Type 1 Interpolation: Quartile Calculations

We interpolate to find the quartiles when given grouped data. We do not use midpoints like for mean and standard deviation, we must use the upper-class boundaries (UCB) and frequencies (ff) instead

Example 1

Age	Frequency
$0 \le w < 5$	5
$5 \le w < 20$	45
$20 \le w < 40$	90
$40 \le w < 65$	130
$65 \le w < 80$	60
$80 \le w < 90$	1

Way 1: Shorter Method

$$median = \frac{n}{2} = \frac{331}{2} = 165.5^{\text{th}} \text{ value}$$

See where 165.5 would insert in the cf column and drop down to next row

Age	Frequency	UCB	cf (running total)
$0 \le w < 5$	5	5	5
$5 \le w < 20$	45	20	50
$20 \le w < 40$	90	40	140
$40 \le w < 65$	130	65	270
$65 \le w < 80$	60	80	330
$80 \le w < 90$	1	90	331

165.5

Apply the formula: $\text{LCB} + \frac{\text{how many in}}{\text{group total}} \times \text{class width}$

$$40 + \frac{165.5 - 140}{130} \times (65 - 40) = 44.9039$$
$$= 44.9$$

Way 2: Longer Method

$$median = \frac{n}{2} = \frac{331}{2} = 165.5^{\text{th}} \text{ value}$$

See where 140.5 would insert in the cf column

Find where the corresponding x value would be in the UCB column and call it x

Upper Class Boundary	cf
5	5
20	50
40	140
65	270
80	330
90	331

x

165.5

zoom in on the rows above and below of where we insert

40	140
x	165.5
65	270

We subtract the distances indicated above

$$\frac{x - 40}{65 - 40} = \frac{165.5 - 140}{270 - 140}$$

$$\frac{x - 40}{25} = \frac{25.5}{130}$$

$$x - 40 = 25\left(\frac{25.5}{130}\right)$$

$$x - 40 = 4.9038$$
$$x = 44.9$$

Note: if you want the lower quartile, upper quartile or pth percentile do the exact same thing, but instead of $\frac{n}{2}$ use $\frac{n}{4}$, $\frac{3n}{4}$, $\left(\frac{p}{100}\right)n$

Example 2

Weight (kg)	Frequency
$1 \le w < 3$	15
$3 \le w < 5$	31
$5 \le w < 6$	45
$6 \le w < 6.5$	37
$6.5 \le w < 7$	21
$7 \le w < 10$	15

Way 1: Shorter Method

$$\frac{n}{2} = \frac{164}{2} = 82^{\text{nd}} \text{ value}$$

See where 82 would insert in the cf column and drop down to next row

Weight (kg)	Frequency	UCB	cf
$1 \le w < 3$	15	3	15
$3 \le w < 5$	31	5	46
$5 \le w < 6$	45	6	91
$6 \le w < 6.5$	37	6.5	128
$6.5 \le w < 7$	21	7	149
$7 \le w < 10$	15	10	164

82

Apply the formula: $\text{LCB} + \frac{\text{how many in}}{\text{group total}} \times \text{class width}$

$$5 + \frac{82 - 46}{45} \times (6 - 5) = 5.8$$

Way 2: Longer Method

$$\frac{n}{2} = \frac{164}{2} = 82^{\text{nd}} \text{ value}$$

See where 82 would insert in the cf column

Find where the corresponding x value would be in the UCB column and call it x

Upper Class Boundary	cf
3	15
5	46
6	91
6.5	128
7	149
10	164

x

82

zoom in on the rows above and below of where we insert

5	46
x	82
6	91

We subtract the distances indicated above

$$\frac{x - 5}{6 - 5} = \frac{82 - 46}{91 - 46}$$

$$\frac{x - 5}{1} = \frac{36}{45}$$

$$x - 5 = 0.8$$

$$x = 5.8$$

Note: if you want the lower quartile, upper quartile or pth percentile do the exact same thing, but instead of $\frac{n}{2}$ use $\frac{n}{4}, \frac{3n}{4}, \left(\frac{p}{100}\right)n$

Type 2 Interpolation: Splitting Up Rows

Example 1

The masses of 140 adult Bullmastiffs are recorded in a table. One dog is chosen at random.

Mass, m (kg)	Frequency
$45 \le m < 48$	17
$48 \le m < 51$	25
$51 \le m < 54$	42
$54 \le m < 57$	33
$57 \le m < 60$	21
$60 \le m < 64$	2

i. Find the probability that the dog has a mass of 54 kg or more

ii. Find the probability that the dog has a mass between 48 kg and 57 kg

The probability that a Rottweiler chosen at random has a mass under 53 kg is 0.54.

iii. Is it more or less likely that a Bullmastiff chosen at random has a mass under 53 kg? State one assumption that you have made in making your decision

Ans.

i. $p(m \ge 54) = \frac{33+21+2}{140} = \frac{56}{140} = 0.4$

ii. $p(48 < x < 57) = \frac{25+42+33}{140} = \frac{100}{140} = 0.71$

iii. Here we need to **interpolate.**

We need to split the pink third row of the table up

Mass	f
45–48	17
48 – 51	25
51–54	42

$\Longrightarrow$

51 – 53	$\frac{2}{3}(42) = 28$
53–54	$\frac{1}{3}(42) = 14$

$$P(\text{Bullmastiff Under } 53) = \frac{17 + 25 + 28}{140} = \frac{70}{140} = 0.5$$

$0.5 < 0.54$ so less likely

Note: we could have interpolated here using the method mentioned for the quartiles, but it is not necessary since we aren't finding an known in the mass column. We are just looking to split the frequencies up.

51	don't care
53	don't care
54	don't care

x 42

$$\frac{53-51}{54-51} = \frac{x}{42}$$

$$x = 28$$

Example 2

The table shows some information about the salaries of a sample of people

i. Work out the proportion of people in the sample who have a salary greater than £40,000

ii. Find an estimate for the median salary

Salary (p) in £1000s	Frequency
0 <p ≤ 10	4
10 < p ≤ 20	9
20 < p ≤ 25	8
25 < p ≤ 35	10
35 < p ≤ 50	12

i.

Salary	f
0 <p ≤ 10	4
10 < p ≤ 20	9
20 < p ≤ 25	17
25 < p ≤ 35	25
35 < p ≤ 50	42

We need to split the last pink column of the table above table up

$\Longrightarrow$

35 – 40	$\frac{5}{15}(12) = 4$
40-50	$\frac{10}{15}(12) = 8$

$$= \frac{8}{4+9+8+10+12} = \frac{8}{43}$$

ii. This is type 1 interpolation, already covered.

Way 1: Shorter Method

Salary (p) in £1000s	Frequency	Upper Bound	cf
0 <p ≤ 10	4	**10**	**4**
10 < p ≤ 20	9	**20**	**13**
20 < p ≤ 25	8	**25**	**21**
25 < p ≤ 35	10	**35**	**31**
35 < p ≤ 50	12	**50**	**43**

$$\frac{43}{2} = 21.5^{th} \text{ value}$$

Apply the formula:

$$\text{LCB} + \frac{\text{how many in}}{\text{group total}} \times \text{class width}$$

$$25 + \frac{21.5 - 21}{10} \times (35 - 25) = 5.8$$

$$25.5$$

$$£25{,}500$$

Way 2: Longer Method

Salary (p) in £1000s	Frequency	Upper Bound	cf
0 <p ≤ 10	4	**10**	**4**
10 < p ≤ 20	9	**20**	**13**
20 < p ≤ 25	8	**25**	**21**
25 < p ≤ 35	10	**35**	**31**
35 < p ≤ 50	12	**50**	**43**

$$\frac{43}{2} = 21.5^{th} \text{ value}$$

Zoom in on the yellow

25	**21**
$\boldsymbol{x}$	**21.5**
35	**31**

$$\frac{x - 25}{35 - 25} = \frac{21.5 - 21}{31 - 21}$$

$$\frac{x - 25}{10} = \frac{0.5}{10}$$

$$x - 25 = 0.5$$

$$x = 25.5$$

$$£25{,}500$$

Example 3 (with a gap)

The table shows the time, to the nearest minute, spend waiting for a taxi by each of 80 people one Sunday afternoon.

Waiting Time (in minutes)	Frequency
2-4	15
5-6	9
7	6
8	24
9-10	14
11-15	12

i. Estimate the number of people with a waiting time between 3.5 minutes and 7 minutes

ii. Use linear interpolation to estimate the median, the lower quartile and the upper quartile of the waiting times

i.

We need to close the gaps between the boundaries first

Waiting Time (in minutes)	Waiting Time (in minutes)	Frequency
2-4	**1.5-4.5**	15
5-6	**4.5-6.5**	9
7	**6.5-7.5**	6
8	**7.5-8.5**	24
9-10	**8.5-10.5**	14
11-15	**10.5-15.5**	12

We need to split the first pink row and third blue row of the table up

Waiting Time	f
1.5-4.5	15
4.5-6.5	9
6.5-7.5	6
7.5-8.5	24
8.5-10.5	14
10.5-15.5	12

$\Longrightarrow$

Splitting the first pink row up gives	
$1.5-3.5$	$\frac{2}{3}(15)=10$
$3.5-4.5$	$\frac{1}{3}(15)=5$

and

Splitting the third blue row up gives	
$6.5-7$	$\frac{0.5}{1}(6)=3$
7-7.5	$\frac{0.5}{1}(6)=3$

$$=5+9+3=17$$

iii. This is type 1

We need to close the gaps first by turning the categories into bounds

Waiting time (minutes)	Frequency
1.5-4.5	15
4.5-6.5	9
6.5-7.5	6
7.5-8.5	24
8.5-10.5	14
10.5-15.5	12

$\Longrightarrow$

Waiting Time (minutes)	Frequency	Upper Bound	cf
1.5-4.5	15	**4.5**	**15**
4.5-6.5	9	**6.5**	**24**
6.5-7.5	6	**7.5**	**30**
7.5-8.5	24	**8.5**	**54**
8.5-10.5	14	**10.5**	**68**
10.5-15.5	12	**15.5**	**80**

Median

$$\frac{80}{2}=40^{th}\text{ value}$$

Way 1:
Zoom in on

7.5	**30**
x	**40**
8.5	**54**

$$\frac{x-7.5}{8.5-7.5}=\frac{40-30}{54-30}$$

$$\frac{x-7.5}{1}=\frac{10}{24}$$

$$x-7.5=\frac{5}{12}$$

$$x=7.92$$

Way 2:

Apply the formula:

$$\text{LCB}+\frac{\text{how many in}}{\text{group total}}\times\text{class width}$$

$$7.5+\frac{40-30}{24}\times(8.5-7.5)$$

$$=7.92$$

Lower Quartile

$$\frac{80}{4}=20^{th}\text{ value}$$

Way 1:
Zoom in on

4.5	**15**
x	**20**
6.5	**24**

$$\frac{x-4.5}{6.5-4.5}=\frac{20-15}{24-15}$$

$$\frac{x-4.5}{2}=\frac{5}{9}$$

$$x-4.5=\frac{10}{9}$$

$$x=5.61$$

Way 2:

Apply the formula:

$$\text{LCB}+\frac{\text{how many in}}{\text{group total}}\times\text{class width}$$

$$4.5+\frac{20-15}{9}\times(6.5-4.5)$$

$$x=5.61$$

Upper Quartile

$$\frac{3(80)}{4}=60^{th}\text{ value}$$

Way 1:
Zoom in on

8.5	**54**
x	**60**
10.5	**68**

$$\frac{x-8.5}{10.5-8.5}=\frac{60-54}{68-54}$$

$$\frac{x-8.5}{2}=\frac{6}{14}$$

$$x-8.5=\frac{6}{7}$$

$$x=9.36$$

Way 2:

Apply the formula:

$$\text{LCB}+\frac{\text{how many in}}{\text{group total}}\times\text{class width}$$

$$8.5+\frac{60-54}{14}\times(10.5-8.5)$$

$$x=9.36$$