02. DESCRIPTIVE ANALYSIS

M5U3P2

WHAT 3 TYPES OF ANALYSIS DO WE HAVE?

3 TYPES OF ANALYSIS

- We have 3 types of data analysis:
  - Descriptive analysis
  - Predictive analysis
  - Prescriptive analysis

WHAT IS DESCRIPTIVE ANALYSIS?

DESCRIPTIVE ANALYSIS

- It's a fundamental analysis
- Gives an overview of data and the relationship between it
- Explains what happened in the past so it will help us make choices for the future

DESCRIPTIVE ANALYSIS

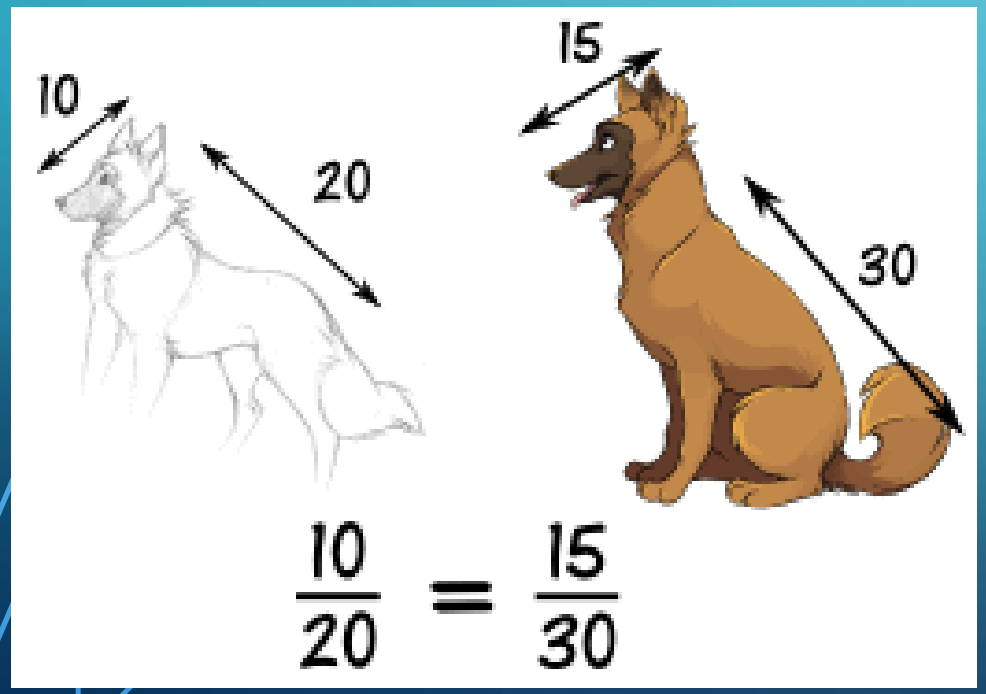
- Uses mathematical calculations and basic statistics such as proportions and percentages
- Measures the central value of the data (median) and distribution of information
- Finds a relationship between the data

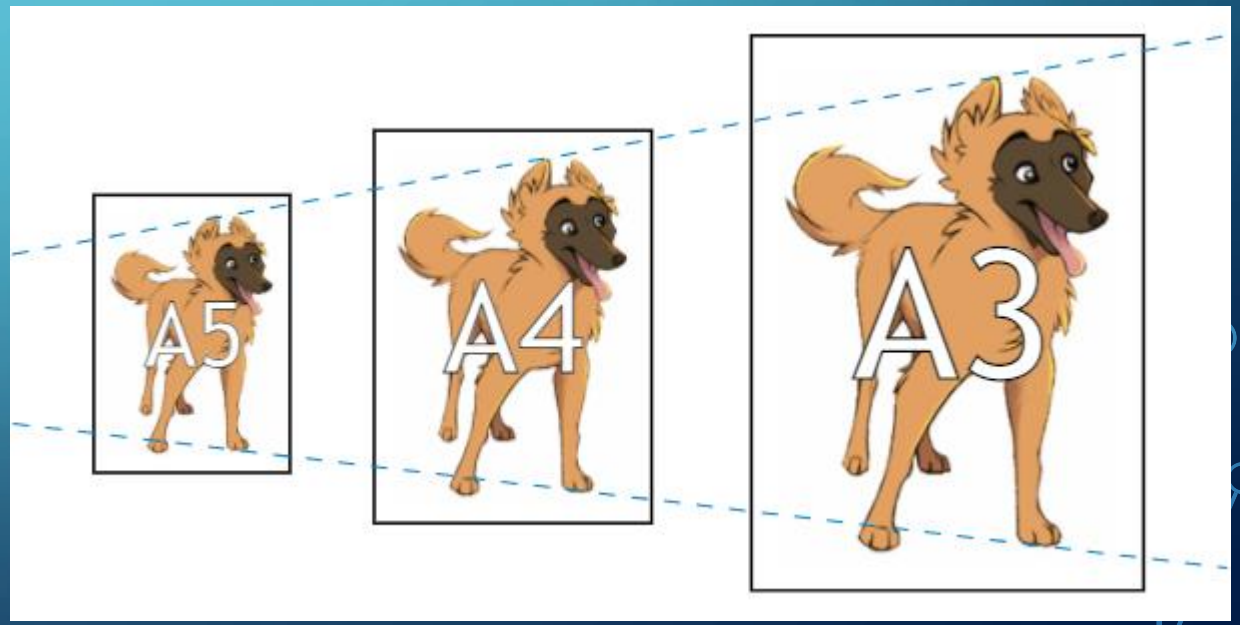
PROPORTION AND PERCENTAGE

- What is proportion (สัดส่วน)?

PROPORTION AND PERCENTAGE

- A proportion (สัดส่วน) is an equation that says that two or more ratios are equal



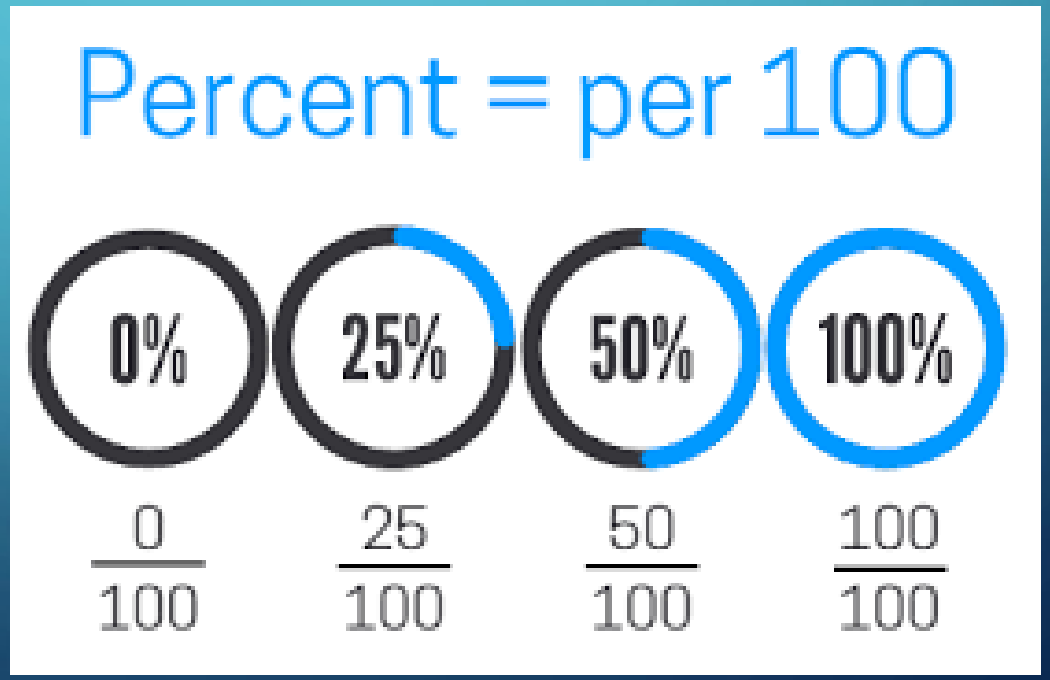


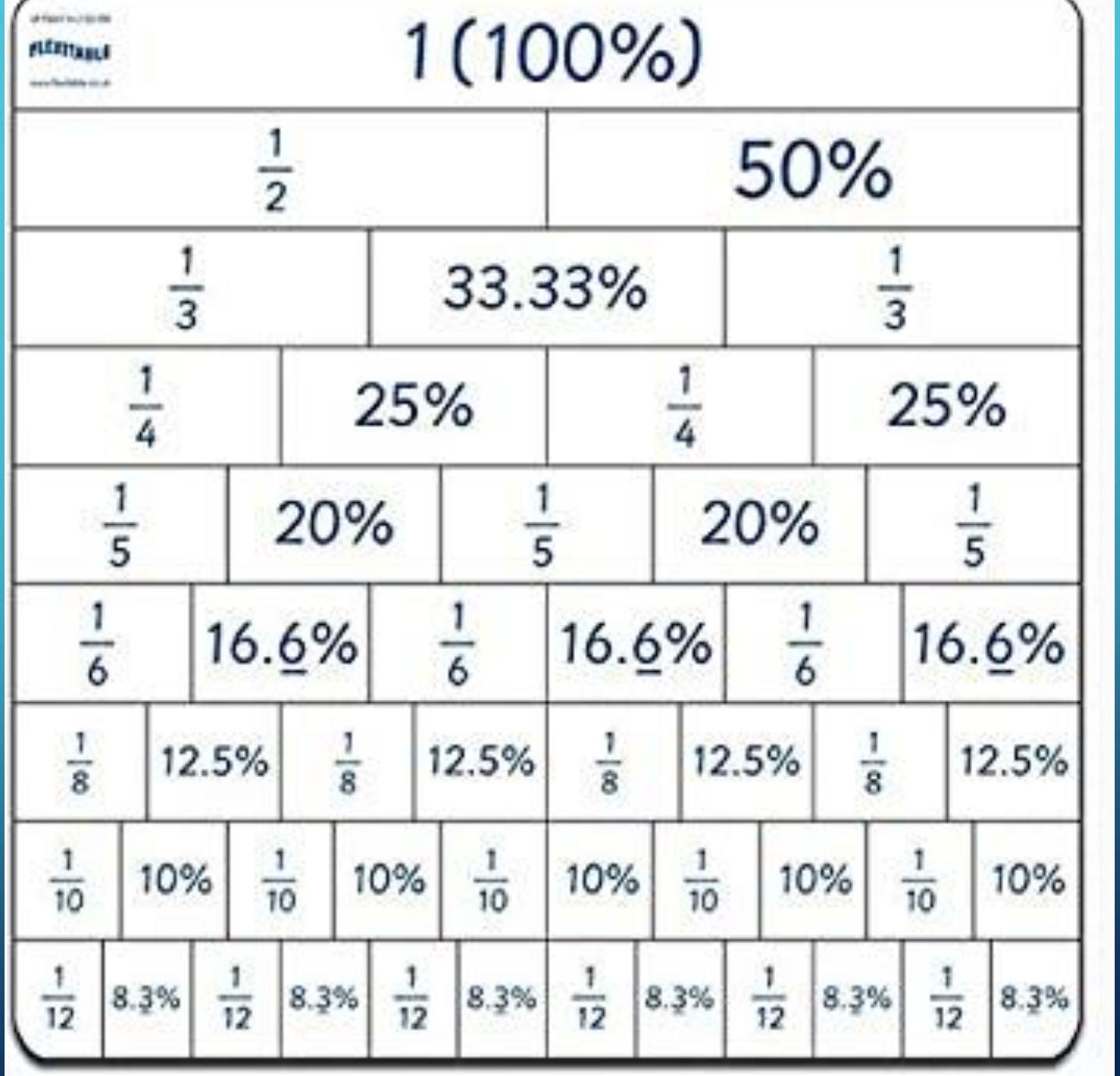
PROPORTION AND PERCENTAGE

What is percentage?

PROPORTION AND PERCENTAGE

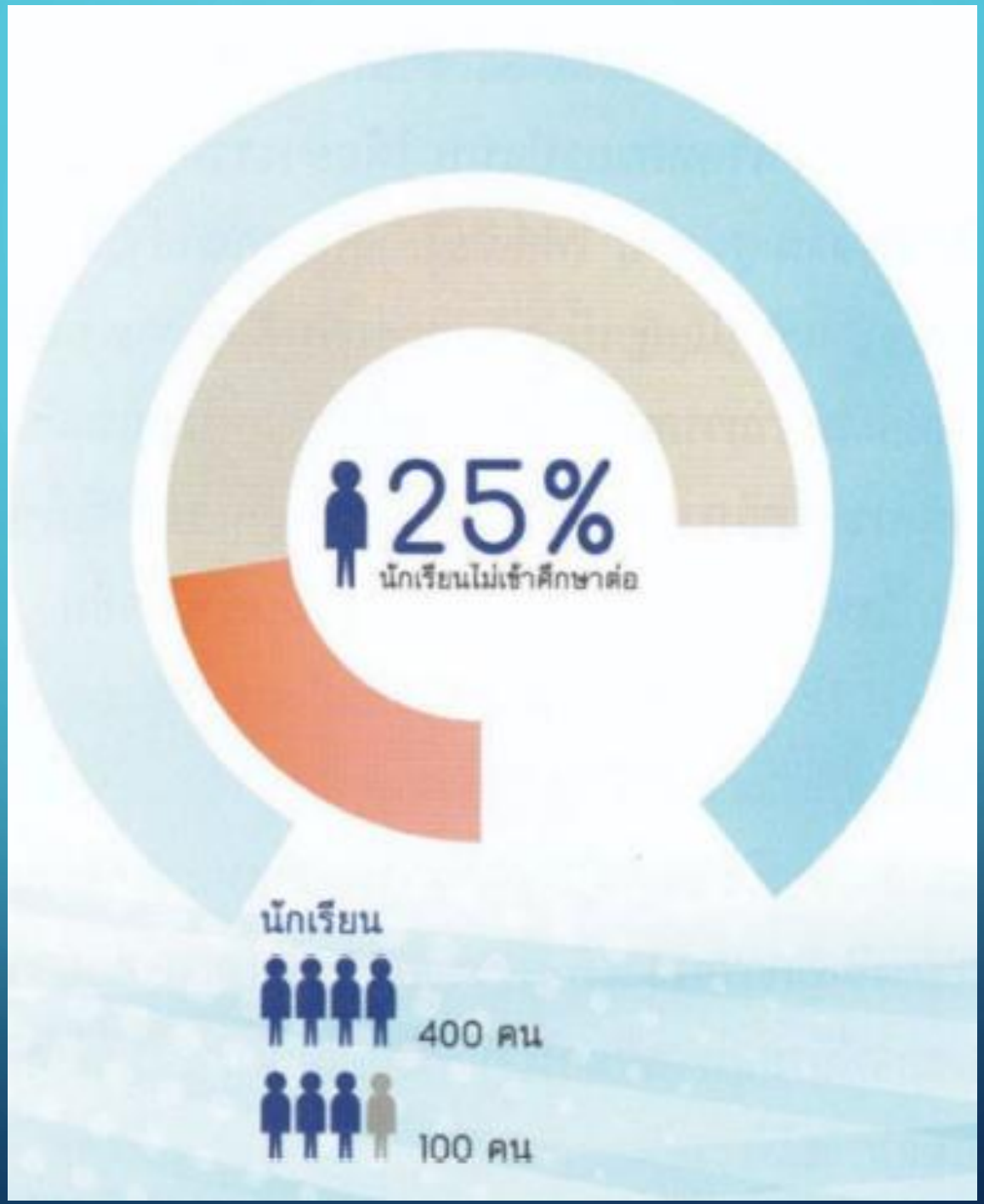
- A percentage is a number or ratio that represents a fraction of 100
- It has a % symbol




1 (100%)
1/2
50%
1/3
33.33%
1/3
1/4
25%
1/4
25%
1/5
20%
1/5
20%
1/5
1/6
16.6%
1/6
16.6%
1/6
16.6%
1/8
12.5%
1/8
12.5%
1/8
12.5%
1/8
12.5%
1/10
10%
1/10
10%
1/10
10%
1/10
10%
1/10
10%
1/12
8.3%
1/12
8.3%
1/12
8.3%
1/12
8.3%
1/12
8.3%
1/12
8.3%

PROPORTION AND PERCENTAGE

- Proportions and percentages make it easier to understand the data because it can be represented with a diagram or a chart
- For example:
  - In a high school survey it was found out that 100 out of 400 students will not continue their further studies
  - How much percent is that?


25%
นักเรียนไม่เข้าศึกษาต่อ
นักเรียน
400 คน
100 คน

WHAT IS CENTRAL/MIDDLE VALUE OF DATA?

CENTRAL/MIDDLE VALUE OF DATA (MEDIAN)

- A median is a value separating the higher half from the lower half of some data
- The median is not the same as average
- Sometimes the median can show the exact middle in some data

EXAMPLE 1: WHAT IS THE AVERAGE SALARY?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

EXAMPLE 1: WHAT IS THE AVERAGE SALARY?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / number of people = average

EXAMPLE 1: WHAT IS THE AVERAGE SALARY?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / number of people = average
(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / 7 = average

EXAMPLE 1: WHAT IS THE AVERAGE SALARY?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / number of people = average
(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / 7 = average
169,700/ 7 = average

EXAMPLE 1: WHAT IS THE AVERAGE SALARY?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / number of people = average
(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / 7 = average
169,700/ 7 = average
169,700/ 7 **= 24,242.85**

EXAMPLE 1: HOW MANY PEOPLE EARN SAME OR MORE THAN AVERAGE SALARY?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / number of people = average
(9,800 + 15,700 + 75,600 + 11,200 + 23,000 + 24,600 + 9,800) / 7 = average
169,700/ 7 = average
169,700/ 7 **= 24,242.85**

EXAMPLE 1: HOW MANY PEOPLE EARN SAME OR MORE THAN AVERAGE SALARY?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

24,242.85 is the average value but it is not the middle value

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

Arrange the numbers from the smallest to the largest

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

Arrange the numbers from the smallest to the largest

Person A						
9,800						

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

Arrange the numbers from the smallest to the largest

Person A	Person G					
9,800	9,800					

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

Arrange the numbers from the smallest to the largest

Person A	Person G	Person D				
9,800	9,800	11,200				

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

Arrange the numbers from the smallest to the largest

Person A	Person G	Person D	Person B			
9,800	9,800	11,200	15,700			

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

Arrange the numbers from the smallest to the largest

Person A	Person G	Person D	Person B	Person E		
9,800	9,800	11,200	15,700	23,000		

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person B	Person C	Person D	Person E	Person F	Person G
9,800	15,700	75,600	11,200	23,000	24,600	9,800

Arrange the numbers from the smallest to the largest

Person A	Person G	Person D	Person B	Person E	Person F	
9,800	9,800	11,200	15,700	23,000	24,600	

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person G	Person D	Person B	Person E	Person F	Person C
9,800	9,800	11,200	15,700	23,000	24,600	75,600

Arrange the numbers from the smallest to the largest
Now we count how many data sets do we have
If we have an odd number we divide it by 2 and round it to a higher number

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person G	Person D	Person B	Person E	Person F	Person C
9,800	9,800	11,200	15,700	23,000	24,600	75,600

Arrange the numbers from the smallest to the largest
Now we count how many data sets do we have
If we have an odd number we divide it by 2 and round it to a higher number

7/2 = 3.5 = 4

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person G	Person D	Person B	Person E	Person F	Person G
9,800	9,800	11,200	15,700	23,000	24,600	75,600

Arrange the numbers from the smallest to the largest
Now we count how many data sets do we have
If we have an odd number we divide it by 2 and round it to a higher number

7/2 = 3.5 = 4

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person G	Person D	Person B	Person E	Person F	Person H	Person C
9,800	9,800	11,200	15,700	23,000	24,600	25,600	75,600

Arrange the numbers from the smallest to the largest
Now we count how many data sets do we have
If we have an even number we divide it by 2 and take that number and the next number

8/2 = 4
4 + 1 = 5

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person G	Person D	Person B	Person E	Person F	Person H	Person G
9,800	9,800	11,200	15,700	23,000	24,600	25,600	75,600

Arrange the numbers from the smallest to the largest
Now we count how many data sets do we have
If we have an even number we divide it by 2 and take that number and the next number
After that we take both numbers and add them together and then divide to get the average value

15,700 + 23,000

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person G	Person D	Person B	Person E	Person F	Person H	Person C
9,800	9,800	11,200	15,700	23,000	24,600	25,600	75,600

Arrange the numbers from the smallest to the largest
Now we count how many data sets do we have
If we have an even number we divide it by 2 and take that number and the next number
After that we take both numbers and add them together and then divide to get the average value

15,700 + 23,000 = 38,700 / 2 = ??

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person G	Person D	Person B	Person E	Person F	Person H	Person C
9,800	9,800	11,200	15,700	23,000	24,600	25,600	75,600

Arrange the numbers from the smallest to the largest
Now we count how many data sets do we have
If we have an even number we divide it by 2 and take that number and the next number
After that we take both numbers and add them together and then divide to get the average value

15,700 + 23,000 = 38,700 / 2 = 19,350

EXAMPLE 1: HOW TO FIND THE MIDDLE VALUE (MEDIAN)?

Person A	Person G	Person D	Person B	Person E	Person F	Person H	Person C
9,800	9,800	11,200	15,700	23,000	24,600	25,600	75,600

Arrange the numbers from the smallest to the largest
Now we count how many data sets do we have
If we have an even number we divide it by 2 and take that number and the next number
After that we take both numbers and add them together and then divide to get the average value

15,700 + 23,000 = 38,700 / 2 = 19,350 → Median
((9,800)*2 + 11,200 + 15,700 + 23,000 + 24,600 + 25,600 + 75,600)/8 = 24,412.5 → Average

HOW TO KNOW WHEN TO USE MEDIAN AND WHEN TO USE AVERAGE?

HOW TO KNOW WHEN TO USE MEDIAN AND WHEN TO USE AVERAGE?

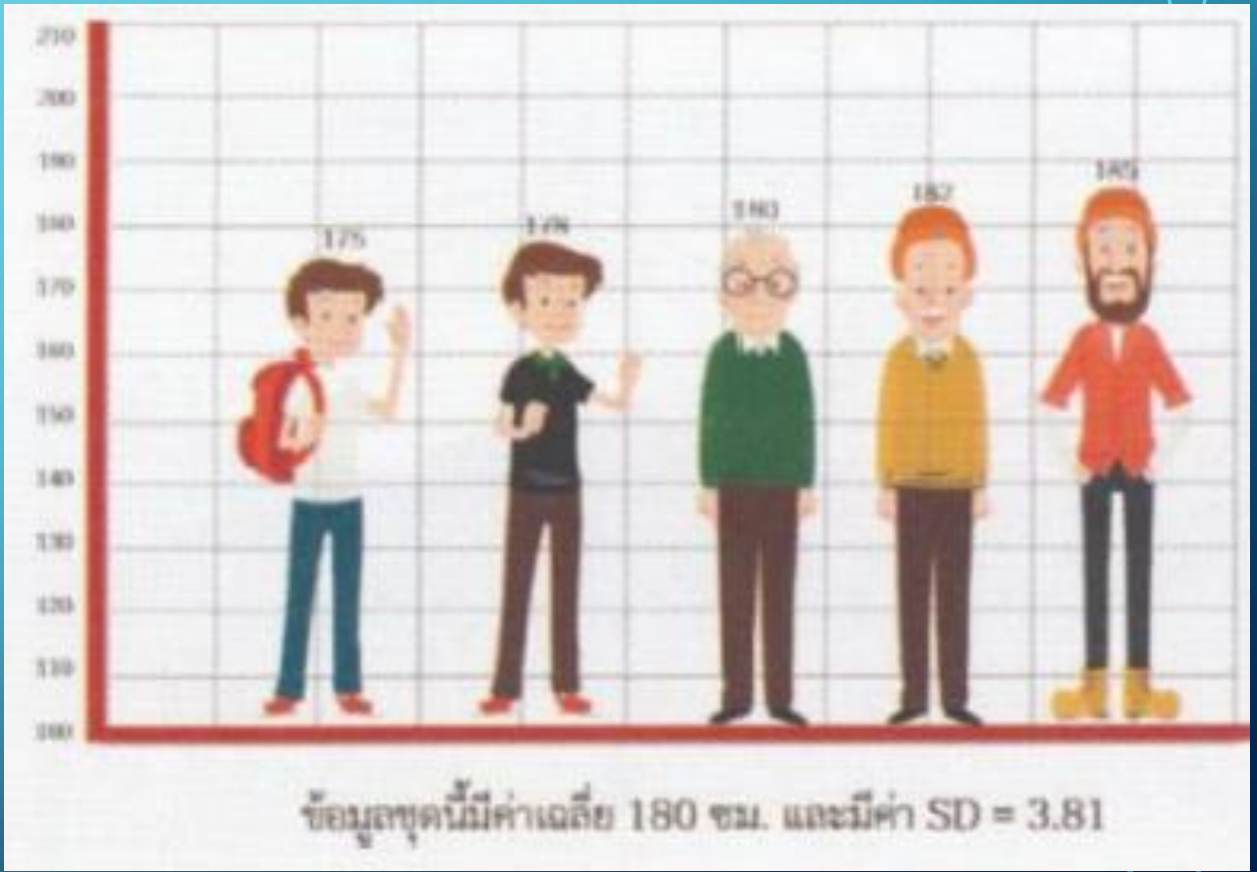
- Median is mostly used for a small spread of information
- Average is used when the spread of information is large
- The standard deviation is used to measure the spread of information

WHAT IS STANDARD DEVIATION?

STANDARD DEVIATION

- It is used to measure the distribution of the dataset
- It is always positive
- If the deviation number is low, that means that the dataset values are close to each other
- If the deviation number is high, that means that the dataset values are spread over a wider range

ข้อมูลชุดนี้มีค่าเฉลี่ย 180 ซม. และมีค่า SD = 22.36



ข้อมูลชุดนี้มีค่าเฉลี่ย 180 ซม. และมีค่า SD = 3.81

WILD CATTLE AND ZEBRA STATISTICS

Year	Wild Cattle (k.)	Wild Zebra (k.)		Year	Wild Cattle (k.)	Wild Zebra (k.)
1.	118.00	124.00		12.	75.84	98.07
2.	106.00	105.00		13.	80.62	128.30
3.	79.31	75.23		14.	86.34	118.22
4.	67.45	78.50		15.	94.06	125.20
5.	57.83	76.16		16.	99.15	130.97
6.	50.66	76.49		17.	96.50	124.31
7.	51.41	83.16		18.	95.47	130.08
8.	55.02	85.11		19.	98.07	121.76
9.	58.16	88.77		20.	97.88	132.40
10.	65.12	108.34		21.	105.74	120.60
11.	81.27	116.03		22.	106.58	135.77

ANALYSIS

- After the foxes have collected the information from last 23 years they have analyzed it

Analysis	Wild Cattle	Zebra
Minimum value	50.66	75.23
Highest value	118.00	135.77
Average	84.30	109.24
Standard deviation	20.83	21.48

ANALYSIS

- The foxes came to a conclusion that:
  - the number of zebras is greater than the number of wild cattle
  - Their standard deviation is very similar

WHAT IF THE NUMBER OF ZEBRAS WAS TWICE LOWER THAN THE NUMBER OF WILD CATTLE?

ANY QUESTIONS?

THE END