Statistically Speaking

An activity of "Fake cancer cures on YouTube"

"Fake cancer cures on YouTube" (page 21, *What's Up* October 2019) warns us about false advertisements on YouTube. Read the article. Then, follow the instructions below.

Instructions: In pairs, complete the discrete frequency distribution table. Column 1 consists of a list of data values (some words from the article) arranged in alphabetical order. In column 2, write tally marks to track the frequency (the total number of times each word appears in the article). Make one vertical line '|' each time the word appears; the fifth time it appears, represent it by a line diagonally across the previous four (i.e. 1 is |, 2 is ||, 3 is |||, 4 is ||||, 5 is 卌, 6 is 卌| and so on). In column 3, write down the frequency based on the tally marks in the previous column. In column 4, calculate the cumulative frequency (the total number of times the word and the words before it appear).

Discrete data values	Tally marks	Frequency	Cumulative Frequency
advertisements	||||	4	4
BBC	卌		6	4+6=10
cancer			
healthcare			
misinformation			
videos	卌 卌 卌 |||		
YouTube			

Based on this frequency distribution table, answer the following.

i.   How many data values will there be in the complete data set?

ii. If the set of the first five data values = {advertisements, advertisements, advertisements, advertisements, BBC}, then write down the set of the first 11 data values.

iii. What is the 13th, 24th and 50th data values of the complete ordered data set?

iv. What will be the middle position(s) of the complete ordered data set?

v.   Find the median of the distribution. Median = Data value(s) positioned in the middle of an ordered data set.

vi. Find the mode of the distribution. Mode = Most frequently repeated data value in the complete data set.

vii. Say whether True or False and explain: It is possible to find the mean for the above data set.