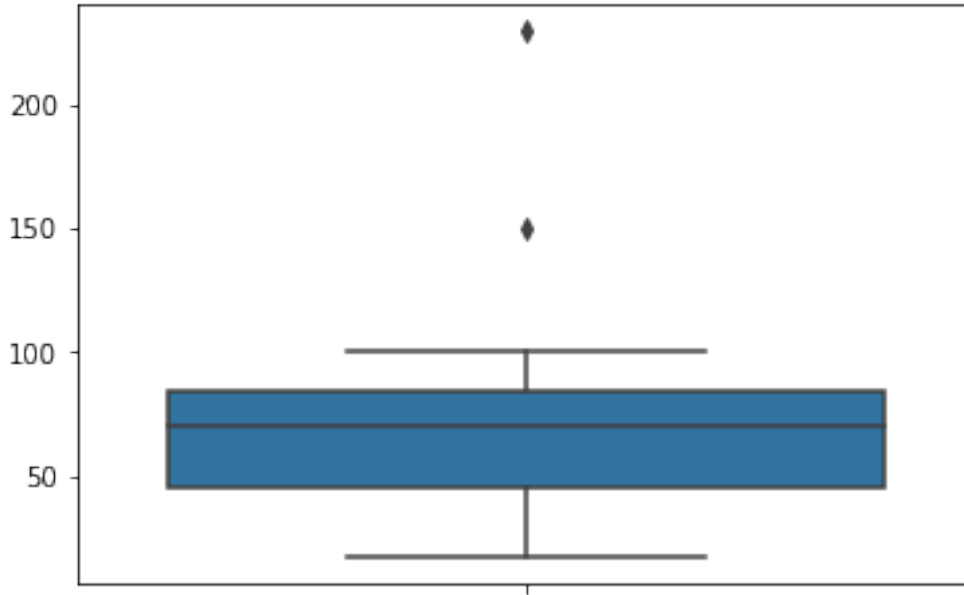
```
import seaborn as sns
```

In [15]:

```
sns.boxplot(a, orient='v')
```

Out[15]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x1a1906e110>



Задача 3 В университет на факультеты А и В поступило равное количество студентов, а на факультет С студентов поступило столько же, сколько на А и В вместе. Вероятность того, что студент факультета А сдаст первую сессию, равна 0.8. Для студента факультета В эта вероятность равна 0.7, а для студента факультета С - 0.9. Студент сдал первую сессию. Какова вероятность, что он учится: а) на факультете А? б) на факультете В? в) на факультете С? Замечание: да, эта задача не на тему описательных статистик, но тема важная, и её стоит иногда освежать в памяти

задача решается по формуле Байеса: допустим число студентов на факультетах А и В =  $n$ , тогда на факультете С число студентов  $2n$

In [16]:

```
pA = 0.8
pB = 0.7
pC = 0.9
n = 10

PA = (2*n*pA)/(n*((2*pA)+pB+pC))
PB = (n*pB)/(n*((2*pA)+pB+pC))
PC = (n*pC)/(n*((2*pA)+pB+pC))

PA, PB, PC
```

Out[16]:

```
(0.5000000000000001, 0.21875000000000003, 0.28125000
000000006)
```

In [ ]:

Задача 1 Даны значения зарплат из выборки выпускников: 100, 80, 75, 77, 89, 33, 45, 25, 65, 17, 30, 230, 24, 57, 55, 70, 75, 65, 84, 90, 150 Используя только встроенные питоновские функции и структуры данных (т.е. без библиотек numpy, pandas и др.) посчитать среднее арифметическое, смещённую и несмещённую оценки дисперсии, среднее квадратичное отклонение для данной выборки. Можно затем посчитать те же значения с использованием библиотек, чтобы проверить себя.

Рассчитаем среднее арифметическое (Ms)

In [3]:

Out[3]:

```
73.14285714285714
```

Смещенная дисперсия (D)

In [4]:

Out [4]:

2135.074829931973

Несмещенная дисперсия (D2)

In [5]:

Out [5]:

2241.8285714285716

Среднее квадратичное отклонение (Sigma)

In [19]:

Out [19]:

47.34795213553139

In [8]:

In [9]:

Out [9]:

73.14285714285714

In [10]:

Out [10]:

2135.0748299319725

In [11]:

Out [11]:

46.206869943028735

Задача 2 Для выборки из задачи 1 найти (также без использования библиотек):

1. медиану, первый и третий квартили, интерквартильное расстояние,
2. выбросы в выборке (используя для этого метод как при построении "усов" из boxplot).

Возможные неоднозначности в вычислении квантилей можно разрешать любым способом.

In [12]:

Out [12]:

(70, 45, 84)

In [13]:

Out [13]:

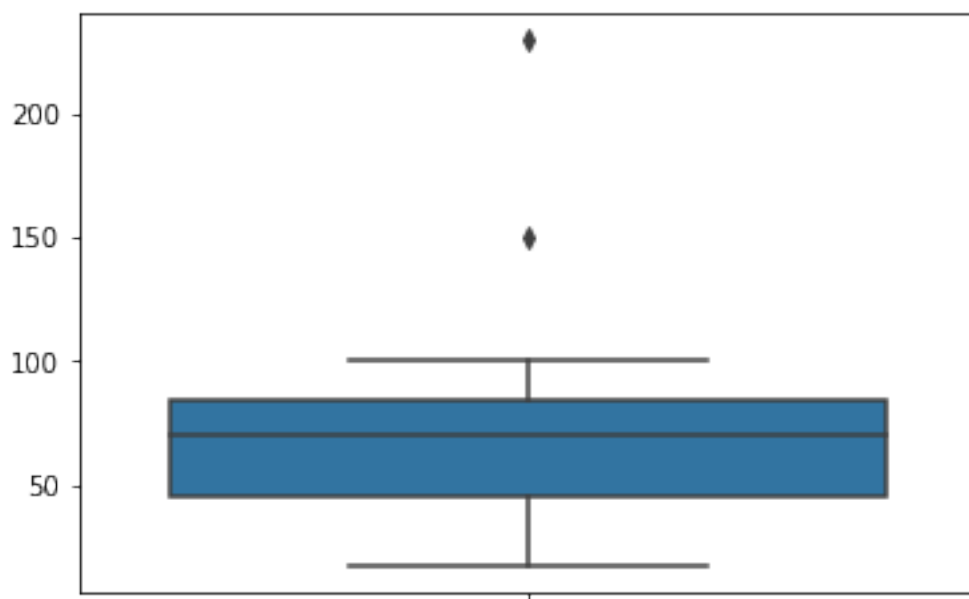
39

In [14]:

In [15]:

Out[15]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a1906e110>
```



Задача 3 В университет на факультеты А и В поступило равное количество студентов, а на факультет С студентов поступило столько же, сколько на А и В вместе. Вероятность того, что студент факультета А сдаст первую сессию, равна 0.8. Для студента факультета В эта вероятность равна 0.7, а для студента факультета С - 0.9. Студент сдал первую сессию. Какова вероятность, что он учится: а) на факультете А? б) на факультете В? в) на факультете С? Замечание: да, эта задача не на тему описательных статистик, но тема важная, и её стоит иногда освежать в памяти

задача решается по формуле Байеса: допустим число студентов на факультетах А и В =  $n$ , тогда на факультете С число студентов  $2n$

In [16]:

Out[16]:

```
(0.5000000000000001, 0.21875000000000003, 0.28125000000000006)
```

In [ ]:



