

Grouped Data

Find the mean, standard deviation, median, upper quartile and 99th percentile for the following grouped table in blue below (on the left).

Number of goals (x)	Frequency (f)
5 ≤ x < 9	2
9 ≤ x < 15	9
15 ≤ x < 21	20
21 ≤ x < 25	13
25 ≤ x < 29	8
29 ≤ x < 35	3

we
must
create
the
following
columns
in yellow

Midpoint x	fx	x ²	fx ²	UCB	cf (running total of f)
$\frac{5+9}{2}=7$	2(7)=14	7 ² =49	2(49)=98	9	2
$\frac{9+15}{2}=12$	9(12)=108	12 ² =144	9(144)=1296	15	2 + 9 = 11
$\frac{15+21}{2}=18$	20(18)=360	18 ² =324	20(324)=6480	21	11 + 20 = 31
$\frac{21+25}{2}=23$	13(23)=299	23 ² =529	13(529)=6877	25	31 + 13 = 44
$\frac{25+29}{2}=27$	8(27)=216	27 ² =729	8(729)=5832	29	44 + 8 = 52
$\frac{29+35}{2}=32$	3(32)=96	32 ² =1024	3(1024)=3072	35	52 + 3 = 55
$\sum fx = 1093$		$\sum fx^2 = 23655$			

For mean and standard deviation we use the **midpoints**

For quartiles (median, lower quartile and upper quartile) we use the **upper class boundaries (UCB)** with the **cumulative frequencies (cf)**

Mean:

$$\frac{\sum fx}{n} = \frac{14 + 108 + 360 + 299 + 216 + 96}{2 + 9 + 20 + 13 + 8 + 3} = \frac{1093}{55} = 19.87$$

Standard Deviation:

$$\sqrt{\frac{\sum fx^2}{n} - \bar{x}^2} = \sqrt{\frac{98 + 1296 + 6480 + 6877 + 5832 + 3072}{2 + 9 + 20 + 13 + 8 + 3} - 19.87^2} = \sqrt{\frac{23655}{55} - 19.87^2} = 5.93$$

Note: You could also have used your calculator to get the mean and standard deviation, which is far quicker. See my full data cheat sheet to find out how

Quartiles:

This is longer. We **MUST INTERPOLATE** using the UCB and cf columns. We CANNOT use the calculator for quartiles with grouped data.

Median:

Find $\frac{n}{2}$

$$\frac{55}{2} = 27.5th \text{ value}$$

Way 1:
See where 27.5 would insert in the cf column and drop down to the next row

Number of goals (x)	f	UCB	cf
5 ≤ x < 9	2	9	2
9 ≤ x < 15	9	15	11
15 ≤ x < 21	20	21	31
21 ≤ x < 25	13	25	44
25 ≤ x < 29	8	29	52
29 ≤ x < 35	3	35	55

Apply the formula:

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

$$15 + \frac{27.5 - 11}{20} \times (21 - 15) = 44.9$$

Way 2:
Find where the 27.5th value **INSERTS** in the cf column and zoom in on the box above and below this with the UCB column also

UCB	cf
15	11
x	27.5
21	31

Now interpolate

$$\frac{x-15}{21-15} = \frac{27.5-11}{31-11} \Leftrightarrow \frac{x-15}{6} = 0.825 \Leftrightarrow x = 19.95$$

Upper Quartile:

Find $\frac{3n}{4}$

$$\frac{3(55)}{4} = 41.25th \text{ value}$$

Way 1:
See where 41.25 would insert in the cf column and drop down to the next row

Number of goals (x)	f	UCB	cf
5 ≤ x < 9	2	9	2
9 ≤ x < 15	9	15	11
15 ≤ x < 21	20	21	31
21 ≤ x < 25	13	25	44
25 ≤ x < 29	8	29	52
29 ≤ x < 35	3	35	55

Apply the formula:

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

$$21 + \frac{41.25 - 31}{13} \times (25 - 21) = 24.2$$

Way 2:
Find where the 41.25th value **INSERTS** in the cf column and zoom in on the box above and below this with the UCB column also

UCB	cf
21	31
x	41.25
25	44

Now interpolate

$$\frac{x-21}{25-21} = \frac{41.25-31}{44-31} \Leftrightarrow \frac{x-21}{4} = 0.788 \Leftrightarrow x = 24.2$$

99th percentile:

Find $\frac{99}{100}n$

$$\frac{99}{100}(55) = 54.45th \text{ value}$$

Way 1:
See where 54.45 would insert in the cf column and drop down to the next row

Number of goals (x)	f	UCB	cf
5 ≤ x < 9	2	9	2
9 ≤ x < 15	9	15	11
15 ≤ x < 21	20	21	31
21 ≤ x < 25	13	25	44
25 ≤ x < 29	8	29	52
29 ≤ x < 35	3	35	55

Apply the formula:

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

$$29 + \frac{54.45 - 52}{3} \times (35 - 29) = 44.9$$

Way 2:
Find where the 54.45th value **INSERTS** in the cf column and zoom in on the box above and below this with the UCB column also

UCB	cf
29	52
x	54.45
35	55

Now interpolate

$$\frac{x-29}{35-29} = \frac{54.45-52}{55-52} \Leftrightarrow \frac{x-29}{6} = 0.817 \Leftrightarrow x = 33.9$$

Note: We sometimes have gaps between the boundaries (between a upper-class boundary of a row and the lower class boundary of the subsequent row)



Here there is no gap

x	f
$5 \leq x < 9$	1
$9 \leq x < 15$	3
$15 \leq x < 21$	2

Here there is a gap

x	f
5-9	1
10-15	3
16-21	2

We close the gap

x	f	Closing the gap gives:
5-9	1	$4.5 \leq x < 9.5$
10-15	3	$9.5 \leq x < 15.5$
16-21	2	$15.5 \leq x < 21$

We then proceed as normal

x	f
$4.5 \leq x < 9.5$	1
$9.5 \leq x < 15.5$	3
$15.5 \leq x < 21$	2