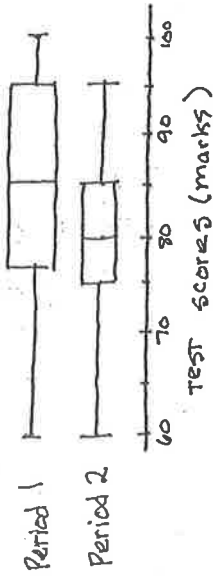


II Five Number Summaries and Box Plots

① Look at the two boxplots of test scores of two classes.



What is the median test score in period 1? —

What is the Period 1 mean score? —

What is the range of test scores in period 1? —

What is the IQR (interquartile range) of Period 1? —

of Period 1 —

of Period 2 —

Can you determine the number of test takers in either Period?

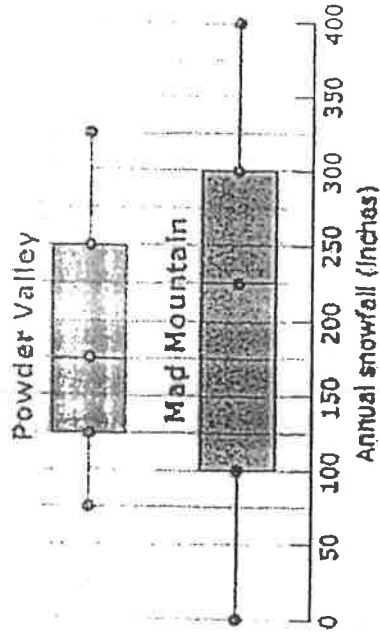
What percentage of the class in Period 1 is above 85 marks? —

What percentage of the class in Per. 2 is above 85 marks? —

What class, in your opinion, had more variation in their scores? —

②

	Min	Q1	median	Q3	max
Powder Valley					
Mad Mountain					



For Mad Mountain, what percent of its years

had less than 225 inches? —

had less than 100 inches? —

had a snowfall between 100 and 225 inches? —

had a snowfall between 100 and 300 inches? —

For Mad Mountain, how many years were measured?

1-variable Statistics - Class Notes - Part 3

1. Draw a box-and-whisker plot for the running times

Minutes to Run 5km				
26	26.1	27.2	27.6	28.9
30.2	30.6	31.1	31.5	32.1
33.4	34	34	34	36.7
45				

Describe the distribution of running times.
(Symmetric, Uniform, skewed left, skewed right)

2. Calculate the inter-quartile range (IQR) _____

3. Calculate the mean run time and the median run time

mean _____ median _____

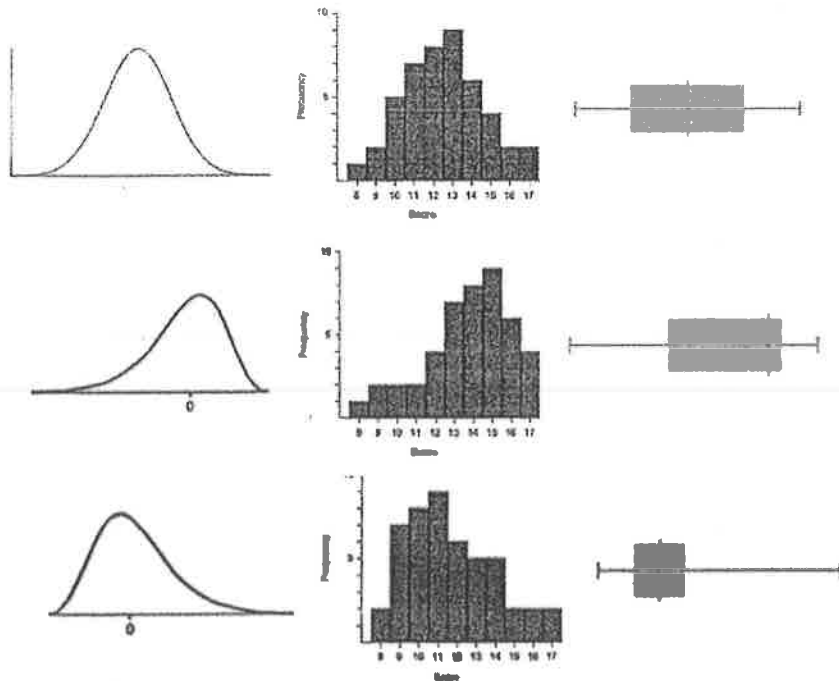
4. Is the mean or median run better to represent the "typical" or center run time?

(It actually depends on the distribution of values which we'll learn about today.)

5. What percent of the runners, at least according to the box plot, is under 34 minutes?



Types of Distribution

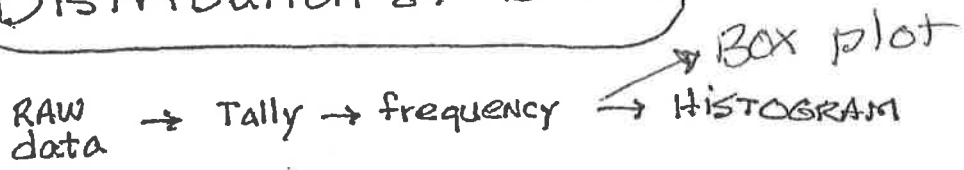


To best represent the "center"

If there are outliers OR the distribution is not symmetric, the median is more reliable.

If the data is symmetric with no outliers, then either the mean or median work equally as well.

Distribution of Data

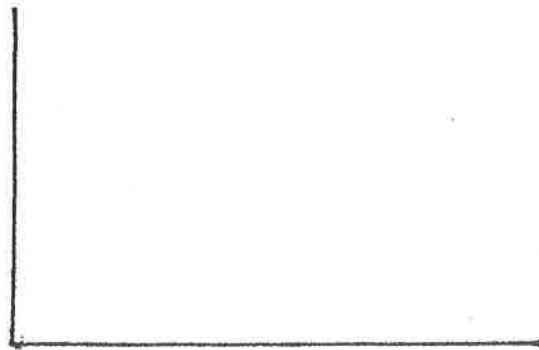


A. Survey of the number of vehicles passing by the store between 8:45 am and 9:00 am over a 30 day period.

27	23	23	35	32	9
30	40	22	24	52	41
17	28	18	18	31	38
13	38	29	24	39	24
46	24	16	44	32	32

Number of cars	Tally	Freq
0-9		
10-19		
20-29		
30-39		
40-49		
50-59		

HISTOGRAM



B. with a GDC

C. combo box Plot histograms

B. Use your GDC to make a histogram of the test averages. (we'll decide on a class on the intervals, but for now use $X_{scale} = 5$ and X_{min} of 40)

142	169	181	165	170
161	188	173	178	182
170	171	195	168	190
172	163	168	186	197
158	160	187	177	174