ESERCIZI

ESERCIZIO_1

La seguente tabella riporta la classificazione delle famiglie italiane secondo il reddito dichiarato (in milioni di lire) nel 1983:

Reddito	Numero di famiglie (in migliaia)
0 ┤ 6	1.128
6 ┤ 10	2.602
10 ┤ 16	5.314
16 ┤ 22	3.768
22 ┤ 30	3.177
30 ┤ 40	1.817
40 ┤ 50	784
50 ┤ 100	556
	19.146

a) Qual è l'unità statistica di riferimento? Qual è il carattere osservato? Di che tipo di carattere si tratta? Quante sono le modalità del carattere della distribuzione data?

b) Calcolare la frequenza cumulata, la frequenza relativa e la frequenza relativa cumulata. In quale classe è compresa una famiglia che ha dichiarato 16.350.000 lire?

c) Costruire la distribuzione del medesimo insieme statistico secondo la classificazione seguente:

0 ┤ 10
10 ┤ 30
30 ┤ 50
più di 50

d) Tra la distribuzione data e quella costruita al punto c) qual è più ricca di informazioni?

ESERCIZIO_2

Si considerino i seguenti valori: 5, 7, 9, 14, 15, 16. Calcolare la media e verificare le due proprietà.

ESERCIZIO_3

Nella tabella seguente è riportata la distribuzione delle famiglie per il numero di componenti in un dato comune:

Numero componenti	**Famiglie**
1	153
2	225
3	335
4	564
5	346
6	133
7	75
8	49
Totale	1.880

Determinare la media aritmetica e la media geometrica.

ESERCIZIO_4

Nella seguente tabella sono riportati gli iscritti ad un'associazione sportiva:

Classi di età	Iscritti
3 - 15	115
15 - 25	156
25 - 40	130
40 - 50	110
50 - 60	90
Oltre 60	38
	639

Determinare la media aritmetica.

ESERCIZIO_5

Si consideri la distribuzione delle unità locali in agricoltura secondo il numero di addetti in Italia al 26 ottobre 1981 (fonte ISTAT):

Numero di addetti	Unità locali	Addetti in complesso
Fino a 2	26.556	36.334
3-9	11.177	48.098
10-49	2.321	43.036
50-99	258	17.355
100-499	132	22.563
500 e oltre	9	7.694
Totale	40.453	175.080

a) Indicare la classe modale della distribuzione;
b) A quale ipotesi si è fatto ricorso per rispondere alla domanda precedente?
c) Calcolare il numero medio aritmetico di addetti per le unità comprese nella classe identificata dall'intervallo 3-9.
d) Calcolare il numero medio aritmetico di addetti per l'insieme per l'insieme delle 40.453 unità locali.

ESERCIZIO_6

Nella seguente tabella sono riportati gli incassi in un giorno di 100 distributori di benzina, sparsi su tutto il territorio nazionale, appartenenti ad una nota compagnia petrolifera; gli incassi sono espressi in milioni di lire.

2,3	4,3	1,6	2,7	6,3	2,6	10,0	9,6	6,8	5,4
11,7	4,7	1,2	2,0	5,7	3,4	6,4	9,1	5,7	1,2
1,4	2,1	3,1	7,0	2,4	2,3	1,6	1,0	1,6	1,3
10,2	1,0	1,6	2,4	2,3	9,7	4,3	3,0	0,9	7,7
9,6	7,3	9,9	3,9	3,2	9,3	4,5	4,6	5,3	6,3
7,3	8,7	6,6	3,4	4,4	1,1	4,9	8,5	6,8	4,2
2,1	6,5	3,5	4,4	9,8	3,9	2,9	5,6	3,4	0,8
10,1	5,4	5,0	9,6	2,2	5,4	3,6	5,4	3,7	2,6
8,6	2,6	2,9	8,3	8,5	10,4	3,7	2,3	10,9	7,3
3,5	3,9	2,3	9,0	8,7	7,5	5,9	2,8	9,7	3,8

Raggruppare i dati in classi chiuse a destra di ampiezza 1. Rappresentare graficamente l'istogramma , calcolare la media aritmetica e la mediana e indicarle sull'istogramma.
Si può parlare di distribuzione bimodale?

ESERCIZIO_7

Costruire un diagramma cartesiano per la seguente serie storica, relativa all'esportazione di agrumi (in migliaia di quintali):

Anni	Migliaia di q
1970	5000
1971	4498
1972	4555
1973	2392
1974	3708
1975	4115
1976	4874
1977	4882
1978	3261
1979	3365
1980	3304
1981	2504

ESERCIZIO_8

Il proprietario di una catena di negozi vuole confrontare il reddito dei suoi 5 negozi che si trovano in città diverse; i ricavi sono riportati nella tabella seguente:

Negozio (città in cui si trova)	Reddito (milioni di lire)
Neg. 1 (Domodossola)	21.500
Neg. 2 (Mestre)	16.800
Neg. 3 (Ancona)	20.500
Neg. 4 (Perugia)	21.100
Neg. 5 (Taranto)	19.700

Disegnare dei grafici che permettano di sostenere:
a) tutti i negozi hanno una resa simile;
b) il negozio di Mestre rende decisamente meno;
c) Se invece si vuole rappresentare il reale stato della situazione senza idee preconcette quale grafico si può utilizzare?

(*Suggerimento:* a,b)Per sostenere le diverse affermazioni si possono usare grafici a nastri dimensionati in modo diverso. c) grafico a torta).

ESERCIZIO_9

Consideriamo la seguente distribuzione percentuale delle famiglie umbre in classi di reddito mensile in milioni riferita al 1998:

Classi	**P_i**
0-5	0,60
5-10	6,12
10-15	13,42
15-25	30,65
25-35	18,77
35-50	16,01
50-100	14,43

Costruire l'istogramma di frequenza.

ESERCIZIO_10

Partendo dalle produzioni per ettaro e le corrispondenti superfici in migliaia di ettari per i principali legumi, calcoliamo la produzione media per ettaro:

Coltivazioni	***Produzione per ha*** **x_i**	***Superficie*** **w_i**
Fagioli	19,40	10,32
Piselli	31,50	5,91
Ceci	14,10	4,73
Lenticchie	7,10	1,08
Totale		22,04

ESERCIZIO_11

Si riporta la distribuzione di 125 atleti per classi di altezza (in cm):

Classi di altezze	**Numero atleti**
171-175	14
176-180	18
181-185	28
186-190	33
191-195	17
196-200	15
Totale	125

Determinare:
a) il primo quartile;
b) il secondo quartile;
c) il terzo quartile.

SOLUZIONI

SOLUZIONE_1

a) L'unità statistica è la famiglia. Il carattere osservato è il reddito dichiarato nel 1983. Esso è un carattere quantitativo continuo. La distribuzione data riporta 8 modalità del carattere.

b)

Reddito	Numero di famiglie (in migliaia)	N_i	f_i	F_i
0 ┤ 6	1.128	1.128	0,06	0,06
6 ┤ 10	2.602	3.730	0,13	0,19
10 ┤ 16	5.314	9.044	0,28	0,47
16 ┤ 22	3.768	12.812	0,20	0,67
22 ┤ 30	3.177	15.989	0,17	0,84
30 ┤ 40	1.817	17.806	0,09	0,93
40 ┤ 50	784	18.590	0,04	0,97
50 ┤ 100	556	19.146	0,03	1,00
	19.146			

Una famiglia che ha dichiarato 16.350.000 lire è compresa nella classe che corrisponde all'intervallo 16 ┤ 22.

c)

Reddito	Numero di famiglie (in migliaia)
0 ┤ 10	3.730
10 ┤ 30	12.259
30 ┤ 50	2.601
più di 50	556
	19.146

d) La distribuzione data è più ricca di informazioni rispetto a quella costruita al punto c).

SOLUZIONE_2

$$x_m = \frac{5+7+9+14+15+16}{6} = 11$$

Proprietà 1)

$$\sum_i (x_i - 11) = (5-11)+(7-11)+(9-11)+(14-11)+(15-11)+(16-11) = 0$$

Proprietà 2)

$$\sum_i (x_i - 11)^2 = 36+16+4+9+16+25 = 106$$

SOLUZIONE_3

x_i	n_i	$x_i\, n_i$
1	153	153
2	225	450
3	335	1.005
4	564	2.256
5	346	1.730
6	133	798
7	75	525
8	49	392
Totale	1.880	7.309

Applicando la formula si ottiene il numero medio di componenti la famiglia per il collettivo considerato:

$$x_m = \frac{7.309}{1.880} = 3{,}88$$

x_i	n_i	$\log x_i$	$n_i \log x_i$
1	153	0,00	0,00
2	225	0,30	67,73
3	335	0,48	159,84
4	564	0,60	339,56
5	346	0,70	241,84
6	133	0,78	103,49
7	75	0,85	63,38
8	49	0,90	44,25
Totale	1.880	-	1.020,10

Da cui:

$$\log G = \frac{1}{1.880} * 1.020{,}10 = 0{,}543$$

Risalendo dal logaritmo al numero si ottiene la media geometrica:

$$G = 3{,}49$$

SOLUZIONE_4

Per il calcolo della media aritmetica è necessario introdurre un'ipotesi di lavoro, cioè che in ogni classe di età, le unità rilevate siano concentrate nel valore centrale della classe.

x_i	x_{i+1}	x_c	n_i	$x_c\, n_i$
3	15	9	115	1.035
15	25	20	156	3.120
25	40	32,5	130	4.225
40	50	45	110	4.950
50	60	55	90	4.950
60	70	65	38	2.470
Totale	-	-	639	20.750

L'età media degli iscritti all'associazione è di circa: $x_m = \frac{20.750}{639} = 32,47$

SOLUZIONE_5

a) La frequenza n_i delle unità locali il cui numero di addetti è compreso nell'i-esimo intervallo dipende dall'ampiezza w_i dell'intervallo stesso, per $i=1, 2, \ldots, k$. Tale ostacolo alla comparabilità delle frequenze può essere superato calcolando la densità di frequenza n_i/w_i:

$x_{i-1} \vdash x_i$	$w_i=x_i-x_{i-1}$	n_i/w_i
1 ⊢3	2	13.278,00
3 ⊢10	7	1.596,71
10 ⊢50	40	58,03
50 ⊢100	50	5,16
100 ⊢500	400	0,33
500 ⊢1000	500	0,02

La classe modale è quella cui corrisponde la densità di frequenza più alta, cioè la classe delle unità locali con al più 2 addetti (si noti che, comunque si scelga l'estremo superiore convenzionale dell'intervallo "500 e oltre", la classe modale non cambia).

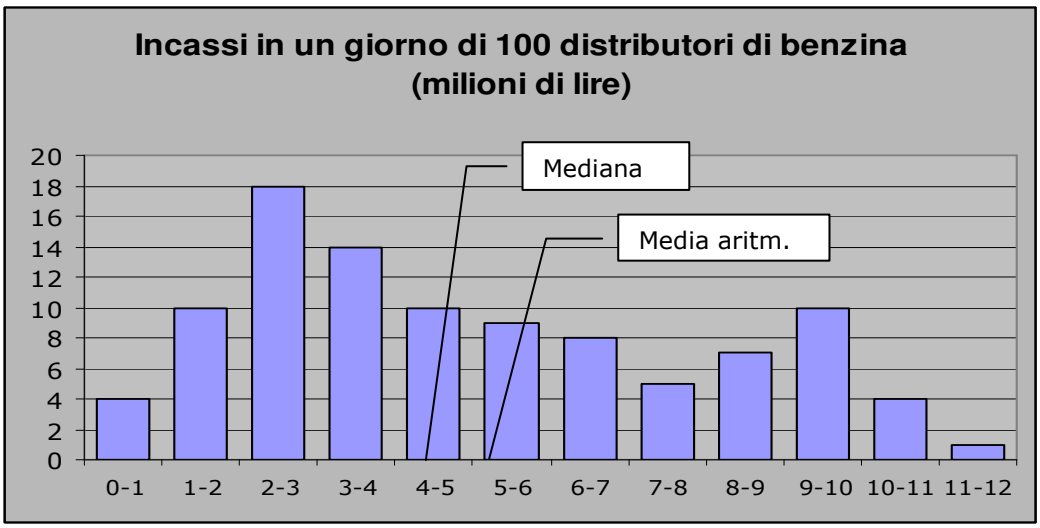
b) Il calcolo della densità di frequenza della classe i-esima equivale a scomporre l'intervallo in subintervalli di ampiezza pari all'unità in cui è espresso il carattere, e a determinare la frequenza che a essi spetta nell'ipotesi di distribuzione uniforme delle unità statistiche.

c) Disponendo dell'informazione sull'ammontare totale del carattere entro ciascuna classe, è possibile calcolare il numero medio aritmetico di addetti per le unità locali della classe 3-9: esso è pari a 48.098/11.177=4,4 addetti.
d) Analogamente, per calcolare il numero medio aritmetico di addetti per l'insieme delle 40.453 unità locali conviene sfruttare l'informazione contenuta nel totale della terza colonna. Tale informazione solitamente non è disponibile e viene stimata con $\hat{x}_i n_i$, dove $\hat{x}_i$ è preso come valore di sintesi della distribuzione del carattere entro la classe i-esima. Così la media aritmetica del numero di addetti è uguale a 175.080/40.453=4,33.

SOLUZIONE_6

Questa è la distribuzione dei dati raggruppati:

Distribuzioni delle frequenze		
Modalità	**Freq. Ass.**	**Freq. Cum.**
0-1	4	4
1-2	10	14
2-3	18	32
3-4	14	46
4-5	10	56
5-6	9	65
6-7	8	73
7-8	5	78
8-9	7	85
9-10	10	95
10-11	4	99
11-12	1	100
Totale	100	



La media aritmetica è pari a 5,03 e la mediana a 4,40.

$$\bar{x} = \frac{(0{,}5*4)+(1{,}5*10)+\ldots+(11{,}5*1)}{100} = 5{,}03$$

$$\tilde{x} = I_m + \left(\frac{0{,}5 - F_{m-1}}{F_m - F_{m-1}}\right)\Delta_m = 4 + \left(\frac{0{,}5-0{,}46}{0{,}56-0{,}46}\right)*1 = 4{,}40$$

I_m estremo inferiore della classe mediana

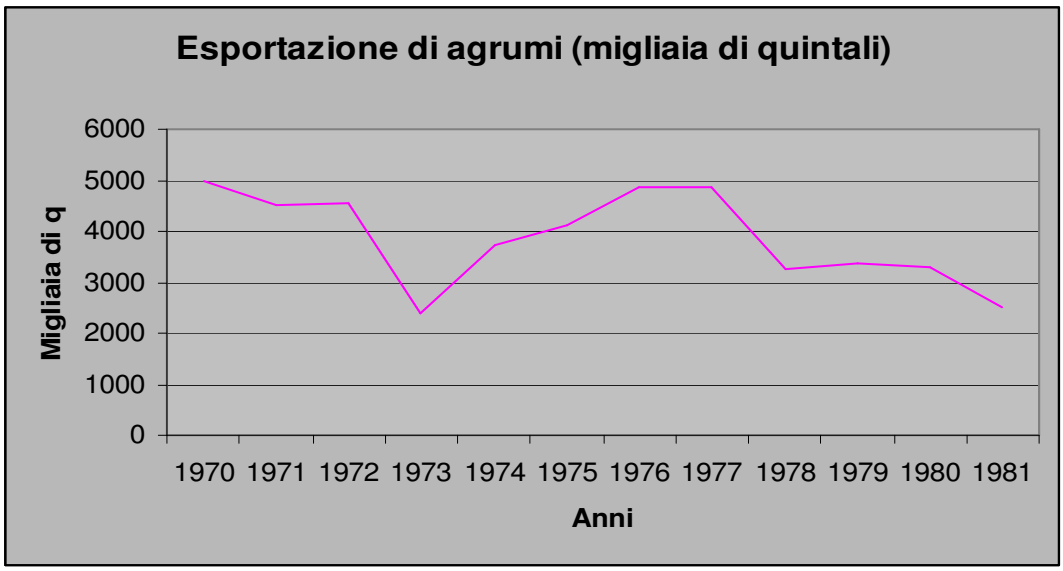
F_{m-1} freq. relat. cumulata fino alla classe precedente a quella mediana

F_m freq. relat. cumulata fino alla classe mediana

Δ_m ampiezza della classe mediana

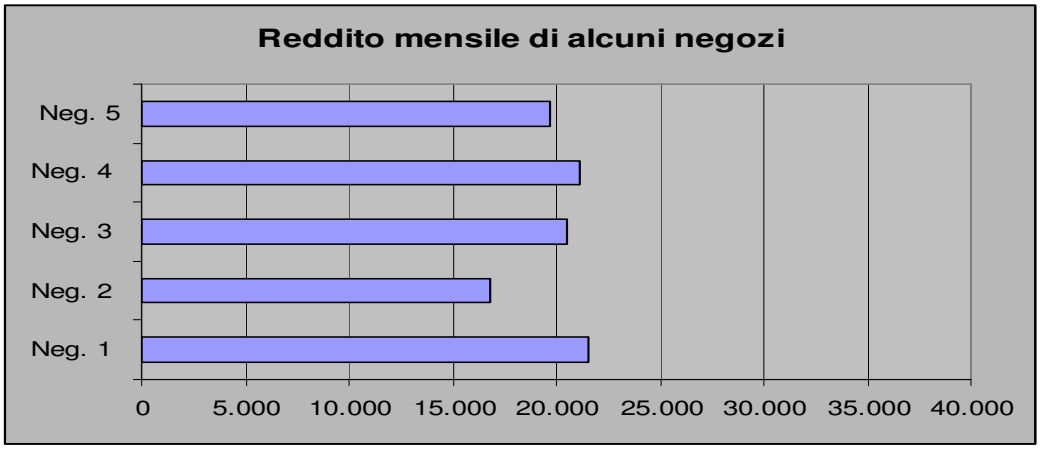
La distribuzione dei dati raggruppati risulta una distribuzione bimodale; le due classi modali sono (2-3) e (9-10). Infatti, anche se queste due classi non hanno la stessa frequenza, possiedono entrambe una frequenza decisamente maggiore rispetto alle classi contigue.

SOLUZIONE_7

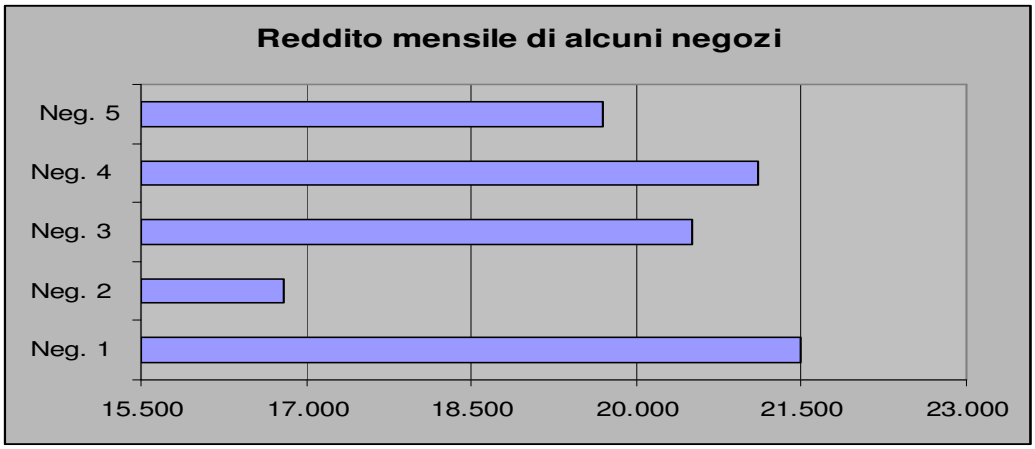


SOLUZIONE_8

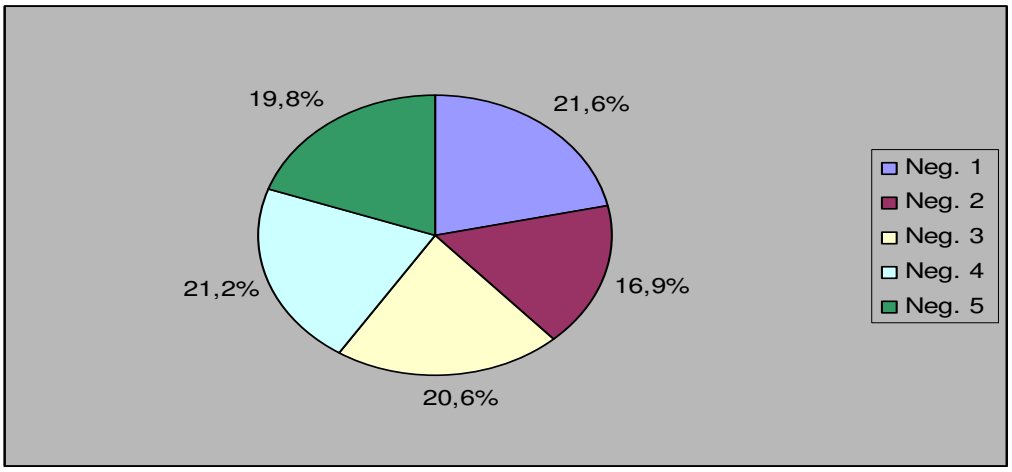
a)



b)



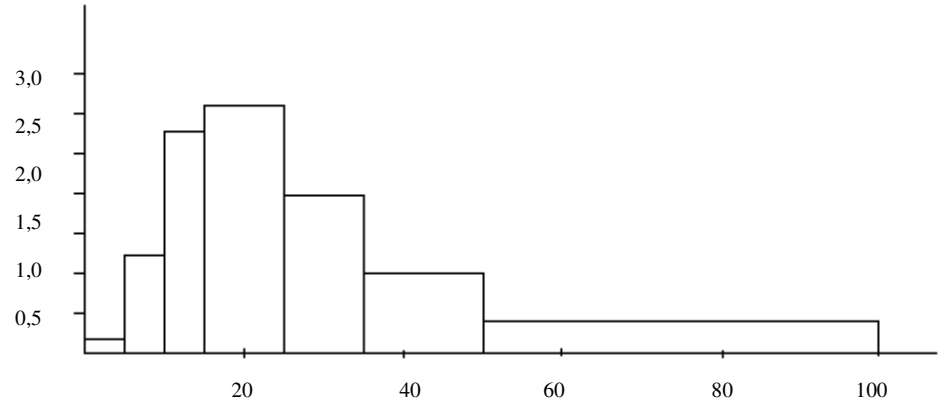
c)



SOLUZIONE_9

Classi	Pi	Ampiezze	Densità
0-5	0,60	5	0,12
5-10	6,12	5	1,22
10-15	13,42	5	2,68
15-25	30,65	10	3,06
25-35	18,77	10	1,88
35-50	16,01	15	1,07
50-100	14,43	50	0,29

Dove $densità = \frac{P_i}{Ampiezza}$



La figura dimostra che la densità cresce rapidamente fino alla classe 15-25 e poi decresce gradualmente con l'aumentare del reddito.

SOLUZIONE_10

Coltivazioni	Produzione per ha xi	Superficie wi	Produzione totale xiwi
Fagioli	19,40	10,32	200,21
Piselli	31,50	5,91	186,17
Ceci	14,10	4,73	66,69
Lenticchie	7,10	1,08	7,67
Totale		22,04	460,73

$$\bar{x} = \frac{460,73}{22,04} = 20,90$$

SOLUZIONE_11

Classi di altezze	Numero atleti	f_i	F_i
171-175	14	0,11	0,11
176-180	18	0,14	0,26
181-185	28	0,22	0,48
186-190	33	0,26	0,74
191-195	17	0,14	0,88
196-200	15	0,12	1,00
Totale	125	1,00	

a) Nella distribuzione data, il primo quartile è l'elemento per cui si ha:

$$¼ = 0,25$$

$$Q_1 = I_{Q_1} + \left(\frac{0,25 - F_{Q-1}}{F_Q - F_{Q-1}}\right)\Delta_Q = 176 + \left(\frac{0,25 - 0,11}{0,26 - 0,11}\right) * 5 = 180,67$$

I_Q **estremo inferiore della classe dove cade il primo quartile**

F_{Q-1} **freq. relat. cumulata fino alla classe precedente a quella in cui cade il primo quartile**

F_Q **freq. relat. cumulata fino alla classe che contiene il primo quartile**

Δ_Q **ampiezza della classe che contiene il primo quartile**

b) Nella distribuzione data, il secondo quartile è l'elemento per cui si ha:

$$½ = 0,50$$

$$Q_2 = I_{Q_2}\left(\frac{0,50 - F_{Q-1}}{F_Q - F_{Q-1}}\right)\Delta_Q = 186 + \left(\frac{0,50 - 0,48}{0,74 - 0,48}\right) * 5 = 186,38$$

c) Nella distribuzione data, il terzo quartile è l'elemento per cui si ha:

$$¾ = 0,75$$

$$Q_3 = I_{Q_3}\left(\frac{0,75 - F_{Q-1}}{F_Q - F_{Q-1}}\right)\Delta_Q = 191 + \left(\frac{0,75 - 0,74}{0,88 - 0,74}\right) * 5 = 191,36$$