

Statistics – Grouped Data

Keywords: Histogram / Frequency Density / Class Interval / Distribution / Cumulative / Frequency / Polygon / Median / Interquartile range / Box Plot / Estimated Mean

Definition / Description:

- Cumulative Frequency** diagram represents a running total of frequencies as a graph
- Box Plot** visually shows the distribution of data by identifying five points in a data set
- Interquartile range** measures the spread of data between the upper and lower quartiles. A small interquartile range shows consistent data
- Histogram** is a graphical representation of data points organised into ranges
- Frequency density** is the frequency per unit for the data in each class. It is used to plot histograms
- Class interval** is the numerical width of any class in a particular distribution, defined as the difference between the upper class limit and the lower class limit
- Estimated Mean** is the average using midpoints of grouped data.

Knowledge points:

- Cumulative Frequency:**
 - Calculate the running total
 - Plot at the Upper boundary of class
 - Join with a smooth curve
- Box Plot:**

5 points from data needed:

 - Minimum value
 - Lower quartile
 - Median
 - Upper quartile
 - Maximum value
- Interquartile range:**

Subtract Lower quartile from Upper quartile
- Histogram:**
 - Find class width of each category
 - Divide frequency by class width to find Frequency Density
- Estimated Mean:**
 - Find midpoint of data range
 - Multiply each frequency by this midpoint to find a breakdown of total
 - Add up breakdown of totals to find final total
 - Divide final total by total frequency

Knowledge point examples:

A cyclist records the number of miles he travels each week

Weeks	Number of miles (frequency)	Cumulative Frequency
1	17	17
2	19	36
3	42	78
4	38	116
5	14	130

Here is some information about the amount of cola contained in a sample of bottles from machine A.

Minimum	Lower quartile	Median	Upper quartile	Maximum
496 ml	502 ml	508 ml	510 ml	514 ml

Draw a box plot to represent this information.

Time taken (t seconds)	Class Width	Frequency	Frequency Density
10 < t < 30	20	5	0.25
30 < t < 40	10	8	0.8
40 < t < 50	10	16	1.6
50 < t < 70	20	24	1.2

Speed (mph)	Frequency (f)	Midpoint (m)	f x m
20 ≤ s < 25	4	22.5	90
25 ≤ s < 30	10	27.5	275
30 ≤ s < 35	12	32.5	390
35 ≤ s < 40	15	37.5	562.5
40 ≤ s < 45	9	42.5	382.5
Total	50		1700

Mean = Total sum ÷ total Frequency
 = 1700 ÷ 50
 = 34 mph

Linked Knowledge Maps: Statistics – ungrouped data Averages