ESERCIZIO

La seguente tabella riporta i dati concernenti 4 caratteri rilevati su 30 abitazioni situate nella città di Roma. I caratteri rilevati riguardano il quartiere, il valore stimato del terreno (in migliaia di Euro), il valore stimato delle ristrutturazioni (in migliaia di Euro) ed il prezzo di vendita (in migliaia di Euro).

ID	Quartiere	Valore stimato del terreno	Valore stimato delle ristrutturazioni	Prezzo di vendita
1	B	44,01	13,62	175,00
2	B	30,28	24,34	187,50
3	B	31,68	97,94	173,50
4	B	15,06	58,84	81,00
5	B	47,73	81,46	125,50
6	B	66,24	25,66	243,00
7	C	15,44	70,31	105,00
8	C	12,22	52,90	87,50
9	C	12,04	34,65	60,00
10	C	30,41	73,21	140,00
11	C	26,52	38,72	84,50
12	D	16,10	66,43	85,00
13	D	55,93	40,92	243,80
14	D	14,59	52,79	82,00
15	D	16,20	78,68	86,80
16	E	59,80	86,12	205,00
17	E	20,80	19,64	78,00
18	E	26,60	6,39	58,00
19	E	10,70	57,66	74,20
20	F	16,93	16,94	39,00
21	F	25,52	37,81	59,90
22	F	14,49	14,95	49,50
23	G	17,95	31,95	61,50
24	G	10,13	10,84	25,40
25	G	5,27	41,27	37,70
26	G	7,55	18,73	36,00
27	G	7,74	18,47	37,20
28	G	7,12	20,13	54,30
29	G	8,37	21,35	45,00
30	G	7,15	16,19	22,40

1) Per il carattere VALORE DEL TERRENO, disegnare la funzione di ripartizione:
   1.1) Per la successione dei primi 5 elementi;
   1.2) Per la distribuzione di frequenza ottenuta suddividendo la successione in 6 classi equiampie.

2) Per lo stesso carattere calcolare i quartili:
   2.1) della distribuzione semplice;
   2.2) della distribuzione delle 6 classi equiampie.

3) Calcolare i decili per il carattere PREZZO.

4) Sia data la seguente nuova distribuzione di frequenze per il carattere VALORE DEL TERRENO:

C_i	n_i	f_i	Fi
≤ 5,27	2	0,05	0,05
]5,27; 15,43]	13	0,33	0,38
]15,43; 25,59]	7	0,18	0,55
]25,59; 35,76]	5	0,13	0,68
]35,76; 45,92]	4	0,10	0,78
]45,92; 56,08]	4	0,10	0,88
]56,08; 66,24]	3	0,08	0,95
> 66,24	2	0,05	1,00
Totali	**40**	**1,00**	

Determinare la media troncata al 10%.

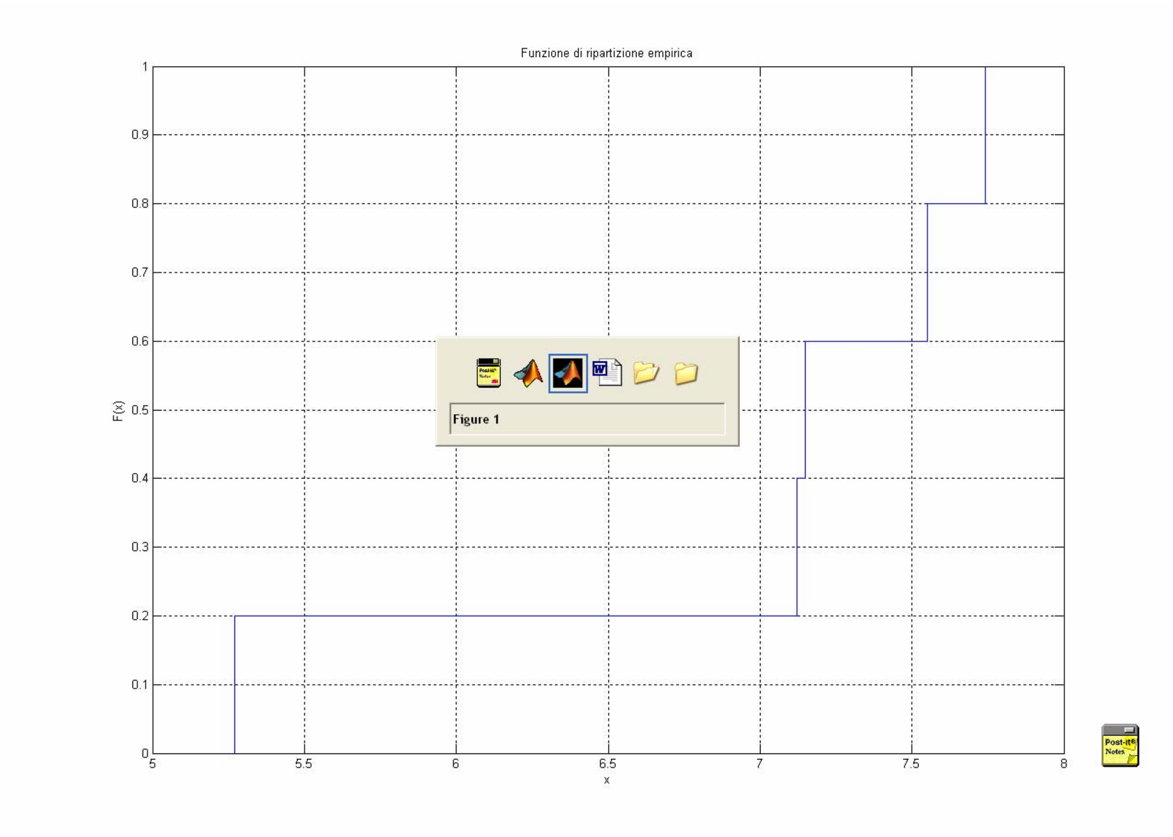
Soluzione

1.1)
Le 30 osservazioni del carattere valore del terreno vanno poste in ordine non decrescente:

1	5,27	16	16,20
2	7,12	17	16,93
3	7,15	18	17,95
4	7,55	19	20,80
5	7,74	20	25,52
6	8,37	21	26,52
7	10,13	22	26,60
8	10,70	23	30,28
9	12,04	24	30,41
10	12,22	25	31,68
11	14,49	26	44,01
12	14,59	27	47,73
13	15,06	28	55,93
14	15,44	29	59,80
15	16,10	30	66,24

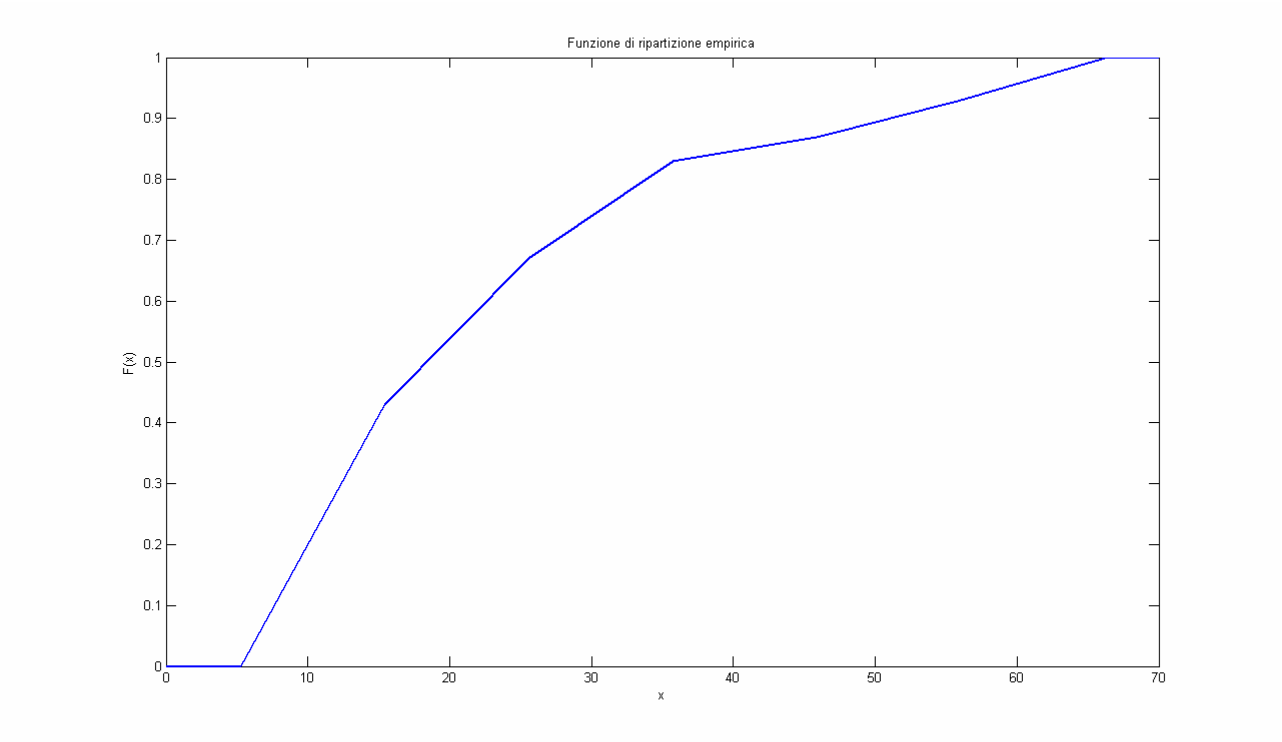
Considerate le prime 5, la funzione di ripartizione è la seguente:

	x_i	f_i	F_i
1	5,27	0,2	0,2
2	7,12	0,2	0,4
3	7,15	0,2	0,6
4	7,55	0,2	0,8
5	7,74	0,2	1



1.2)
La distribuzione ottenuta suddividendo le osservazioni in 6 classi equiampie è la seguente:

C_i	n_i	f_i	Fi
[5,27; 15,43]	13	0,43	0,43
]15,43; 25,59]	7	0,23	0,67
]25,59; 35,76]	5	0,17	0,83
]35,76; 45,92]	1	0,03	0,87
]45,92; 56,08]	2	0,07	0,93
]56,08; 66,24]	2	0,07	1,00
Totale	**30**	**1,00**	



2.1)
I quartili della distribuzione semplice vanno individuati determinando prima la mediana (secondo quartile) e poi le mediane delle due metà della distribuzione che si trovano a sinistra (primo quartile) e a destra (terzo quartile) della mediana.
Il numero totale di osservazioni N è pari: la mediana sarà la semisomma dei due valori che occupano le due posizioni centrali, ossia:

$$Q_2 = Me = \frac{x_{\frac{N}{2}} + x_{\frac{N}{2}+1}}{2} = \frac{x_{15} + x_{16}}{2} = \frac{16{,}10 + 16{,}20}{2} = 16{,}15$$

Per il calcolo di Q_1 e Q_3 si fa riferimento alle due semidistribuzioni di N = 15 elementi:

$$Q_1 = x_{\frac{N+1}{2}} = x_8 = 10{,}70$$

$$Q_3 = 30{,}28$$

2.2)
I quartili di una distribuzione in classi saranno determinati in base alla formula generica:

$$x_{Px} = x_{Px-1} + \frac{x_{Px} - x_{Px-1}}{F_{Px} - F_{Px-1}}\left(F_{desiderata} - F_{Px-1}\right)$$

in cui, individuata la classe di riferimento, si sostituirà ad $F_{desiderata}$ il valore 0,25 per Q_1, 0,5 per Q_2 e 0,75 per Q_3.
In base alla frequenza relativa cumulata, il primo quartile si trova nella prima classe. Quindi:

$$Q_1 = 0 + \frac{10{,}16}{0{,}43}\left(0{,}25 - 0\right) = 5{,}9$$

Il secondo quartile si trova nella seconda classe, quindi:

$$Q_2 = 15{,}43 + \frac{10{,}16}{0{,}67 - 0{,}43}\left(0{,}5 - 0{,}43\right) = 18{,}39$$

Infine, il terzo quartile si trova nella terza classe, quindi:

$$Q_3 = 25{,}59 + \frac{10{,}16}{0{,}83 - 0{,}67}\left(0{,}75 - 0{,}67\right) = 30{,}67$$

3)

I 9 decili del carattere PREZZO vanno determinati suddividendo la numerosità totale N = 30 in dieci parti uguali, ottenendo il numero costante di unità che separano ciascun decile dal successivo, ossia $\frac{30}{10} = 3$.
Dunque, il k^{mo} decile è determinato come semisomma degli elementi di posto 3k e 3k+1, con k = 1, …, 9:

$$d_k = \frac{x_{3k} + x_{3k+1}}{2}$$

Data la successione non decrescente del carattere PREZZO:

1	22,40	16	81,00
2	25,40	17	82,00
3	36,00	18	84,50
4	37,20	19	85,00
5	37,70	20	86,80
6	39,00	21	87,50
7	45,00	22	105,00
8	49,50	23	125,50
9	54,30	24	140,00
10	58,00	25	173,50
11	59,90	26	175,00
12	60,00	27	187,50
13	61,50	28	205,00
14	74,20	29	243,00
15	78,00	30	243,80

i 9 decili sono:

$$d_1 = \frac{x_3 + x_4}{2} = \frac{36 + 37,2}{2} = 36,60$$

$$d_2 = \frac{x_6 + x_7}{2} = \frac{39 + 45}{2} = 42,00$$

$$d_3 = \frac{x_9 + x_{10}}{2} = \frac{54,3 + 58}{2} = 56,15$$

$$d_4 = \frac{x_{12} + x_{13}}{2} = \frac{60 + 61,5}{2} = 60,75$$

$$d_5 = \frac{x_{15} + x_{16}}{2} = \frac{78 + 81}{2} = 79,50$$

$$d_6 = \frac{x_{18} + x_{19}}{2} = \frac{84,5 + 85}{2} = 84,75$$

$$d_7 = \frac{x_{21} + x_{22}}{2} = \frac{87,5 + 105}{2} = 96,25$$

$$d_8 = \frac{x_{24} + x_{25}}{2} = \frac{140 + 173,5}{2} = 156,75$$

$$d_9 = \frac{x_{27} + x_{28}}{2} = \frac{187,5 + 205}{2} = 196,25$$

4)
In generale la media troncata secondo una percentuale α si ottiene attribuendo alle osservazioni un peso π pari a:

$$\pi_\alpha = \begin{cases} 0 & \left\{1 \le i \le \frac{\alpha}{2} \cdot N\right\} \cup \left\{\left(n - \frac{\alpha}{2} \cdot N\right) \le i \le N\right\} \\ 1 & \left(\frac{\alpha}{2} \cdot N\right) < i < \left(n - \frac{\alpha}{2} \cdot N\right) \end{cases}$$

Data la distribuzione:

C_i	n_i	f_i	Fi
≤ 5,27	2	0,05	0,05
]5,27; 15,43]	13	0,33	0,38
]15,43; 25,59]	7	0,18	0,55
]25,59; 35,76]	5	0,13	0,68
]35,76; 45,92]	4	0,10	0,78
]45,92; 56,08]	4	0,10	0,88
]56,08; 66,24]	3	0,08	0,95
> 66,24	2	0,05	1,00
Totali	**40**	**1**	

il peso $\pi_{0,1}$ per la media troncata al 10% $\mu_{0,1}$ è:

$$\pi_{0,1} = \begin{cases} 0 & \{1 \le i \le 0{,}05 \cdot N\} \cup \{(n - 0{,}05 \cdot N) \le i \le N\} \\ 1 & (0{,}05 \cdot N) < i < (n - 0{,}05 \cdot N) \end{cases}$$

cioè:

$$\pi_{0,1} = \begin{cases} 0 & \{1 \le i \le 2\} \cup \{(39) \le i \le 40\} \\ 1 & 2 < i < 39 \end{cases}$$

In tal caso, essendo le frequenze relative nella prima e nell'ultima classe pari proprio a 0,05, basta eliminare queste due classi e procedere con il calcolo della media nelle altre 6 (corrispondenti al 90% della distribuzione).

C_i	c_i	n_i	f_i	Fi
]5,27; 15,43]	10,35	13	0,36	0,36
]15,43; 25,59]	20,51	7	0,19	0,56
]25,59; 35,76]	30,67	5	0,14	0,69
]35,76; 45,92]	40,84	4	0,11	0,81
]45,92; 56,08]	51,00	4	0,11	0,92
]56,08; 66,24]	61,16	3	0,08	1,00
Totali		**36**	**1,00**	

$$\mu_\alpha = \frac{\sum_{i=1}^{K} c_i n_i}{N_\alpha}$$

in cui K è il numero di classi effettive (dopo il troncamento) e N_α è il numero di unità effettive che entrano nel calcolo della media, ossia:

$$N_\alpha = N - 10\%\ N = 40 - 4 = 36$$

Dunque:

$$\mu_\alpha = \frac{(10{,}35 \times 13) + (20{,}51 \times 7) + (30{,}67 \times 5) + (40{,}84 \times 4) + (51 \times 4) + (61{,}16 \times 3)}{36} = 27{,}29$$

o, equivalentemente:

$$\mu_\alpha = (10{,}35 \times 0{,}36) + (20{,}51 \times 0{,}19) + (30{,}67 \times 0{,}14) + (40{,}84 \times 0{,}11) + (51 \times 0{,}11) +$$
$$+ (61{,}16 \times 0{,}08) = 27{,}29$$