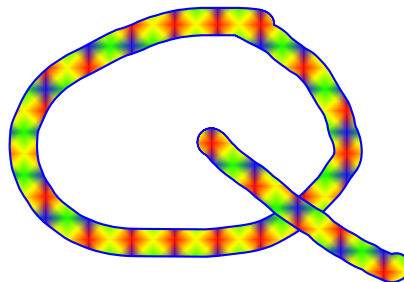


Happy Friday

Let me know about HW  
questions before class starts



Questions on HW



**EXERCISE 6A**

p. 160

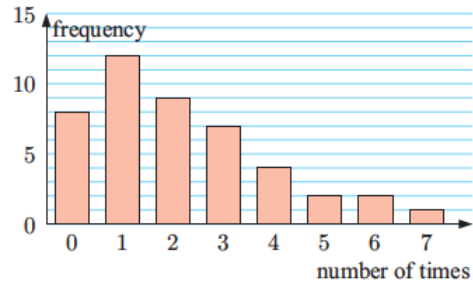
- 1 Classify the following variables as categorical, quantitative discrete, or quantitative continuous
  - a the number of brothers a person has
  - b the colours of lollies in a packet
  - c the time children spend brushing their teeth each day
  - d the height of trees in a garden
  - e the brand of car a person drives
  - f the number of petrol pumps at a service station
  - g the most popular holiday destinations
  - h the scores out of 10 in a diving competition

**EXERCISE 6A**

- ① a quantitative discrete      b categorical  
 c quantitative continuous      d quantitative continuous  
 e categorical      f quantitative discrete      g categorical  
 h quantitative discrete      i quantitative continuous  
 j quantitative continuous      k quantitative continuous  
 l categorical      m quantitative discrete

p. 165 ●●● 6B ●... 4

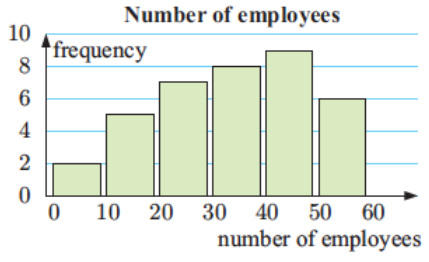
- 4 A random sample of people were asked “How many times did you eat at a restaurant last week?” A column graph was used to display the results.
- a How many people were surveyed?
  - b Find the mode of the data.
  - c How many people surveyed did not eat at a restaurant at all last week?
  - d What percentage of people surveyed ate at a restaurant more than three times last week?
  - e Describe the distribution of the data.



- a 45      b 1 time      c 8      d 20%  
 e positively skewed, no outliers

p. 167... GC

2 A selection of businesses were asked how many employees they had. A column graph was constructed to display the results.



- a How many businesses were surveyed?
- b Find the modal class.
- c Describe the distribution of the data.
- d What percentage of businesses surveyed had less than 30 employees?
- e Can you determine the highest number of employees a business had?

- 2 a 37    b 40 - 49 employees    c negatively skewed
- d  $\approx 37.8\%$
- e No, only that it was in the interval 50 - 59 employees.

p. 167... GC

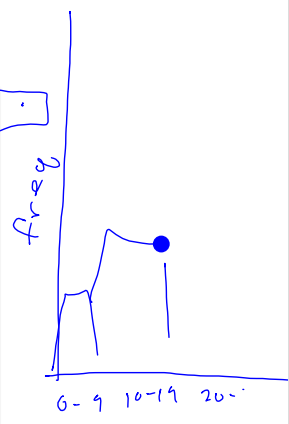
3 A city council does a survey of the number of houses per street in a suburb.

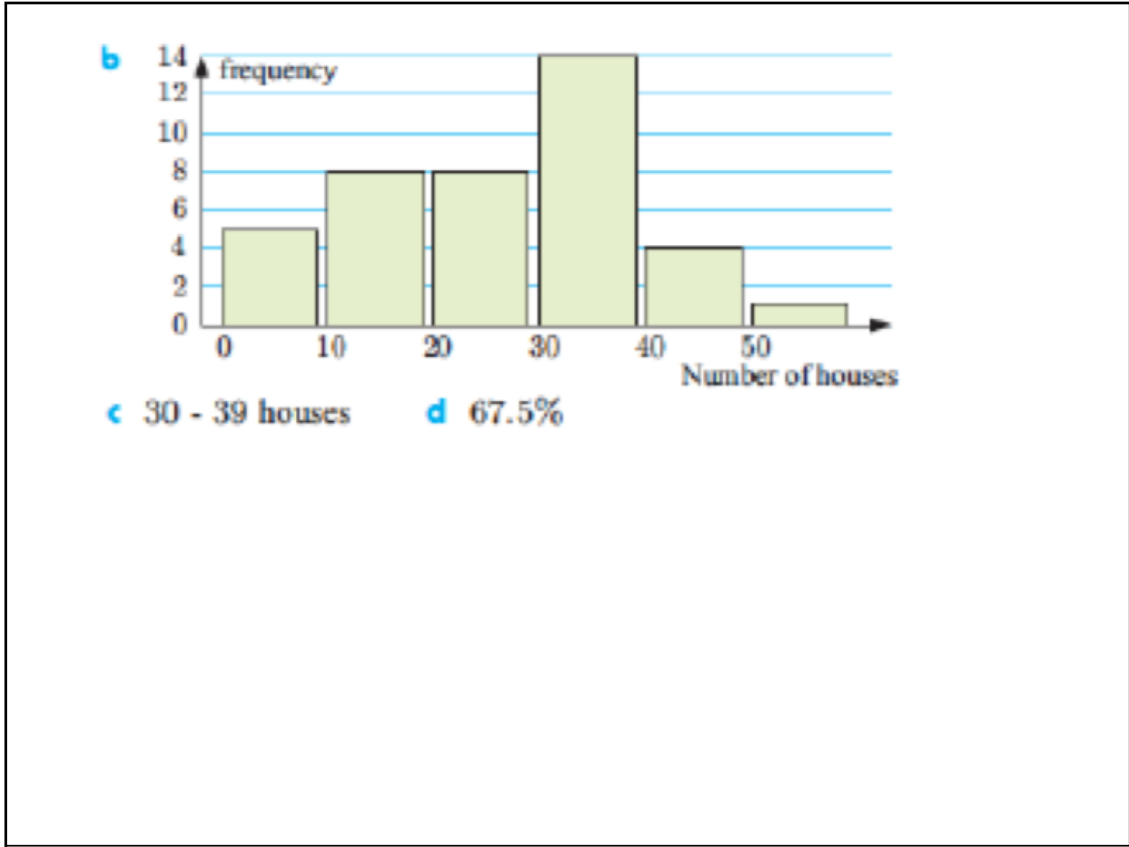
42 15 20 6 34 19 8 5 11 38 56 23 24 24  
 35 47 22 36 39 18 14 44 25 6 34 35 28 12  
 27 32 36 34 30 40 32 12 17 6 37 32

- a Construct a frequency table for this data using class intervals 0 - 9, 10 - 19, ..., 50 - 59.
- b Hence draw a column graph to display the data.
- c Write down the modal class.
- d What percentage of the streets contain at least 20 houses?

a

Number of houses	Tally	Frequency
0 - 9		5
10 - 19		8
20 - 29		8
30 - 39		14
40 - 49		4
50 - 59		1
Total		40





**EXERCISE 6D**

p.168 •• 6D



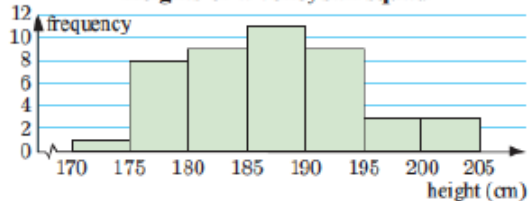
**1** A frequency table for the heights of a volleyball squad is given alongside.

- a** Explain why 'height' is a continuous variable.
- b** Construct a frequency histogram for the data. Carefully mark and label the axes, and include a heading for the graph.
- c** What is the modal class? Explain what this means.
- d** Describe the distribution of the data.

Height ( $H$ cm)	Frequency
$170 \leq H < 175$	1
$175 \leq H < 180$	8
$180 \leq H < 185$	9
$185 \leq H < 190$	11
$190 \leq H < 195$	9
$195 \leq H < 200$	3
$200 \leq H < 205$	3

**a** Height is measured on a continuous scale.

**b** **Heights of a volleyball squad**



- c**  $185 \leq H < 190$  cm. This is the class of values that is most often.
- d** slightly positively skewed

Write this score next  
to your name

Hans Rosling  
9/8/16 (9)

example

**\*\* Showing work is a requirement on all assignments..... but use common sense.**

May 2018  
IB Studies scores

---

	4 4	5			
	4 4 4	5 5	6		
	4 4 4	5 5	6 6	7	
	4 4 4	5 5	6 6	7	
2	4 4 4	5 5	6 6	7	

35 students

may 2018	may 2017
4 4      5	4      5      7
4 4 4    5 5 6	4      5 5      7
4 4 4    5 5 6 6 7	4 4    5 5      7
4 4 4    5 5 6 6 7	3 4 4    5 5    6 7
2 4 4 4    5 5 6 6 7	
35 students	23 students

Today

Measures of Center

Mean, Median, Mode

which one is best?

Hans Rosling

# THE **MEDIAN**

IS ANOTHER KIND OF CENTER: THE "MIDPOINT" OF THE DATA, LIKE THE "MEDIAN STRIP" IN A ROAD.





TO FIND THE MEDIAN  
VALUE OF A DATA SET,  
WE ARRANGE THE DATA  
IN ORDER FROM  
SMALLEST TO LARGEST.  
THE MEDIAN IS THE  
VALUE IN THE MIDDLE.

3 5 7 7 38  
↑

THE MEDIAN

IF THE NUMBER OF POINTS IS *EVEN*—IN WHICH CASE THERE IS NO MIDDLE, WE  
AVERAGE THE TWO VALUES AROUND THE MIDDLE... SO IF THE DATA ARE

3 5 7 7 WE AVERAGE 5 AND 7 TO GET  $\frac{5 + 7}{2} = 6$   
↑  
MIDDLE  
SPACE

<i>Height (H cm)</i>	<i>Frequency</i>
$170 \leq H < 175$	1
$175 \leq H < 180$	8
$180 \leq H < 185$	9
$185 \leq H < 190$	11
$190 \leq H < 195$	9
$195 \leq H < 200$	3
$200 \leq H < 205$	3

Mode

Modal Class

So Now use your GDC  
to calculate the

Median  
and the Mean

Calculating the mean of a data set is easy but there is a

Notation

you need to know

Data

5 7 3 38 7 ...

Data

	①	②	③	④	⑤	...	i
$x$	5	7	3	38	7	...	$x_i$

the 4<sup>th</sup> piece of data could be denoted as  $\rightarrow x_4 = 38$

$x_5 = 7$

↑  
all of the data

## THE MEAN (OR "AVERAGE")

THE MEAN OR AVERAGE VALUE IS REPRESENTED BY  $\bar{x}$ . IT'S OBTAINED BY ADDING ALL THE DATA AND DIVIDING BY THE NUMBER OF OBSERVATIONS:

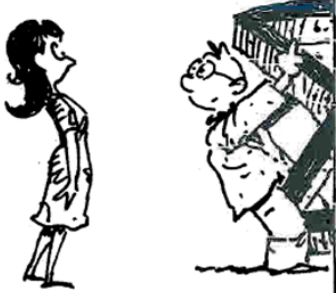
$$\bar{x} = \frac{\text{SUM OF DATA}}{n}$$

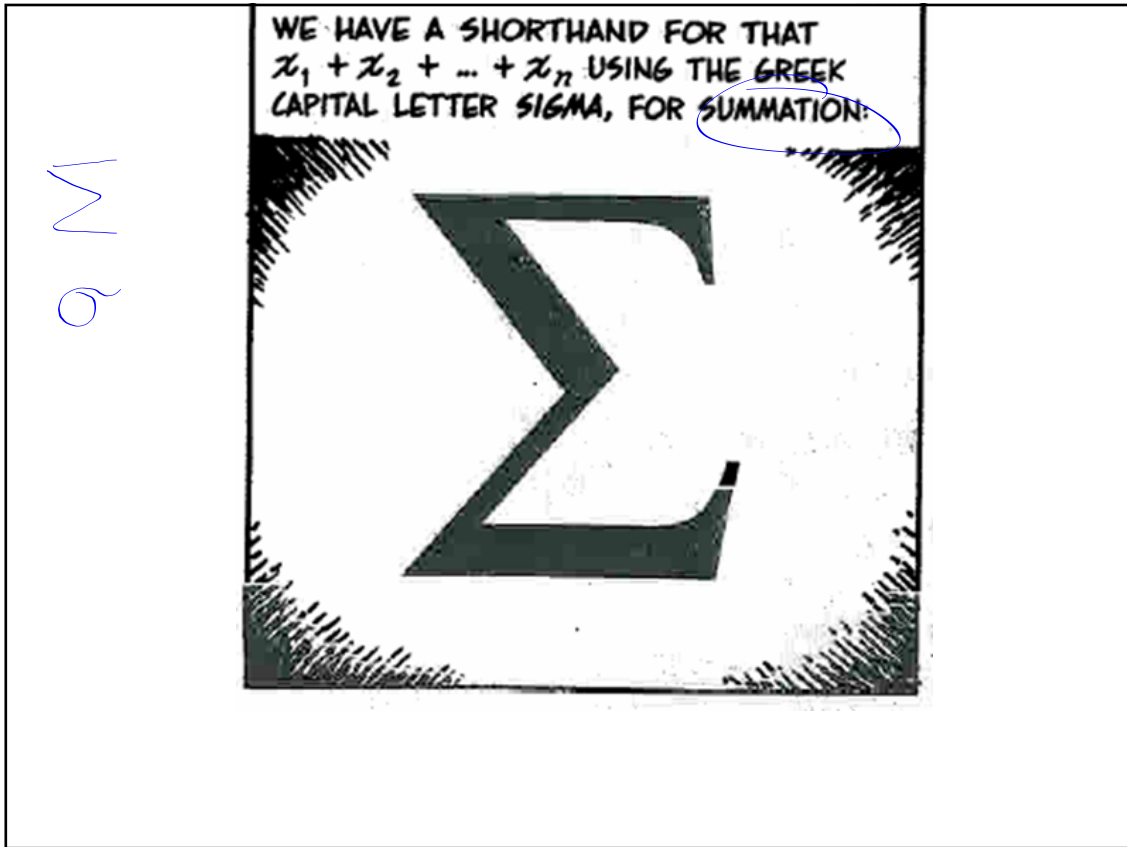
$$= \frac{x_1 + x_2 + \dots + x_n}{n}$$

FOR OUR EXAMPLE,

$$\bar{x} = \frac{5 + 7 + 3 + 38 + 7}{5} = \frac{60}{5}$$

**= 12 HOURS**





FOR THE SUM  $x_1 + x_2 + \dots + x_n$  WE WRITE

$$\sum_{i=1}^n x_i$$

AND READ IT AS "THE SUM OF  $x_i$  AS  $i$  GOES FROM 1 TO  $n$ ."

SAY IT TEN TIMES AND YOU'LL NEVER FORGET IT...

$x_1$     $x_2$     $x_3$     $x_4$

↓   ↓   ↓   ↓

15   6   10   3   27

$\sum_{i=1}^4 x_i = 15 + 6 + 10 + 3 = 34$

Mean

$$\bar{X} = \frac{\sum_{i=1}^K x_i}{n}$$

Write down

# of pieces of data a.k.a frequency


$$\sum f_i = n$$

IN THE CASE OF OUR 92 PENN STATE STUDENTS, THE MEAN WEIGHT IS

$$\frac{\sum_{i=1}^{92} x_i}{92} = \frac{13,354}{92}$$

=

145.15 POUNDS



but

Sometimes data is repeated and grouped and a variation of the formula is needed.

Pick up the handout

**Calculating the Mean of non continuous data that has been grouped**

A boy rolled a die 50 times with the following results:

$$\begin{aligned}\bar{x} &= \frac{9 \cdot 1 + 10 \cdot 2 + \dots}{50} \\ &= \frac{177}{50} \\ &= 3.54\end{aligned}$$

Score $x$	Frequency $f$
1	9
2	10
3	5
4	8
5	7
6	11
	50

$$\bar{x} = \frac{\sum f_i x_i}{n}$$

Calculate Mean score from the above data and show critical totals?

Now use the GDC  
Lists as a spreadsheet.

everyone needs to be  
able to do this.

Now repeat but use  
the shortcut  
(add to reference sheet)



Calculate Mean of data with frequencies

< enter data  $L_1$ , freq.  $L_2$

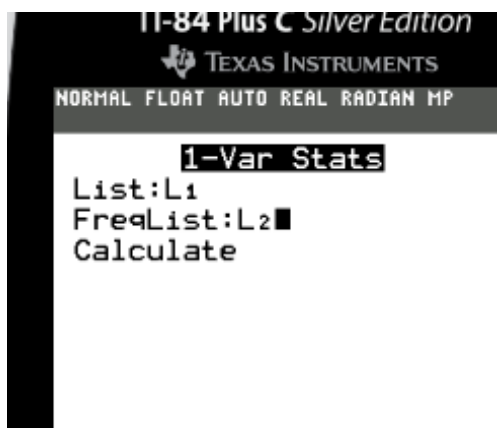
**STAT** **CALC** 1-Variable Stats  $L_1, L_2$

$L_1$   
 $L_2$

for those that have the  
Newer type calculations

- remove the  $L_2$  when finished

$L_1$   
 ~~$L_2$~~



**Calculating the Mean of *non continuous data* that has been grouped**

1. A boy rolled a die 50 times with the following results:

Score	Frequency
1	9
2	10
3	5
4	8
5	7
6	11

Calculate Mean score from the above data and show critical totals?

2. Find the mean of the heights using a

computer spreadsheet:

Use the extra column of the table



Height	Frequency	
10cm	3	
15cm	8	
20cm	7	
25cm	6	
TOTAL		



2. Find the mean of the heights using a computer spreadsheet:

Use the extra column of the table

Height	Frequency	
10cm	3	
15cm	8	
20cm	7	
25cm	6	
TOTAL		

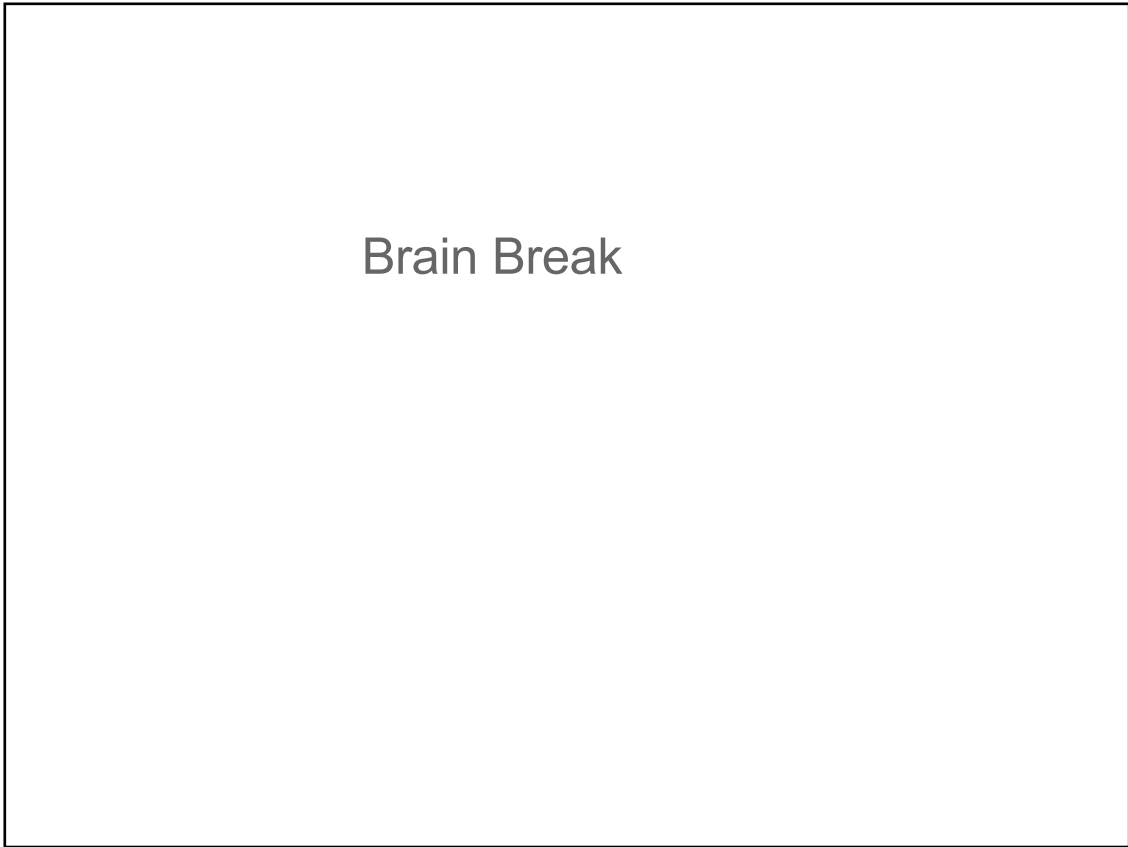
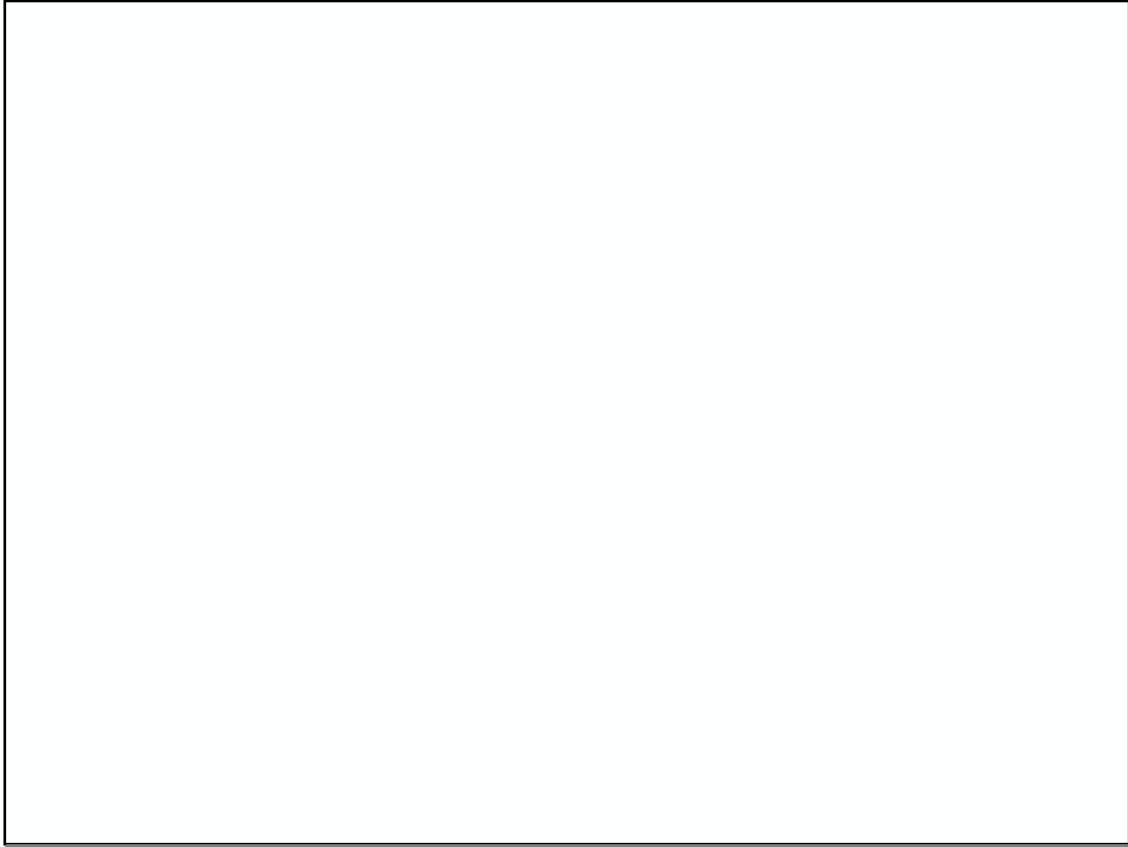
real  
Spreadsheet

2	Bazer, Madison R.			
3	Cervantes-Frank, Valenti	13	Melconian, Alexandra M.	
4	Conaghan, Whitney A.	14	Nashawi, Lynn	22
5	Duhaime, Hope	15	Pelayo, Miriam E.	23
6	Gittins, Benjamin T.	16	Roome, Calvin S.	24
7	Kennedy, John B. III	17	Sain, Margaret M.	25
8	Kinner, Nicole A.	18	Saunders, Morgan C.	26
9	Leach, Camille N.	19	Suryanata, Natania	27
10	Lugo, Isaiah P.	20	Thomas, Morgan A.	28
11	McKenzie, Caleb J.	21	Todayl, Andrew C.	29
12	McNair, Morgan A.			30
				Tort, Luis E.
				Ulm, Sophie E.
				Vasquez, Isaac R
				Villada-Youel, Ste
				Wagner, Ethan P.
				Wagner, Tristan A
				Warner, Liam G.
				Wood, Jacob
				Yeh, Nathan H.

A boy rolled a die 50 times with the following results:

Score	Frequency
1	9
2	10
3	5
4	8
5	7
6	11

Calculate Mean score from above data?



# Assignment

Study pp 170-177

Do.... p. 173....7, 8abc, 11

- page 174... Investigation #2 (1 to 4)  
page 178....1 and 2

<b>AP Statistics Tutorial</b>	<b>AP Statistics Tutorial</b>
<b>Exploring Data</b>	Welcome to Stat Trek's free, online Advanced Placement (AP) Statistics tutorial. It has been carefully developed to help you master the Advanced Placement Statistics Examination. > <a href="#">Begin lesson 1</a>
▶ <a href="#">The basics</a>	
▶ <a href="#">Charts and graphs</a>	
▶ <a href="#">Regression</a>	
▶ <a href="#">Categorical data</a>	
<b>Planning a Study</b>	<b>About the Tutorial</b>
▶ <a href="#">Surveys</a>	This tutorial provides accurate and complete coverage of the AP Statistics curriculum. Specifically, the AP Statistics curriculum and this tutorial cover the following topics:
▶ <a href="#">Experiments</a>	
<b>Anticipating Patterns</b>	<ul style="list-style-type: none"><li>▪ <b>Exploring data.</b> Using graphical and numerical techniques to study patterns of data. Emphasizes interpreting graphical information and descriptive statistics.</li><li>▪ <b>Sampling and experimentation.</b> How to plan and conduct a study. Focuses on clarifying</li></ul>

**AP Statistics Tutorial**

- Exploring Data
  - ▶ The basics
  - ▶ Charts and graphs
  - ▶ Regression
  - ▶ Categorical data
- Planning a Study**
  - ▶ Surveys
  - ▶ Experiments
- Anticipating Patterns
  - ▶ Probability

**Planning a Study**

- ▼ Surveys
  - Data collection
  - Sampling methods
  - Bias in surveys

