L4 Statistics

Concept 1: Measure of Center

A **measure of center** is a value that tries to describe a set of data using a single number, which is a typical value within the data set that represents its center point. Ideally, most numbers in the data set lie around the center. There are three measures of the "center" of the data: mean, median, and mode.

1. **Mean**: The "average". It can be found by adding all the numbers in the data set and then dividing by the number of values in the set.

For example, the height of four people is 1.5m, 1.6m, 1.7m, and 1.8m. The average height of the four people is (1.5+1,6+1,7+1,8)÷4 = 1.65m

2. **Median**: The middle value when the data items are listed in order.

For example, the height of five people is 1.5m, 1.6m, 1.7m, 1.8m, and 1.9m. The median height is 1.7m.

If there is an even number of values, the median is the mean of the two values in the middle.

For example, the height of four people is 1.5m, 1.6m, 1.7m, and 1.8m. The median height is (1.6+1.7)÷2 = 1.65m

3. **Mode**: The data value that appears most often. A set of data can have multiple modes.

For example, the height of five people is 1.5m, 1.5m, 1.7m, 1.9m, and 1.9m. The mode height is 1.5m and 1.9m

Math Exploration 1

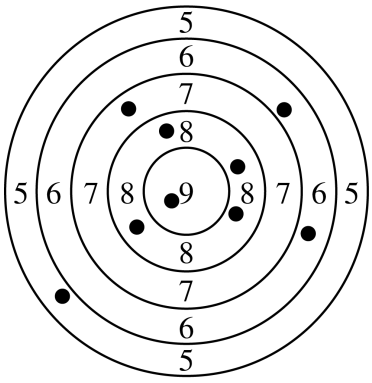
1 Which of the following statements is correct?

A. The average weight of the students in Andy's group is 36kg; the average weight of the students in Betty's group is 34kg. Andy must be heavier than Betty.

B. There are 43 students in Class 4 (1). The average score of the final math exam is 90 points. There must be no students failing in this class.

C. Lili's average score in the final examination of science, mathematics and Spanish is 91 points, and her total score is 273.

2 Jack made 9 shots in total in a shooting practice session. The result is shown in the picture below. What is the mean, median and mode of Jack's shooting result(round down to second digit after the decimal)?

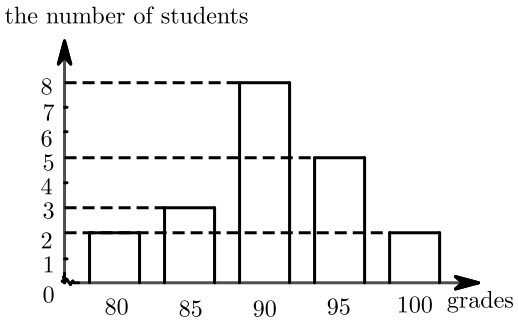


3 Amy took three tests last week, the mean score is 96, shown as the following below:

Subjects	Math	Physics	Chemistry
Score	93	98	

She got _____ score in chemistry test.

4 The following graph shows 20 students' math competition grades. Find the mode, median, and mean.



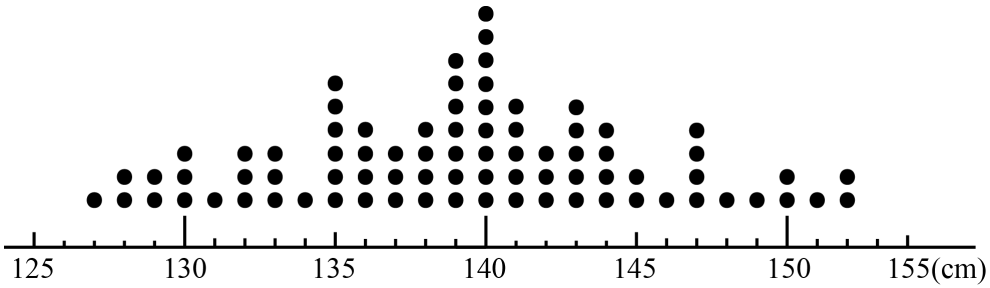
Mode: _____

Median: _____

Mean: _____

5 The monthly salary of the manager of a company is $5000; the monthly salary of two vice managers is $2000; and the average monthly salary of other 15 employees is $1200. The average monthly salary of the 18 people in this company is $ _____ ; The median of this set of data is _____ , and the mode is _____ .

6 The box plot below shows the height of the fourth grade students in a school.



As shown in the chart, among the fourth grade students in this school:

(1) The highest height is _____ cm, the shortest height is _____ cm.

(2) Most students are _____ cm tall, there are _____ of them.

Concept 2: Measure of Variation

A **measure of variation** is a value that tries to describe the spread in a data set. Ideally, the measure of variation tells you how far apart the numbers in the data set are from the center.

There are three typical measures of variations: range, mean absolute deviation, and interquartile range.

1. Range: How much a given data set is stretched out from smallest to largest. In a set of data, Range = largest number - smallest number.

For example, the height of four people is 1.5m, 1.6m, 1.7m, and 1.8m. The range of height is 1.8-1.5=0.3m

2. Mean Absolute Deviation(MAD): The average distance between each data point and the mean. In a set of data, MAD can be found by the following steps:

(1). Calculate the mean

For example, the height of four people is 1.5m, 1.6m, 1.7m, and 1.8m. The mean is 1.65m.

(2). Calculate the absolute distance between each data point and the mean

Data point	Distance from mean
10	$\|10-16\|=6$
15	$\|15-16\|=1$
15	$\|15-16\|=1$
17	$\|17-16\|=1$
18	$\|18-16\|=2$
21	$\|21-16\|=5$

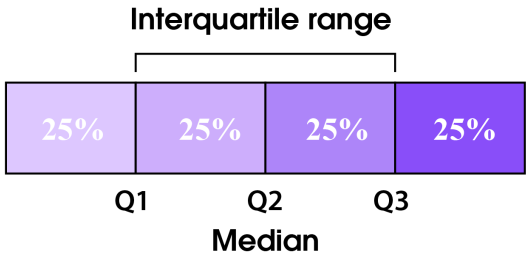
(3). Add all the distances together

(0.15+0.05+0.05+0.15)=0.4

(4). Divide the sum by the number of data points in the set.

0.4÷4=0.1

3. Interquartile Range(IQR): The spread of the middle half of the data set when all data points are ordered from low to high into four equal parts called **Quartiles**.



It can be found by the following steps:

(1). Find out the value of Q1 and Q3, which is the median of the first half and second half of the dataset. If there is an odd number of data points in the set, include the median of the data in both halves.

For example, the height of five people is 1.4m, 1.5m, 1.6m, 1.7m, and 1.8m.

The first half of the data set is: 1.4m, 1.5m, 1.6m(median); The second half of the data set is 1.6m(median), 1.7m, 1.8m.

Q1 = 1.5m; Q3 = 1.7m

(2). IQR = Q3 -Q1.

IQR = 1.7-1.5=0.2

Math Exploration 2

7 Find the range, lower quartile, and upper quartile of the data.

$6, 17, 26, 28, 30, 40, 13, 37, 33, 25$

(1) Range: _____

(2) Lower Quartile: _____

(3) Upper Quartile: _____

(4) IQR: _____

8 The box plot below shows the scores of a math test. Find the following values:

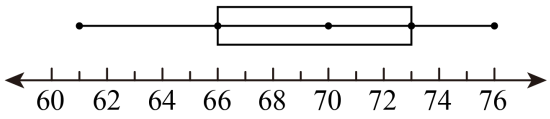
(1) Minimum: _____

(2) Maximum: _____

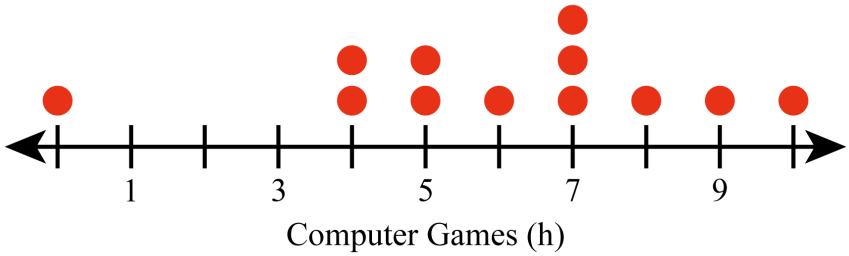
(3) Median: _____

(4) Range: _____

(5) Interquartile range: _____



9 Look at the dot plots of the number of hours that students play computer games each week.



(1) What is the mean of this set of data in dot plot?

(2) What is the MAD?

10 Find the MAD based on the data set below.

25	26	26	24	24	25

Concept 3: Population of Samples

When information is being gathered about a group, the entire group of objects, individuals, or events is called the **population**. A **sample** is part of the population that is chosen to represent the entire group.

A sample in which every person, object, or event has an equal chance of being selected is called a **random sample**.

When a sample does not accurately represent the population, it is called a **biased sample**.

Math Exploration 3

11 In order to understand the daily exercise of students in a middle school, which of the following data should be used?

A. Investigate the seventh grade students daily exercise.

B. Investigate the daily exercise of girls in this school.

C. Investigate the daily exercise of male students in this school.

D. Investigate 100 students each from seventh, eighth and ninth grades, in this school.

12 Which of the following methods of investigation fits with the goal?

A. To understand the time the middle school students spent on studying everyday in the city, investigate the entire population.

B. In order to ensure the successful launch of the carrier rocket, test some parts.

C. In order to know the diet of people who live in America, investigate the diet of the entire population in the area

D. Use sample to test the qualification of car accessories.

13 Which of the following methods of investigation fits with the goal?

A. In order to know the service life of a batch of tubes, test all tubes.

B. In order to understand the visual acuity of junior high school students in a city, use samples.

C. In order to understand the treatment of household garbage by residents in a state, investigate everyone who lives in that state.

D. In order to know the quality of rocket equipment parts, select sampling survey.

14 Which of the following methods of investigation is correct?

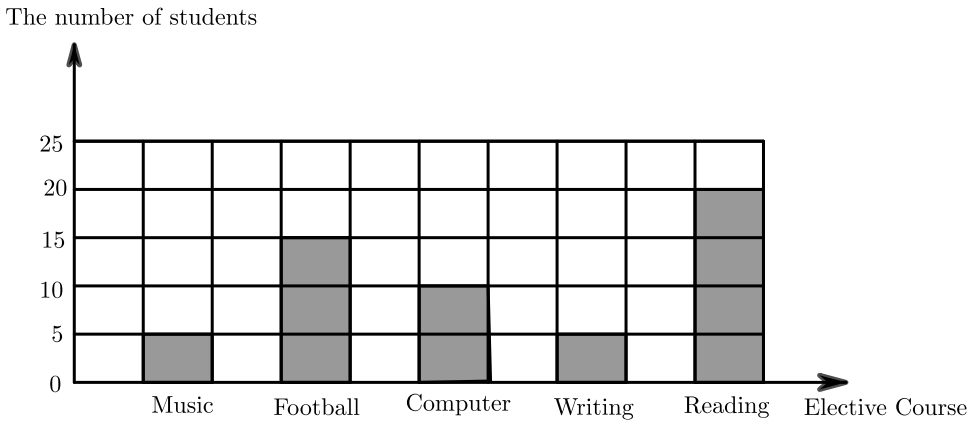
A. To know how many people visit amusement parks each year, researchers surveyed the number of people who attend during spring break.

B. In order to investigate the heights of students in a school, randomly select 50 school basketball team students to do the investigation.

C. In order to know the rating of a certain movie, select a single row number of the audience to conduct a survey.

D. In order to understand the weight of middle school students in A town, randomly select 100 students in a dessert shop to do the investigation.

15 A survey surveyed 55 students about which elective course they chose, as shown in the picture below.



(1) Complete the table.

Elective Course	Music	Football	Computer	Writing	Reading
Students					

(2) Which elective course is the most popular? _____

(3) _____ and _____ were choosen by the same number of students.

(4) Eli wasn't in school today, but he chose the elective course. Do you know what course he's most likely in? Why?

Homework

16 For the following statements of "Mean", which one is correct? ( ) .

A. The car's average speed is 60 kilometers per hour, so the car keep the speed of 60 kilometers per hour all the time.

B. The average weight of the school football players is $70\ \text{kg}$, some players will be heavier than $70\ \text{kg}$, some players will be lighter than 70kg.

17 Which measures can reflect the most frequent cases of a dataset?

A. Mean B. Median C. Mode

18 In the set $157, 170, 150, 156, 162, 157, 168, 155, 157, 158$, the median is _____ , and the mode is _____ .

19 George asked 10 classmates to record the number of hours they spent doing sports during one week. The table shows the data he collected.

10	11	22	17	7
17	20	19	12	31

Calculate the mean, median and mode.

20 Find the range, lower quartile, and upper quartile of the data.

$94, 34, 98, 78, 50, 85, 69, 74$

(1) Range: _____

(2) Lower Quartile: _____

(3) Upper Quartile: _____

(4) IQR: _____

21 Find the MAD based on the data set below.

101	98	102	100	99

22 In the following survey, which can not use the sample?

A. Investigate the product quality of important aircraft parts.

B. Research the ratings of a TV series.

C. Investigat the effectiveness of the vaccine.

D. Investigate how many trees are in a forest.

23 Which of the following investigation method is correct?

A. In order to find out audiences' opinion of the film, Sam randomly interviewed 8 Grade 3 students in his school.

B. In order to find out how much time students in the whole school spend on math homework, Emily surveyed 3 friend.

C. To find out how much teenagers across the country sleep, researches should investigate every teenagers in tha country.

D. In order to find out the quality of satellite parts, inspectors test every parts.

24 Of the following surveys, which one should inverstigate the entrie population?

A. Understand the usage lifespan of a batch of Ipad.

B. Understand the ratings of a TV program.

C. Know the test scores of a class of students.

D. Know the number of fish in a fish pond.

25 Which of the following investigation method is correct?

A. To find out how citizens feel about the film, Zoey randomly interviewed 8 Grade seven students at a school.

B. In order to find out how much time students in the whole school spend on reading every week, Andrew surveyed 20 his friends.

C. To find out how much teenagers and children across the country sleep, statisticians used a census.

D. In order to understand the quality of the cookies, inspectors ramdomly select 10 samples among every 100 population.