7.1 Representing and summarising data

Name: _____ Class: _____ Date: _____ **Total: 25 marks**

KEY WORDS histogram • median

Representing and summarising data

Objective

Work with histograms of unequal class widths (frequency density, bar-height ratios), estimate counts over sub-intervals, and interpolate quartiles in grouped data (objectives 2.1–2.3).

1 Key knowledge and worked examples

Histograms. Height = frequency ÷ class width; area = frequency. Compare bars by ratio, not by height alone when widths differ. To estimate a count over part of a class, scale the class frequency by the fraction of the width. Quartiles in grouped data: find the containing class from cumulative frequencies, then $Q = a + w \cdot \dfrac{\text{target} - F}{f}$.

Worked example – bar heights and quartiles (WST01 June 2025 Q5). Speeds of 225 cars in classes 20–25 (54), 25–30 (90), 30–40 (60), 40–55 (15), 55–70 (6): densities 10.8, 18, 6, 1, 0.4. If the tallest bar (25–30) is 10 cm, the 20–25 bar is $\frac{10.8}{18} \times 10 = 6$ cm (the 30–40 bar also has density 6 – only the ratio argument scores). Cars between 35 and 57: $60 \times \frac{5}{10} + 15 + 6 \times \frac{2}{15} = 45.8 \to 46$. Median (112.5th value): $25 + 5 \times \frac{112.5-54}{90} = 28.25$; Q_3 (168.75th): $30 + 10 \times \frac{168.75-144}{60} = 34.125$.

2 Practice

2.1 A histogram class of width 4 contains 24 items; an adjacent class of width 2 contains 18 items. Which bar is taller, and by what factor? [2]

2.2 For a grouped sample of 80 with cumulative frequencies 20, 50, 68, 80 at class boundaries 0, 10, 20, 30: estimate the median. [3]

2.3 Estimate the number of items between the values 12 and 26 given the classes 10–20 (30 items) and 20–30 (18 items), assuming uniform spread. [4]

3 Exam-style questions

3.1 *(Edexcel IAL WST01 June 2025 Q5, condensed.)* The speeds of a random sample of 225 cars are summarised: $20 \le x < 25$ (54), $25 \le x < 30$ (90), $30 \le x < 40$ (60), $40 \le x < 55$ (15), $55 \le x < 70$ (6). A histogram is drawn and the tallest bar is 10 cm. (a) Calculate the height of the second-tallest bar. (b) Estimate the number of cars with speeds between 35 and 57 km/h. (c) Use linear interpolation to estimate the median speed. (d) Given the lower quartile is 25.14, find the upper quartile and the interquartile range. [11]

3.2 *(Original.)* For the same sample, (a) describe the skewness using the quartiles, and (b) comment on a schoolmate's claim that a Normal distribution models these speeds well. [5]

Solutions

2.1 Densities: $24/4 = 6$ versus $18/2 = 9$: the width-2 class's bar is taller, by a factor $\frac{9}{6} = 1.5$.

2.2 The median is the 40th value, in the second class (cumulative $20 \to 50$): $10 + 10 \times \frac{40-20}{30} = 16.7$.

2.3 $30 \times \frac{8}{10} + 18 \times \frac{6}{10} = 24 + 10.8 = 34.8$, so about **35** items.

3.1 (a) Densities 10.8, 18, 6, 1, 0.4: tallest is 25–30; second tallest is 20–25: $\frac{10.8}{18} \times 10 = 6$ cm. (b) $60 \cdot \frac{5}{10} + 15 + 6 \cdot \frac{2}{15} = 30 + 15 + 0.8 = 45.8$, so 46 cars. (c) The median is the 112.5th value, in 25–30 (cumulative $54 \to 144$): $25 + 5 \times \frac{112.5-54}{90} = 28.25$. (d) Q_3 is the 168.75th value, in 30–40: $30 + 10 \times \frac{168.75-144}{60} = 34.125$; $IQR = 34.125 - 25.14 = 8.985$.

3.2 $(Q_3 - Q_2) = 5.875 > (Q_2 - Q_1) = 3.11$, so positively skewed (the estimated mean $30.6 >$ median 28.25 agrees). A Normal model is not suitable: the distribution is not symmetric.

Local review guidance: densities computed as f/w before any comparison; ratio argument for heights; interpolation fractions using the correct cumulative frequency. This is authored classroom guidance.